

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA E GESTÃO DO
CONHECIMENTO**

FLÁVIO MARCELO RISUENHO DOS SANTOS

**AVALIAÇÃO DA QUALIDADE DA INFORMAÇÃO DOS CONTROLES DE
EVENTOS DE FALHA E MANUTENÇÃO DE EQUIPAMENTOS INDUSTRIAIS**

Florianópolis
2009

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA E GESTÃO DO
CONHECIMENTO**

FLÁVIO MARCELO RISUENHO DOS SANTOS

**AVALIAÇÃO DA QUALIDADE DA INFORMAÇÃO DOS CONTROLES DE
EVENTOS DE FALHA E MANUTENÇÃO DE EQUIPAMENTOS INDUSTRIAIS**

Dissertação submetida à Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Engenharia e Gestão do Conhecimento.

Orientador:

LUIZ FERNANDO JACINTHO MAIA

Florianópolis

2009

Catálogo na fonte pela Biblioteca Universitária da
Universidade Federal de Santa Catarina

S237a Santos, Flávio Marcelo Risuenho dos
Avaliação da qualidade da informação dos controles
de eventos de falha e manutenção de equipamentos
industriais [dissertação] / Flávio Marcelo Risuenho dos
Santos ; orientador, Luis Fernando Jacintho Maia. -
Florianópolis, SC, 2009.
175 f.: il.

Dissertação (mestrado) - Universidade Federal de Santa
Catarina, Centro Tecnológico. Programa de Pós-graduação
em Engenharia e Gestão do Conhecimento.

Inclui referências

1. Engenharia e gestão do conhecimento. 2. Equipamento
industrial - Confiabilidade (Engenharia). 3. Indicador de
qualidade. 4. Qualidade da informação. 5. CommonKads. I.
Maia, Luiz Fernando Jacintho. II. Universidade Federal de
Santa Catarina. Programa de Pós-Graduação em Engenharia e
Gestão do Conhecimento. III. Título.

CDU 659.2

FLÁVIO MARCELO RISUENHO DOS SANTOS

**AVALIAÇÃO DA QUALIDADE DA INFORMAÇÃO DOS CONTROLES DE
EVENTOS DE FALHA E MANUTENÇÃO DE EQUIPAMENTOS INDUSTRIAIS**

Esta Dissertação foi julgada adequada para obtenção do Título de "Mestre em Engenharia e Gestão do Conhecimento", Especialidade em Engenharia do Conhecimento e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento.

Florianópolis, 24 de agosto de 2009.

Prof. Roberto Carlos dos Santos Pacheco, Dr.
Coordenador do Curso (EGC)

Banca Examinadora:

Prof. Luiz Fernando Jacintho Maia, Dr.
Orientador
UFSC

Prof. Francisco Antonio Pereira Fialho, Dr.
Membro da Banca
UFSC

Prof. Manoel Rosa de Oliveira Lino, Dr.
Membro da Banca
UFSC

Prof^ª. Raquel Borba Balceiro, Dra.
Membro da Banca
Petrobras

Aos meus Pais, que sempre me incentivaram
a trilhar o caminho do conhecimento.

AGRADECIMENTOS

A Deus, por ter provido os meios necessários para que eu pudesse concluir essa jornada.

À minha esposa, pela paciência para suportar meu espírito inquieto e pelo apoio para que eu pudesse realizar este trabalho.

À minha família, por sempre renovar minhas forças quando sobrevinha o desânimo.

Ao meu pai, pelas sugestões para aprimorar a qualidade do meu texto.

À Petrobras, em especial ao Paulo Roberto Viana, por ter me concedido a oportunidade e o apoio necessários para cursar o mestrado no EGC.

Ao meu orientador, sem o qual não seria possível minha atuação junto ao EGC.

Ao Prof. Dr. Richard Perassi Luiz de Sousa, pelo intenso debate sobre a natureza do conhecimento tratado pela Engenharia do Conhecimento.

A todos os professores do EGC, cujas contribuições foram imprescindíveis para a realização deste trabalho.

A toda a comunidade do EGC, pela intensa troca de idéias e pontos de vista que permitiram enriquecer este trabalho.

A todos os demais que, direta ou indiretamente, contribuíram para a realização desta dissertação.

RESUMO

Este trabalho propõe a construção e uso de um indicador para monitorar e controlar a qualidade das informações que podem ser obtidas a partir dos controles de eventos de falha e manutenção em equipamentos da indústria petrolífera. Para isso, através de pesquisa bibliográfica, identifica: os cenários das áreas de confiabilidade e manutenção industrial, os principais usos das informações desses registros em estudos de confiabilidade e os atributos de qualidade da informação pertinentes ao processo de coleta e acompanhamento dessas informações. A tarefa de monitoração é então adaptada para esse contexto, gerando um indicador evolutivo para a qualidade de coleta capaz de representar a situação da qualidade dos registros, o qual é elaborado com base no conhecimento de especialistas desse processo. Este indicador traduz, em uma visão gerencial, a situação da qualidade dos controles de eventos de falha e manutenção, tornando possível o acompanhamento periódico da qualidade das informações desse processo e, por consequência, diminuindo a incerteza das informações geradas a partir de seus dados e melhorando a qualidade potencial das decisões tomadas com base nessas informações.

Palavras-chave: Indicador de Qualidade. Qualidade da Informação. Manutenção Industrial. Confiabilidade. Representação de Conhecimento. CommonKads.

ABSTRACT

This work proposes the design and use of an indicator to monitor and control the information quality that can be obtained from the controls of failure and maintenance events in oil industry equipments. Through bibliographical research, this work identifies: the scene of the reliability and industrial maintenance areas; the main uses of the recorded information in reliability studies; and the quality attributes of the pertinent information to the collection and monitoring process of that information. The monitoring task is then adapted for that context, generating an evolutionary indicator for the collection quality that is capable of representing the registrations quality status and which is elaborated through the knowledge of process' specialists. This indicator translates, in a managerial vision, the quality status of failure and maintenance events controls, allowing the periodic evaluation of the process information quality and, because of this, decreasing the uncertainty about the information generated from this data and improving the potential quality of the decisions taken based on that information.

Keyword: Quality Indicator. Information Quality. Industrial Maintenance. Reliability. Knowledge Representation. CommonKads.

LISTA DE FIGURAS

Figura 2.1 – Evolução da Manutenção e do desenvolvimento de Bancos de Dados de Confiabilidade durante o século XX	22
Figura 2.2 – Fronteira para Motores Elétricos	24
Figura 2.3 – Diagrama esquemático de um sistema de comunicação	48
Figura 2.4 – Conjunto de modelos do CommonKADS	60
Figura 2.5 – Tipologia das tarefas intensivas em conhecimento	63
Figura 2.6 – Principais assuntos para a qualidade de dados	68
Figura 2.7 – Modelo de maturidade da qualidade de dados do Gartner	70
Figura 2.8 – Modelo de um sistema de gestão da qualidade baseado em processo	73
Figura 2.9 – Relacionamento entre os fatores de qualidade e as decisões de projeto de um <i>data warehouse</i>	82
Figura 2.10 – Componentes da qualidade da informação	89
Figura 2.11 – Framework de qualidade da informação	93
Figura 2.12 – Subconjunto de dados com acurácia	95
Figura 2.13 – Etapas de <i>data profiling</i>	95
Figura 3.1 – Eventos que demandam a execução de um serviço de manutenção	114
Figura 3.2 – Especificação de método para o método guiado por dados para monitoramento	155
Figura 3.3 – Estrutura de inferência para o esquema de tarefa de monitoramento	156

LISTA DE QUADROS

Quadro 2.1 – História Parcial da Coleção de Dados de Confiabilidade	17
Quadro 2.2 – Redução do plano de manutenção em relação ao requerido pelo DC-8	19
Quadro 2.3 – Classificação por Tipo para Motores Elétricos	25
Quadro 2.4 – Subdivisão de Motores Elétricos	25
Quadro 2.5 – Características técnicas de motores elétricos	26
Quadro 2.6 – Exemplo de dados de Local de Instalação (Tag)	27
Quadro 2.7 – Exemplo de dados de equipamento	28
Quadro 2.8 – Atividades de manutenção	29
Quadro 2.9 – Dados de manutenção a serem registrados	31
Quadro 2.10 – Métodos de detecção de falhas	32
Quadro 2.11 – Modos de falhas de motores elétricos	33
Quadro 2.12– Mecanismo de falhas	35
Quadro 2.13– Causa da falha	37
Quadro 2.14 – Parâmetros de confiabilidade e manutenção para os níveis da taxonomia	38
Quadro 2.15 – Dados solicitados para o registro de uma falha	39
Quadro 2.16 – Dimensões da qualidade para valores de dados	77
Quadro 2.17 – Relações hierárquicas multidimensionais dos conceitos utilizados para avaliação da qualidade dos dados	78
Quadro 2.18 – Categorias e Dimensões da qualidade dos dados	80
Quadro 2.19 – Perspectiva, Dimensões e Critérios da Qualidade dos Dados propostos por Redman	81
Quadro 2.20 – Dimensões e critérios de qualidade da informação por Königer e Reithmayer	83
Quadro 2.21 – Dimensões da qualidade da Informação no modelo PSP/IQ (1998)	84
Quadro 2.22 – Dimensões e critérios de qualidade da informação na metodologia TIQM ...	85
Quadro 2.23 – Dimensões de qualidade da informação utilizadas com a técnica GQM	86
Quadro 2.24 – Critérios de qualidade da informação para <i>Web</i>	86
Quadro 2.25 – Critérios de qualidade da informação	87
Quadro 2.26 – Dimensões de qualidade da informação por Loshin	88
Quadro 2.27 – Dimensões da qualidade da Informação no modelo PSP/IQ (2002)	90
Quadro 2.28 – Os 16 critérios de qualidade da informação e seus opostos por Eppler	92

LISTA DE QUADROS

Quadro 2.29 – Atributos de qualidade da informação do <i>framework</i> NCW	98
Quadro 2.30 – Exemplo de taxonomia orientada a resultados de atributos de qualidade de dados ou informação em uma seqüência econômica de avaliação	99
Quadro 2.31 – Principais diferenças entre fabricação de produtos e produção de informações	104
Quadro 3.1 – Atributos relativos à qualidade de coleta segundo Liu e Chi (2002)	119
Quadro 3.2 – Matriz quantitativa de conseqüência para os registros de falha	122
Quadro 3.3 – Matriz de relevância dos registros de falha	123
Quadro 3.4 – Matriz de relevância dos registros de manutenção	123
Quadro 3.5 – Relevância do Equipamento	124
Quadro 3.6 – Matriz de Relevância dos registros de falha	125
Quadro 3.7 – Matriz de Relevância dos registros de manutenção	126
Quadro 3.8 – Ficha do atributo Relevância	126
Quadro 3.9 – Ficha do atributo Controle de Duplicação	128
Quadro 3.10 – Ficha do atributo Completeza	131
Quadro 3.11 – Ficha do atributo Atualidade	135
Quadro 3.12 – Ficha do atributo Acurácia	144
Quadro 3.13 – Ficha do atributo Reputação	148
Quadro 3.14 – Ficha do Indicador de Qualidade da Informação – Coleta de dados	152

SUMÁRIO

RESUMO	5
ABSTRACT	6
LISTA DE FIGURAS	7
LISTA DE QUADROS	8
1 INTRODUÇÃO	12
1.1 Apresentação do Problema de Pesquisa	12
1.2 Objetivo Geral do trabalho	13
1.3 Objetivos Específicos do trabalho	13
1.4 Escopo do Trabalho	14
1.5 Aderência ao Programa de Pós-Graduação do EGC	14
2 FUNDAMENTAÇÃO TEÓRICA	15
2.1 Históricos das áreas de Manutenção e Bancos de Dados de Confiabilidade	15
2.2 Registros de Manutenção e Confiabilidade Conforme a Norma ISO 14224	23
2.3 A Informação e o Conhecimento no Contexto da Gestão e Engenharia do Conhecimento	41
2.3.1 Abordagem hierárquica DIKW	42
2.3.2 Deficiências da DIKW	43
2.3.3 Revisão Epistemológica	44
2.3.4 Implicações Relevantes	51
2.3.5 O Conhecimento do “Nível do Conhecimento”	55
2.3.6 O Conhecimento dos “Sistemas Baseados em Conhecimento”	57
2.3.7 Diferenciando as áreas de Informação e Conhecimento	58
2.4 A Modelagem de Conhecimento no CommonKADS	60
2.4.1 Tarefas Intensivas de Conhecimento	63
2.5 Qualidade da informação	66
2.5.1 Definições para a qualidade da informação	70
2.5.2 Conferências sobre qualidade de dados ou de informação	71
2.5.3 Gestão da qualidade na NBR ISO 9000	72
2.5.4 Principais abordagens para o tratamento da qualidade da informação	74
2.5.5 Algumas lições sobre medição de qualidade da informação	109
3 CONSTRUÇÃO EVOLUTIVA DE UM INDICADOR DE QUALIDADE DE INFORMAÇÃO APLICADO NA COLETA DE DADOS DE INTERESSE PARA BANCOS DE DADOS DE CONFIABILIDADE E MANUTENABILIDADE	111
3.1 Indicadores	111
3.2 Entendendo o ambiente de Manutenção	112
3.3 Delegando propriedade e responsabilidade sobre os dados	115
3.4 Traduzindo a percepção de qualidade em um conjunto de atributos	116
3.5 Mapeando variáveis para os atributos	121
3.5.1 Relevância	121
3.5.2 Controle de Duplicação	128
3.5.3 Completeza	130
3.5.4 Atualidade	134
3.5.5 Acurácia	141

3.5.6 Reputação	146
3.6 Primeira versão do indicador de qualidade da informação da fase de coleta	151
3.7 Ajustando a métrica de avaliação à percepção de qualidade da empresa	153
3.8 Conduzindo futuras evoluções baseadas em conhecimento	154
4 CONCLUSÃO	158
4.1 Conclusões	158
4.2 Sugestões para Trabalhos Futuros	160
5 REFERÊNCIAS	162

1 - INTRODUÇÃO

1.1 - Apresentação do Problema de Pesquisa:

No último século a evolução dos produtos e meios de produção passou a demandar uma maior disponibilidade, confiabilidade e manutenibilidade. Para atingir esses objetivos, era necessário acompanhar o desempenho das máquinas e produtos tanto em testes de laboratório quanto no campo.

A partir da análise dos dados obtidos houve a evolução tanto da atividade de manutenção, que busca manter a função do bem, quanto do próprio processo de coleta. Hoje a atividade de manutenção conta com várias técnicas de análise, para as quais é fundamental a existência de dados históricos que possam caracterizar o desempenho dos bens mantidos.

Entretanto, o processo de coleta continua sendo um problema. Uma solução alternativa que pode ser empregada consiste no uso de dados divulgados por outras empresas. Para auxiliar esse processo de troca de informações entre empresas da área de petróleo foi publicada em 1999 a norma ISO 14224, que teve uma segunda versão liberada em 2006.

Essa norma especifica vários dados que devem ser registrados durante o processo de manutenção, definindo inclusive alguns critérios para a sua estruturação. Além disso, a norma destaca em seu item 7 a necessidade do acompanhamento da qualidade dos dados a serem coletados.

Essa questão do acompanhamento da qualidade de dados e da informação tem sido bastante abordada em diversos trabalhos, publicados principalmente a partir da década de 1990, nos quais normalmente é destacada a necessidade de medição da qualidade para o gerenciamento da informação (PAUTKE e REDMAN, 1990; REDMAN, 1992; WANG e KON, 1993; FOX *et al.*, 1994; WANG *et al.*, 1995b; REDMAN, 1996; ENGLISH, 1999; LOSHIN, 2001; REDMAN, 2001; LIU e CHI, 2002; OLSON, 2003; EPPLER, 2006). Entretanto, as características ou dimensões que definem o que é qualidade e, portanto, como esta pode ser aferida, variam conforme o processo analisado. Talvez por isso, a construção de indicadores sempre foi deixada a cargo de cada organização que deseja aferir a qualidade de seus dados.

A avaliação da qualidade da informação presente nos registros de manutenção usualmente é feita apenas de modo pontual, como parte de um estudo de confiabilidade ou manutenibilidade realizado para sanar alguma dificuldade já manifesta. Essa avaliação, que

demanda um conhecimento tanto do comportamento esperado do equipamento quanto do observado, é uma parte essencial do estudo e, junto com o processo de coleta e preparação dos dados, absorve normalmente a maior parte do tempo empregado (ZHANG *et al.*, 2005; ZHU e WU, 2005).

A preparação de um modelo de dados apropriado em conjunto com a avaliação periódica da qualidade das informações registradas no sistema de gestão da manutenção poderia reduzir significativamente este tempo mas, dado o volume de dados que deveria ser analisado e o perfil requerido do profissional, os custos de manter uma análise manual como atividade rotineira são normalmente proibitivos. Por isso, existe a necessidade de uma forma automatizada de verificação da qualidade.

Para preencher essa lacuna, é preciso identificar, em função dos objetivos previstos para os dados, quais são requisitos de qualidade que devem ser atendidos, e como eles podem ser avaliados de forma sistemática.

Essa avaliação da qualidade demanda um grande conhecimento do processo industrial monitorado e seus equipamentos, do processo de gestão implementado, do sistema de informações em uso e dos objetivos dos dados. Assim, este trabalho visa responder a seguinte questão:

“Como resguardar o conhecimento utilizado para avaliar a qualidade das informações que podem ser obtidas a partir dos registros da atividade de manutenção?”

1.2 - Objetivo Geral do trabalho

Elaborar uma estrutura que permita resguardar o conhecimento utilizado para a avaliação da qualidade das informações obtidas a partir de dados de manutenção de equipamentos.

1.3 - Objetivos Específicos do trabalho

- Estabelecer as informações que deverão ser obtidas dos dados de manutenção de equipamentos industriais da área de petróleo.
- Identificar quais registros são essenciais para essas informações.
- Identificar os conhecimentos utilizados para avaliação da qualidade das informações obtidas por intermédio destes registros.

- Propor uma estrutura que permita utilizar esse conhecimento para avaliação da qualidade das informações obtidas por intermédio destes registros.

1.4 - Escopo do Trabalho

Dado a ampla gama de fatores que influem tanto na definição das informações que devem ser registradas quanto na avaliação de sua qualidade, este trabalho irá discutir em detalhes apenas a avaliação da qualidade da informação dos registros de manutenção de um motor elétrico, conforme os requisitos especificados na norma ISO 14224 (2006).

O gerenciamento da qualidade de informação, conforme destacado em vários trabalhos, demanda a implementação de uma abordagem sistêmica que promova principalmente uma mudança de cultura e valores na organização onde for implantado. Entretanto, este trabalho não irá considerar qualquer outro aspecto envolvido nesse processo além da medição da qualidade da informação relativa ao registro das falhas e manutenções. E devido ao limite de tempo disponível, não serão realizados testes de aplicação nem pesquisa para avaliação da aplicabilidade do indicador. Além disso, a avaliação da qualidade da informação será baseada apenas em fatores técnicos e objetivos. Não serão considerados aspectos subjetivos como motivação dos funcionários, qualidade da comunicação interpessoal, etc.

1.5 – Aderência ao Programa de Pós-Graduação do EGC

Avaliar a qualidade da informação para sua utilização em processos de conhecimento demanda tanto uma percepção específica sobre o contexto no qual essa informação é gerada e utilizada quanto a construção de uma definição de trabalho para os próprios conceitos de qualidade e informação.

Neste trabalho, uma vez estabelecido o processo de análise de qualidade de informação, é proposto o desenvolvimento de uma estrutura que permita automatizá-lo parcialmente. Isso exige um diálogo entre conhecimentos de várias disciplinas e sua recontextualização para o desenvolvimento de uma aplicação baseada em conhecimento.

Esse foco na codificação do conhecimento para a construção de uma ferramenta de gestão das informações oriundas dos registros da atividade de manutenção é condizente com os objetivos do Programa EGC, que é voltado para ensino, pesquisa e desenvolvimento em codificação, gestão e disseminação dos conhecimentos em organizações e na sociedade.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Históricos das áreas de Manutenção e Bancos de Dados de Confiabilidade

No decorrer do século XX, a manutenção deixou de ser uma atividade secundária para tornar-se “uma atividade de importância estratégica para as empresas” (CARVALHO, 1993). Segundo Moubray (2001), essa evolução pode ser dividida, a partir da década de 30, em três gerações: a primeira que abrange o período até a Segunda Guerra Mundial; a segunda que se estende do início da década de 50 até o choque do petróleo do início da década de 70; e a terceira que inicia em meados da década de 70 e continua até os dias atuais.

Na Primeira Geração da Manutenção, como a indústria ainda não era altamente mecanizada, os períodos de inatividade dos equipamentos à espera de recuperação de falhas não eram muito importantes. Em geral os equipamentos eram simples e superdimensionados, o que os tornava confiáveis e fáceis de consertar. Como conseqüências, a necessidade de pessoal especializado era menor, as tarefas de manutenções periódicas eram restritas a serviços como limpeza e lubrificação e a expectativa para a área de manutenção se limitava a realização de reparos após a ocorrência das falhas (MOUBRAY, 2001).

Esse cenário mudou drasticamente durante a segunda guerra mundial, quando o aumento da demanda por produtos de todos os tipos em conjunto com a redução da mão de obra disponível, levaram a um grande incremento da mecanização da indústria (MOUBRAY, 2001).

Este novo ambiente deu origem a Segunda Geração da Manutenção, quando o crescimento da dependência das máquinas tornou relevante o impacto causado pelo tempo parado à espera de manutenção, e levou a idéia de que as falhas poderiam e deveriam ser prevenidas, que foi formalizada com a criação do conceito de manutenção preventiva. Na década de 60, essa manutenção consistia basicamente na realização de revisões gerais em intervalos fixos (MOUBRAY, 2001).

Além disso, os significativos montantes de capital imobilizado em bens fixos em conjunto com o rápido crescimento dos custos de capital que ocorreram neste período, fomentaram o desenvolvimento de várias iniciativas que visavam maximizar a vida desses bens, como o desenvolvimento de tarefas de coleta de dados para acompanhamento e divulgação de taxas de falhas.

Fragola (1996) destaca que as atividades de coleta de dados empíricos e desenvolvimento de bases de dados para predição de probabilidade de eventos futuros são bem antigas, podendo ser rastreadas pelo menos até o século 17, mas que somente com os desenvolvimentos tecnológicos do século 20 este trabalho ganhou uma nova dimensão. Diversas atividades de coletas de dados de confiabilidade desenvolvidas no século XX e destacadas em Fullwood (1999) são apresentadas no Quadro 2.1.

Segundo Fragola (1996), um dos primeiros projetos que utilizaram bases de dados de confiabilidade e risco foi o de automóveis, e o processo de coleta de dados começou quando o automóvel passou a ser utilizado como uma alternativa de meio de transporte para grande parte da população. Essa coleta sistemática de dados realizada para obter a frequência de manutenção desses automóveis foi a atividade de coleta de dados de confiabilidade destacada por Fullwood (1999) para o período de 1900 a 1930.

Para a década de 30, quando a aviação comercial despontou, Fullwood (1999) salienta a atividade de coleta de dados de confiabilidade referente a falhas de motores, realizada para embasar substituições na aviação comercial.

Na década de 40, a atividade evidenciada por Fullwood (1999) é a coleta realizada por militares para análise da frequência de falha de partes com vistas a identificar e tratar os "elos fracos" em seus equipamentos e sistemas. É no final desta década que, segundo Pecht (2001), a confiabilidade de sistemas e produtos emerge como uma área independente. Já O'Connor *et al.* (2002) e Saleh e Marais (2006) consideram que isto ocorreu apenas na década de 50.

Quanto a isso, Fullwood (1999) menciona apenas que com o aumento da complexidade de aviões, tanques, veículos e embarcações, desenvolvidos após a 2ª guerra mundial, houve a necessidade da engenharia de confiabilidade para mantê-los funcionando, e que antes disso os equipamentos normalmente eram avaliados apenas qualitativamente, por serem confiáveis ou não.

Foi a partir da década de 50 também que a NASA e os militares incluíram especificações de confiabilidade em pedidos de aquisição e, dessa forma, fomentaram o processo de coleta e análise de dados com acurácia estatística, embora isso elevasse seus custos (FULLWOOD, 1999). Como frutos desse cenário ocorreram as primeiras iniciativas de divulgação pública de dados de taxas de falha.

Organizações industriais como a *Radio Corporation of America (RCA)*, *General Electric (GE)* e *Motorola* publicaram catálogos contendo taxas de falha de componentes obtidas a

partir de dados de teste e de campo. Essas iniciativas são classificadas por Fragola (1996) como a Primeira Geração de Banco de Dados de Confiabilidade.

Quadro 2.1 – História Parcial da Coleção de Dados de Confiabilidade.

Fonte: adaptado de FULLWOOD (1999), p. 152, Tabela 4.1-1.

Período	Ramo	Tipo ou uso de Dados	Progresso
1900-1930	Automóveis	Frequência de manutenção	Coleta sistemática para obter a frequência
1930-1940	Aviação Comercial	Falhas de motor versus número de motores	Coleta para embasar as substituições
1940-1950	Militar	Frequência de falha de partes para tratar o problema dos "elos fracos".	Coleta para análise de sistemas.
1950-1960	Militar / Espaço (Titan)	Martin: Dados de confiabilidade de falhas de partes genéricas, modos de falhas durante o tempo de vida.	Taxas exponenciais, modos correlacionados com o tipo e uso da unidade $1E^{-6}/hr.$
1960-1965	Espaço Militar	Mil HDBK 217-1961 US Navy	Dados eletrônicos abrangentes
1960-1965	Espaço Militar	RADC Notebook 1961 USAF	Surveys e taxas relacionadas aos modos de falha
1960-1965	Espaço Militar	Taxas de falhas (ARINC, COLLINS, Boeing, etc.)	Criação de sistemas para coleta de dados
1965-1970	Espaço Civil	Trabalhos para o desenvolvimento do programa Apollo	PRACA (Problem Report And Corrective Action) e GIDEP
1965-1970	Comercial	Dados de falha e alerta de distribuidores comerciais baseados em estatísticas de segunda ordem	Rastreamento de medidas de confiabilidade em operação ligadas ao MSG.
1965-1975	Usina de força (comercial)	Coleção de frequência de falha	Dados coletados sistematicamente a partir de dados de plantas nucleares
1965-1975	Usina de força (comercial)	Relatório - Regulatory Guide 1.16 futura Licensee Event Report (LER)	Aplicação do método híbrido Delphi no desenvolvimento de banco de dados.
1965-1975	Usina de força (comercial)	Erros Humanos	Desenvolvimento de uma abordagem de coleta de dados abrangente para taxas de falhas de componentes
1965-1975	Usina de força (comercial)	Relatório de performance de unidade	Coleção abrangente de falhas de componentes de sistemas de dados nos EUA.
1975-1982	Usina de força (comercial)	Taxas de falha de equipamentos de segurança calculadas a partir de dados das plantas - NPRD	Coleção sistemática de dados a partir de dados de plantas nucleares
1975-1982	Usina de força (comercial)	Taxas de falha de equipamentos eletro-eletrônicos genéricos – IEEE 500 (1977)	Aplicação do método híbrido Delphi no desenvolvimento de banco de dados.
1975-1982	Usina de força (comercial)	Taxas de falhas de bombas, válvulas e diesel obtidas por dados de plantas	Coleção abrangente de dados de falhas de componentes.
1975-1982	Usina de força (comercial)	Dados de sistemas de utilidades	Coleção abrangente de dados de falhas de componentes nos EUA e na Europa.
1982-1997	Usina de força nuclear	NUREG/CR-4550	Dados usados para a NUREG-1150 e boa parte do IPE
1982-1997	Planta Química	CCPS, 1989	Uma coleção de dados de várias fontes para uso em segurança de plantas químicas

Fragola (1996) considera que, provavelmente, a primeira publicação amplamente difundida foi o *Martin Titan Handbook*, lançado em 9 de julho de 1959 com o título “*Procedure and Data for Estimating Reliability and Maintainability*”. Esta foi a primeira fonte conhecida a padronizar as taxas de falhas em termos de falhas por 10^6 horas. Continha várias taxas de falhas genéricas, mas não fornecia informações importantes, como a população incluída na amostra de testes e o número de falhas observado. Além disso, a representação exponencial da confiabilidade obtida a partir dos dados coletados foi estabelecida de modo ad hoc.

Foi com base nesse trabalho que surgiu a Segunda Geração dos Bancos de Dados de Confiabilidade, desenvolvida nas décadas de 60 e 70 e instigada em grande parte pelos militares nos EUA. Deste período, Fragola (1996) destaca o *MIL-Handbook-217*, o *Failure Rate Data Bank* (FARADA) e o *RADC Non-Electronic Reliability Notebook 's*, dos quais algumas de suas características são destacadas a seguir.

O *MIL-Handbook-217*, assim como o *Titan Handbook*, também utilizou a representação de taxas de falhas por 10^6 horas e o fator de ajuste de taxas de falhas em função de redundâncias ou condições operacionais (FRAGOLA, 1996). Foram publicadas revisões de **A** a **F** entre 1962 e 1995.

Inicialmente chamado de FARADA, o programa patrocinado em conjunto pelos comandantes do *US Army Material Command*, *Air Force Logistics Command* e *Air Force Systems Command*, encorajava a troca de dados sobre equipamentos e foi posteriormente rebatizado como GIDEP (*Government/Industry Data Exchange Program*). Os dados incluídos no banco de dados eram coletados a partir de dados de campo, de ensaios acelerados de vida realizados em laboratório e de testes de demonstração de confiabilidade. Essa foi a primeira fonte a utilizar um sistema de gerenciamento de banco de dados (FRAGOLA, 1996).

O *US Air Force Rome Air Development Center* (RADC) localizado em Rome, New York, desenvolveu o *RADC Non-Electronic Reliability Notebook 's* que na década de 70 continha taxas de falhas de mais de 300 tipos de componentes e partes. Dessa fonte destacam-se dois pontos: o primeiro é o uso de tabelas separadas para apresentar os dados oriundos de campo, de testes, de check-out, de demonstrações de confiabilidade e de dados de fabricantes, e o segundo é o fato de ser a primeira fonte de dados conhecida a perceber o problema de unir conjuntos de dados não homogêneos, e evitar esse problema usando o Teste F de Fisher, que verifica a homogeneidade dos dados (FRAGOLA, 1996).

Ainda no final da década de 50, o grande número de aviões comerciais existentes proporcionou dois importantes fatores para a evolução da atividade de manutenção: a

existência de um amplo conjunto de dados de falha e manutenção disponíveis e um custo total de manutenção grande o suficiente para justificar uma revisão minuciosa nos reais resultados das práticas existentes. Ao mesmo tempo a *Federal Aviation Agency* (FAA) tomava conhecimento que o uso de revisões gerais programadas não era eficaz para controlar a ocorrência de alguns tipos de falhas (MOUBRAY, 2001).

Então, Moubray (2001) relata que em 1960, a FAA e as companhias aéreas dos EUA formaram uma força tarefa para investigar as capacidades e potencialidades da manutenção preventiva e chegaram a duas importantes descobertas:

- Revisões Gerais programadas tinham pouco efeito na confiabilidade total de um item complexo, exceto quando este possuía um modo de falha dominante.
- Existem muitos itens para os quais não existe nenhuma forma efetiva de manutenção programada.

Diversos programas de confiabilidade foram então implementados nas companhias aéreas dos EUA e a análise e organização das lições aprendidas forneceu subsídios para a reavaliação da filosofia de manutenção vigente. Com o desenvolvimento do Boeing 747, foi criado um Grupo de Padronização de Manutenção ("*Maintenance Steering Group*") para supervisionar o desenvolvimento de seu plano de manutenção. O produto desse trabalho, conhecido como MSG 1 (*Maintenance Steering Group 1*), foi publicado em 1968 pela *Air Transport Association* em *Washington DC*, e aprimorado em 1970 com a publicação do MSG 2, que foi usado para o Lockheed L-1011 e para o Douglas DC-10, obtendo resultados expressivos, conforme o Quadro 2.2 (MOUBRAY, 2001).

Quadro 2.2 – Redução do plano de manutenção em relação ao requerido pelo DC-8
Fonte: adaptado de MOUBRAY (2001), p. 320.

Quantidade de itens que requerem <i>overhaul</i>	
DC-8	339 itens
DC-10	7 itens
H/H requerido antes da revisão geral de 20.000 horas de operação	
DC-8	4.000.000 homem/hora
Boeing 747 original	66.000 homem/hora

Esse estudo mostrou que o conceito vigente na época, de que as falhas ocorriam de acordo com um padrão conhecido como curva da banheira, era inadequado pois, na verdade, existiam 6 padrões distintos de falhas. Para tratá-los foram desenvolvidas novas técnicas, como a manutenção por monitoramento, e fortalecidas novas áreas, como a Engenharia da Manutenção e a Engenharia da Confiabilidade.

Essa mudança de paradigma em conjunto com o desenvolvimento teórico e tecnológico desse período levou, em meados da década de 70, ao surgimento da Terceira Geração da Manutenção. Dentre as várias novidades dessa geração da manutenção, destacam-se: a incorporação dos conceitos das ciências comportamentais; o surgimento e desenvolvimento da Terotecnologia e da Logística; a oficialização do TPM (*Total Productive Maintenance*) na empresa japonesa Nippon Denso, em 1971; o lançamento em 1978 do documento de referência inicial do RCM (*Reliability Centered Maintenance*); a divulgação das pesquisas que identificaram os seis padrões de falha; a intensa evolução dos conceitos de confiabilidade e qualidade total pelo mundo; o expressivo crescimento de novos conceitos e técnicas de manutenção; a introdução da TPM e RCM no Brasil no final da década de 80; a revolução da informação e conhecimento centrado em computação; a microeletrônica; os microcomputadores; a radiodifusão; as redes industriais; e a Internet (ALKAIM, 2003).

As principais mudanças ocorridas ao longo das três Gerações da Manutenção podem ser agrupadas em três tópicos principais: crescimento das expectativas de manutenção (novas expectativas); melhor entendimento sobre como os equipamentos falham (novas pesquisas) e uma larga evolução nas técnicas de gerenciamento de manutenção (novas técnicas e conceitos) (MOUBRAY, 2001).

Essa evolução também alcançou os Bancos de Dados de Confiabilidade, levando ao surgimento da sua Terceira Geração na década de 80, quando várias simplificações que comprometiam a qualidade das informações obtidas passaram a ser tratadas, dentre elas: a natureza não-homogênea das sub-populações; a diferenciação entre as taxas de falha baseadas em unidades de tempo e as baseadas em unidades de demanda; e a categorização dos modos de falha (FRAGOLA, 1996).

Dentre os bancos de dados de confiabilidade da terceira geração, destaca-se o OREDA (*Offshore Reliability Data*), que apresenta dados de equipamentos da indústria de exploração e produção de petróleo e gás natural. Iniciado em 1981, como um *Joint Industry Project* (JIP) de diversas companhias do Mar do Norte e do Mar Adriático, hoje possui dados relativos a

mais de 15.000 equipamentos em 250 instalações, com mais de 33.000 dados de falhas e 54.000 registros de manutenção (OFFSHORE RELIABILITY DATA (OREDA), 2009).

Nesse período foram elaboradas para comercialização quatro edições (publicadas em 1984, 1992, 1997 e 2002) e existe previsão de uma nova edição para este ano (2009).

O OREDA também serviu de base para a criação da Norma ISO 14224, que fornece orientações para o desenvolvimento de Bancos de Dados de Confiabilidade, visando à compatibilidade dos dados para troca de informações entre as empresas da área de óleo e gás.

Esse é um importante passo para obter um maior entendimento a respeito do comportamento dos equipamentos dessa área, conforme atesta a experiência da aviação comercial descrita anteriormente.

Fundamentalmente, é importante destacar que as expectativas da manutenção não mais se restringem à programação e execução de tarefas, uma vez que, para o estabelecimento de um bom programa de confiabilidade e manutenabilidade, é imprescindível o controle e registro de todas as atividades e eventos relevantes ocorridos.

Uma representação cronológica da evolução das áreas de manutenção e de bancos de dados de confiabilidade durante o século XX é apresentada na Figura 2.1 que destaca também alguns dos eventos relevantes desse período.

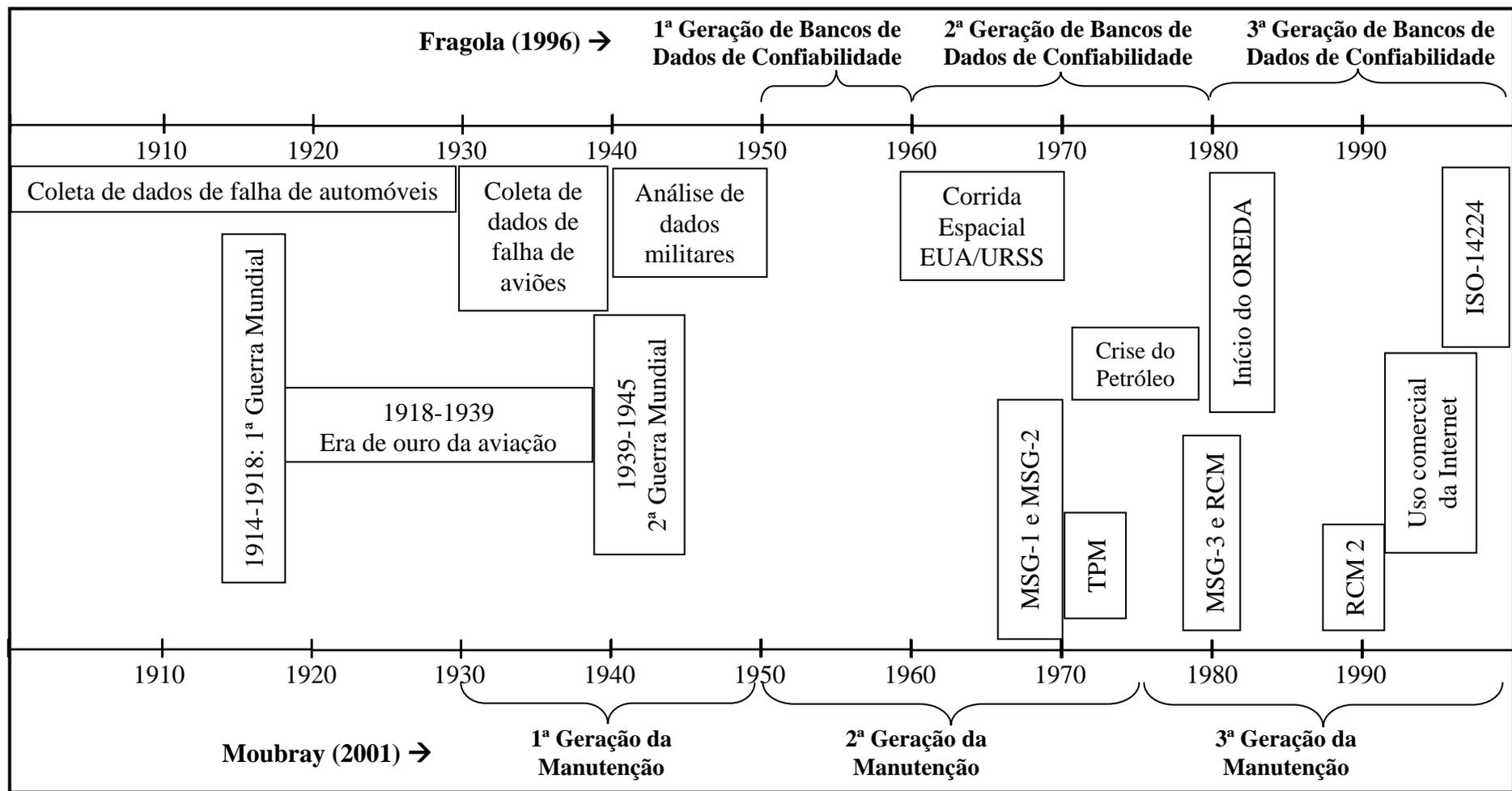


Figura 2.1 – Evolução da Manutenção e do desenvolvimento de Bancos de Dados de Confiabilidade durante o século XX.

2.2 Registros de Manutenção e Confiabilidade Conforme a Norma ISO 14224

A norma ISO 14224, *Petroleum, petrochemical and natural gas industries – Collection and Exchange of maintenance data for equipment*, foi elaborada pelo comitê técnico ISO/TC 67, *Materials, equipment and offshore structures for petroleum, petrochemical and natural gas industries*, e publicada inicialmente em 1999.

Essa edição foi cancelada com a publicação da segunda edição, em 15 de dezembro de 2006, a qual foi modificada tecnicamente e expandida. Além disso, dos seus anexos que antes eram apenas recomendações, o anexo B, *Interpretation and notation of failure and maintenance parameters*, passou a ser normativo. Os demais anexos fornecem recomendações para o uso de dados de confiabilidade e manutenção em várias aplicações.

Seu objetivo é prover um ponto de vista comum que facilite às companhias da área de petróleo, petroquímica e gás natural, a troca de informações oriundas de plantas, fabricantes ou empreiteiros. Mais do que a simples comparação de resultados obtidos, quanto mais dados forem compartilhados por diferentes empresas em diversos ambientes, maior será o entendimento sobre o comportamento de seus sistemas e equipamentos, o que permitirá tanto reduzir a margem de incerteza em estudos de riscos quanto concentrar as ações de manutenção em atividades mais efetivas para garantir de modo rentável a confiabilidade e disponibilidade esperadas.

A ISO 14224 é composta por 9 capítulos. O primeiro estabelece o escopo da norma; o segundo traz algumas referências normativas relevantes; o terceiro possui alguns termos e definições importantes; o quarto contém a lista de abreviações utilizadas; o quinto detalha as aplicações e limitações da norma; o sexto lista diversos benefícios proporcionados pela coleta e compartilhamento de dados de confiabilidade e manutenção; no sétimo são esboçadas algumas diretrizes para o controle da qualidade dos dados; o oitavo versa sobre os critérios para o estabelecimento de fronteiras para os equipamentos, sobre a classificação taxonômica a ser adotada e sobre questões relativas ao registro de tempo de eventos relevantes; e o nono detalha quais os dados recomendados a serem coletados, abrangendo três categorias; dados de equipamento, dados de manutenção e dados de falha.

É importante notar que essa norma não tem intenção de detalhar como um programa de manutenção deve ser criado ou controlado, mas especifica que dados precisam ser gravados para permitir a realização de estudos de confiabilidade e manutenibilidade.

Parte desses dados são estruturados em um sistema de classificação com 9 níveis. Os 5 primeiros representam uma categorização de alto nível que identifica as aplicações existentes

nas plantas, unidades e áreas abrangidas pela norma, de forma a retratar o contexto operacional do equipamento. Por sua vez, os últimos 4 níveis descrevem o equipamento.

Para o objetivo deste trabalho, serão estudados apenas os dados solicitados pela norma para o acompanhamento do ciclo de vida operacional de um motor elétrico, o que abrange os capítulos oitavo e nono, o exemplo exposto no anexo A (*Equipment-class attributes*), e os anexos B (*Interpretation and notation of failure and maintenance parameters*) e D (*Typical requirements for data*).

A primeira informação essencial é a definição da fronteira (item 8.1 da ISO 14224) do que consiste um motor elétrico (Figura 2.2). Isto inclui tanto a identificação das interfaces quanto da subdivisão interna do equipamento. O Quadro 2.3 apresenta os tipos de motores elétricos e o Quadro 2.4 detalha as subdivisões previstas pela ISO 14224.

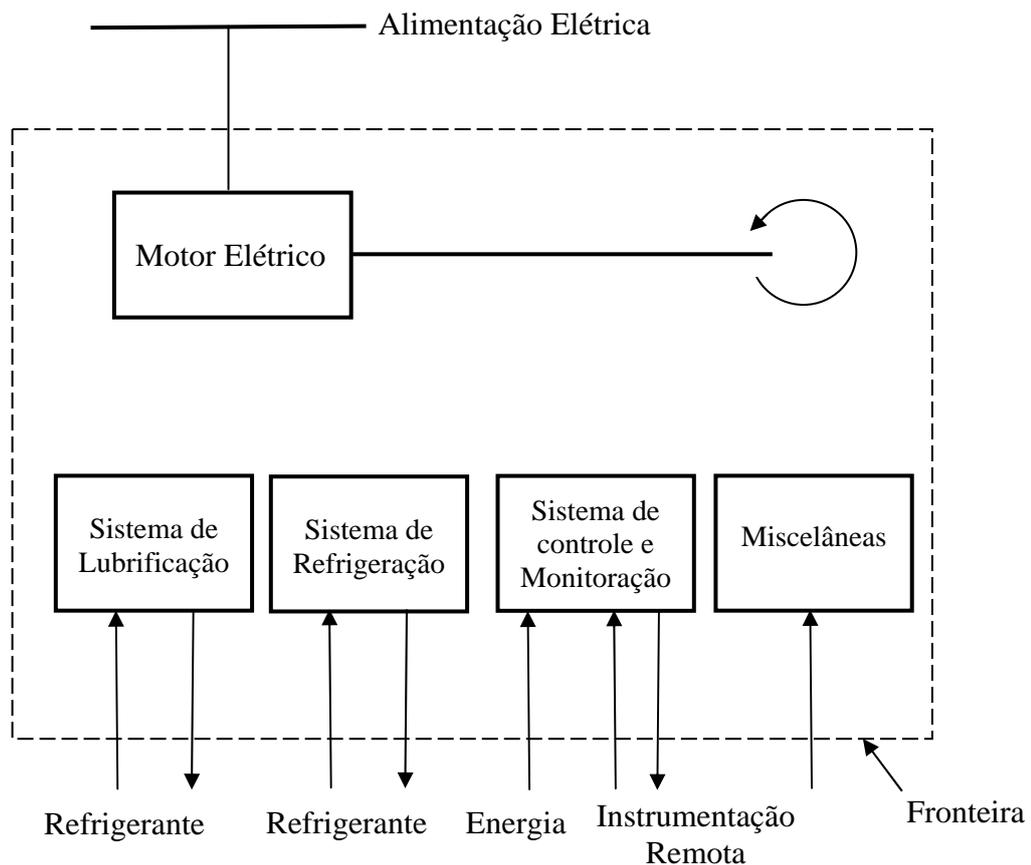


Figura 2.2 – Fronteira para Motores Elétricos.
Fonte: ISO 14224, Figura A.4, p. 45.

Quadro 2.3 – Classificação por Tipo para Motores Elétricos

Fonte: ISO 14224, Figura A.14, p. 45

Classe de Equipamento – Nível 6		Tipo de equipamento	
Descrição	Código	Descrição	Código
motor elétrico (<i>Electric motor</i>)	EM	Corrente alternada (<i>Alternating current</i>)	AC
		Corrente contínua (<i>Direct current</i>)	DC

Quadro 2.4 – Subdivisão de Motores Elétricos

Fonte: ISO 14224, Figura A.15, p. 46

Classe	Motor Elétrico				
Subunidades	Corpo do motor elétrico	Controle e monitoração ^a	Sistema de lubrificação	Sistema de refrigeração	Miscelânea
Itens Manuteníveis	Rotor/Estator Excitação do Rotor Rolamento Radial Rolamento Axial	Dispositivo de atuação Unidade de Controle Fonte de alimentação interna Monitoração Sensores ^b Válvulas Cabeamento Tubulação Selagem	Reservatório Bomba Motor Filtro Refrigerador (<i>cooler</i>) Válvulas Tubulação Óleo	Ventilador Filtros Trocador de calor Bomba Motor Válvulas Tubulação	Tampa (<i>Hood</i>)

^a Normalmente não existe um sistema de controle extra para motores elétricos. Isso se aplica, por exemplo, para motores Ex-p, que tem sua pressão interna monitorada, ou para grandes motores que possuem monitoração de temperatura.

^b Deve ser especificado o tipo de sensor, por exemplo, de pressão, de temperatura, de nível, etc.

Uma vez definido o que abrange a classe motor elétrico, é preciso identificar quais são suas características técnicas mais relevantes. Para motores elétricos, a ISO 14224 recomenda que sejam gravadas as informações descritas no Quadro 2.5.

As informações representadas na Figura 2.2 e nos Quadros 2.3, 2.4 e 2.5 definem o que deve ser considerado como parte de um motor elétrico. Todas as classes que serão abrangidas por um banco de dados de confiabilidade e manutenibilidade devem ser especificadas desse modo, com o cuidado de evitar sobreposições entre as fronteiras das classes. Para auxiliar nesse trabalho, o Anexo A da ISO 14224 possui várias fronteiras recomendadas para diversas classes de equipamentos.

Além dos dados de Classe (nível 6), Subunidade (nível 7) e Item manutenível (nível 8), a ISO 14224 ainda apresenta um último nível, que é o de Parte (nível 9) do equipamento,

correspondendo às peças ou materiais que o compõem. Como este nível é muito específico, seu monitoramento é opcional.

Conforme já mencionado, além dos dados relativos aos equipamentos é necessário registrar também seus contextos operacionais, que são identificados pelas informações relativas à Indústria (Nível 1), Categoria de Negócio (Nível 2), Categoria de Instalação (Nível 3), Categoria de Planta ou Unidade (Nível 4) e Seção ou Sistema (Nível 5). O Quadro 2.6 apresenta um exemplo de preenchimento dessas informações.

Quadro 2.5 – Características técnicas de motores elétricos
Fonte: ISO 14224, Tabela A.16, p. 46

Nome	Descrição	Unidade ou lista de valores	Prioridade
Máquina acionada	Classe, tipo e identificação do equipamento acionado	Especificar	Alta
Tipo de proteção Ex	Marcação Ex do equipamento para atmosferas explosivas	Norma ABNT NBR IEC 60079-0.	Alta
Grau de Proteção	Grau de proteção.	Normas ABNT NBR IEC 60529 e 60034-5.	Média
Tipo de Motor	Tipo	Indução, comutador (d.c.), síncrono	Média
Velocidade	Velocidade de Projeto	RPM (Rotações por minuto)	Média
Potência – projeto	Potência máxima de saída – projeto	Kilowatt	Média
Voltagem	Voltagem de alimentação – projeto	Volts	Média
Classe de isolamento – rotor ^a	<i>Insulation class</i> de acordo com a norma IEC 60034-1	Y, A, E, B, F, H	Média
Elevação de temperatura – rotor ^a	<i>Temperature rise</i> de acordo com a norma IEC 60034-1	Y, A, E, B, F, H	Média
Classe de isolamento – estator	<i>Insulation class</i> de acordo com a norma IEC 60034-1	Y, A, E, B, F, H	Média
Elevação de temperatura – estator	<i>Temperature rise</i> de acordo com a norma IEC 60034-1	Y, A, E, B, F, H	Baixa
Variador de velocidade	Especifica se possui ou não.	Sim/Não	Baixa
Power – operating	Especifica a potência aproximada na qual o motor opera na maior parte do tempo monitorado.	Kilowatt	Baixa

^a Irrelevante para motores de indução.

Além dos dados específicos de motor elétrico, descritos no Quadro 2.5, existe um conjunto de dados que deve ser preenchido para todas as classes de equipamento, o qual é apresentado no Quadro 2.7. Todos esses dados são relativos ao 6º nível da classificação taxonômica.

Quadro 2.6 – Exemplo de dados de Local de Instalação (Tag)

Fonte: adaptação da Tabela 5 da ISO 14224, p. 25

Categoria principal	Nível Taxonômico	Hierarquia Taxonômica	Definição	Exemplo
Dados relativos ao contexto operacional (Indústria e Locação)	1	Indústria	Tipo principal da indústria	Petróleo
	2	Categoria de negócio*	Tipo de negócio ou de fluxo processado	E&P – Exploração e Produção de petróleo (<i>Upstream</i>)
	3	Categoria de instalação ¹	Tipo de instalação (<i>facility</i>)	Produção <i>offshore</i> de óleo e gás
	3	Código da instalação*	Nome ou Código que identifica a instalação	Campo de Caratinga
	3	Localização Geográfica	Local onde se encontra a instalação	Rio de Janeiro, Brasil
	4	Proprietário	Código ou Nome do Proprietário	Petrobras
	4	Categoria de Planta ou Unidade ² *	Tipo de planta ou unidade	FPSO (Floating production storage and offloading)
	4	Código da Planta ou Unidade*	Nome ou Código que identifica a Unidade	P-48
	5	Seção ou Sistema ³	Seção ou sistema principal da planta	Sistema de água potável (Utilidade)
	6	Localização / Identificação do equipamento (número do tag)	Código que identifica um equipamento em um diagrama de processo	M-B-512201-A
	6	Descrição do equipamento	Descrição do serviço do equipamento	Motor da bomba para o vaso hidrofórico
	6	Data de comissionamento	Data em que o Tag foi liberado para operação	04.02.2005

¹. A tabela A.1 da ISO 14224 sugere algumas opções para esse campo.

². A tabela A.2 da ISO 14224 sugere algumas opções para esse campo.

³. A tabela A.3 da ISO 14224 sugere algumas opções para esse campo.

* Conjunto mínimo de dados exigidos pela ISO 14224.

As informações tratadas até este ponto pertencem à categoria de dados de equipamentos e permitem caracterizar um equipamento e seu contexto operacional. As outras duas categorias especificadas no capítulo 9 da ISO 14224, dados de manutenção e dados de falha, descrevem os dados de serviços de manutenção e eventos de falha ocorridos nesses equipamentos e que devem ser registrados.

Quadro 2.7 – Exemplo de dados de equipamento
 Fonte: adaptação da Tabela 5 da ISO 14224, p. 26

Categoria principal	Hierarquia Taxonômica	Definição	Exemplo
Atributos de Equipamento	Classe do equipamento ¹ *	Classe que agrupa equipamentos similares	Motor Elétrico
	Tipo do equipamento ² *	Identifica subgrupos relevantes nas classes	Corrente alternada
	Identificação única de equipamento	Número de série do equipamento	BI60363
	Fabricante *	Nome do fabricante do equipamento	WEG
	Modelo	Identificação do modelo	160M_440V_2P_20CV
	Dados específicos da Classe	Conjunto de atributos relevantes para cada classe	Ver Figura 2.7
Dados de Operação	Estado/modo normal de operação *	Descreve o regime normal de funcionamento do equipamento	Em funcionamento
	Data do início da operação	Data em que o equipamento começou a ser utilizado	10.03.2005
	Data do início da operação no Tag *	Data em que o equipamento começou a operar no Tag atual	10.06.2006
	Tempo monitorado (h) – Calculado *	Tempo de monitoração do equipamento em horas	33.520h
	Tempo em operação (h)	Tempo de operação medido ou calculado em horas	18.754h
	Número de demandas / partidas*	Número de vezes que o equipamento foi solicitado – tanto para operação quanto para testes	108
Informações adicionais	Parâmetros operacionais relevantes	Conjunto de dados operacionais relevante para cada classe	Motor abrigado
	Informações adicionais em texto livre	Campo para entrada de informações relevantes em forma de texto livre (Usar quando for necessário)	Motor elétrico de indução; trifásico padrão; uso industrial; Forma construtiva B3D.
	Origem dos dados	Identificar as Folhas de dados e diagramas de processo utilizados (Usar quando for necessário)	FD-5122.01-023-549-A

¹. A tabela A.4 da ISO 14224 sugere algumas opções para esse campo.

². Ao longo do Anexo A da ISO 14224 são apresentadas algumas opções para esse campo.

* Conjunto mínimo de dados exigidos pela ISO 14224.

As informações requeridas pelos registros de falha e manutenção normalmente são comuns a todas as classes de equipamentos, exceto em casos como os equipamentos submarinos, que requerem alguns tipos específicos de dados. Esses casos não são discutidos neste trabalho.

Para registrar um serviço de manutenção, é necessário identificar o tipo e a atividade da manutenção executada. A ISO 14224 reconhece dois tipos básicos de manutenção: Corretiva (C), executada para restabelecer um item após a sua falha; e Preventiva (P), que visa prevenir a ocorrência de falhas em um item. Quanto às atividades, são 12 ao todo, cujas descrições são apresentadas no Quadro 2.8.

Quadro 2.8 – Atividades de manutenção
Fonte: adaptação da Tabela B.5 da ISO 14224

Atividade	Descrição	Exemplos	Uso
Reposição	Substituição por outro, reformado ou novo, do mesmo fabricante e modelo	Substituição de um rolamento usado, de um mancal desgastado, etc.	C, P
Reparo	Ação manual de manutenção executada para restaurar o item ao seu estado ou aparência original	Reconectar, soldar, tampar, etc.	C
Modificação	Substituição, reforma ou alteração de um equipamento, ou parte, por outro diferente em relação ao tipo, marca, material ou projeto	Instalação de um filtro com diâmetro de malha menor, substituição de uma bomba por um tipo diferente, reconfiguração, etc.	C, P
Ajuste	Devolver algum parâmetro a sua faixa de tolerância	Alinhar, ajustar, calibrar, balancear, etc.	C, P
Readaptação (<i>Refit</i>)	Pequenos serviços / reparos para devolvê-lo a sua aparência aceitável, interna e externa	Polir, limpar, esmerilhar, pintar, revestir, lubrificar, trocar o óleo, etc.	C, P
Verificação	A investigação da causa da falha não demanda nenhuma ação mais complexa do que um reinício ou restauração da configuração inicial	Particularmente relevante para falhas funcionais de equipamentos como detectores de fogo e gás	C
Serviço	Pequenos serviços periódicos, normalmente executados sem desmontar o item	Limpeza, reabastecimento de consumíveis, ajustes e calibrações	P
Teste	Testes periódicos de função ou performance	Teste funcional do detector de gás, acurácia do indicador, etc.	P
Inspeção	Inspeção ou verificação periódica: exame detalhado do item, com ou sem desmontagem, normalmente pelo uso dos sentidos	Todos os tipos de verificações gerais. Inclui pequenos serviços como parte da tarefa de inspeção	P
Revisão geral	Revisão geral (<i>overhaul</i>)	Ação abrangente com grande quantidade de desmontagem e substituição de peças	C, P
Combinação	Inclui várias das atividades acima	Quando existir, é preferível registrar a atividade dominante	C, P
Outros	Outra atividade diferente das especificadas acima	–	C, P

A ISO 14224 destaca algumas observações sobre essas atividades de manutenção:

- Nas manutenções corretivas, esta informação descreve o tipo de ação de restauração que foi executada.
- As atividades reparo, substituição, revisão geral ou modificação são consideradas predominantes quando executadas em conjunto com readaptação ou ajuste.
- Um reparo é uma ação executada para corrigir poucas falhas, normalmente executada no local onde o equipamento está instalado.
- As atividades de verificação incluem tanto as circunstâncias em que a causa da falha foi descoberta, mas nenhuma ação de manutenção foi considerada possível ou necessária para ser executada, quanto quando nenhuma causa para a falha pôde ser encontrada.
- Uma revisão geral é usada para identificar quando vários reparos são realizados para a correção de diversas falhas ou de uma grande falha, requerendo um extenso trabalho ou uma renovação completa de uma subunidade do equipamento. É considerada como uma ação de manutenção de 3º escalão, efetuada por pessoal especializado normalmente em uma oficina. Também pode ser realizada de forma preventiva, quando são trocadas várias peças que normalmente sofrem desgaste durante sua operação.
- Quando a atividade combinação for escolhida, as atividades executadas devem ser listadas em outro campo.
- A modificação é uma alteração do equipamento original na qual o projeto original foi alterado ou um item foi substituído por outro de diferente modelo ou fabricante.
- Quando a modificação é de natureza significativa, não é considerada como uma ação da manutenção, apesar de poder ser executada pela equipe de manutenção ou em cooperação com a mesma.
- Se o equipamento tiver sido totalmente substituído por um novo ou for modificado é recomendável tratá-lo como um novo equipamento, reiniciando a contagem dos parâmetros de tempo para esse tag ou equipamento. Entretanto isso não se aplica para equipamentos de baixa complexidade ou naqueles em que a reposição é considerada como uma tarefa normal de manutenção.

O registro completo de uma manutenção executada deve conter informações sobre a identificação do serviço executado, a caracterização desse serviço, os recursos utilizados e os tempos envolvidos. Todos os dados solicitados pela norma para um registro de manutenção estão presentes no Quadro 2.9.

Quadro 2.9 – Dados de manutenção a serem registrados
 Fonte: adaptação das Tabelas 8 e D.4 da ISO 14224

Categoria	Dados a serem gravados	Descrição
Identificação do serviço	Número do registro de Manutenção *	Chave que identifica a tarefa de manutenção
	Identificação/Tag do Equipamento *	Número de tag ou número de série
	Número do registro de falha *	Chave que identifica um registro de falha
Caracterização do serviço	Datas da manutenção *	Datas relevantes como: Data Desejada/Prevista, Data programada, Data de execução e Data de registro da atuação.
	Categoria de manutenção *	Categoria principal (corretiva ou preventiva)
	Prioridade da manutenção	Alta, média ou baixa
	Periodicidade (manutenção oriunda de plano)	Tempo calendário ou operacional (Não utilizado para manutenções corretivas)
	Atividade de Manutenção	Identificada conforme a tabela B.5 da ISO 14224
	Impacto da manutenção na operação da planta	Nenhum, parcial ou total
	Subunidade atendida pela manutenção	Nome da subunidade que recebeu manutenção (pode ser omitida em manutenções preventivas)
	Item(s) manutenível(is) atendido(s) pela manutenção	Especificar o Item(s) manutenível(is) que recebeu manutenção (pode ser omitido em manutenções preventivas)
	Sobressalente	Peças trocadas ou reparadas
	Localização do sobressalente	Disponibilidade de sobressalentes (Local, distante, fabricante, etc.)
Recursos Utilizados	Homem-hora de manutenção, por disciplina	Homem-hora de manutenção, por disciplina (mecânica, elétrica, instrumentação, etc.)
	Homem-hora de manutenção, total	Valor total empregado de Homem-hora de manutenção
	Equipamentos de manutenção utilizados	Talha, medidor padrão, etc.
Tempos da atividade	Tempo de manutenção ativa *	Tempo durante o qual a atividade de manutenção foi efetivamente executada
	<i>Down time</i> *	Tempo em que um item permanece desativado em função de uma tarefa de manutenção
	<i>Lead Time</i>	Tempo aguardando a chegada de materiais para uma tarefa de manutenção
	<i>Turn Around Time (TAT)</i>	Tempo que uma solicitação de manutenção leva para ser concluída
	Atrasos da manutenção	Identificação das Causas do <i>Down time</i> . (logística, condições climáticas, andaime, ausência de sobressalentes, etc.)
Observações	Informações Adicionais	Fornecer mais detalhes sobre a ação de manutenção e recursos utilizados

* Conjunto mínimo de dados exigidos pela ISO 14224.

Além do registro do serviço de manutenção, é requerido também o registro das falhas detectadas, contendo os dados de identificação e caracterização das mesmas. Dentre os dados de caracterização estão o método de detecção, o modo de falha, o mecanismo de falha e a causa da falha.

A ISO 14224 prevê apenas 10 categorias de detecção de falhas, listadas no Quadro 2.10.

Quadro 2.10 – Métodos de detecção de falhas
Fonte: adaptação da Tabela B.4 da ISO 14224

Número	Origem ¹	Descrição	Atividade
1	Manutenção Periódica	A falha descoberta exige a troca de uma peça ou a execução de um procedimento que não estavam previstos na ordem de manutenção	Atividades programadas
2	Teste Funcional	Falha descoberta ao tentar ativar uma determinada função e não receber a resposta esperada. Este é o método típico para detecção de falhas ocultas	
3	Inspeção	Falha descoberta durante uma inspeção planejada, como uma inspeção visual ou um ensaio não destrutivo	
4	Monitoração Periódica de Condição ²	Falha descoberta durante uma monitoração de condição periódica (termografia, análise de vibração, etc.) ³	
5	Monitoração Contínua de Condição ²	Falha descoberta por um sistema de monitoração contínua de condição ³	Monitoração contínua
6	Interferência na Produção	Falha descoberta pelos efeitos causados na produção como: distúrbio, redução, etc.	
7	Observação casual	Falha detectada por uma pessoa ao observar a planta, com o uso dos seus sentidos (ao perceber, por exemplo, uma temperatura ou ruído maior que o normal)	Ocorrências casuais
8	Manutenção Corretiva	Falha detectada durante a execução de uma manutenção corretiva	
9	Falha sob demanda	A falha se manifesta quando o equipamento recebe uma demanda para iniciar uma operação (por exemplo, uma válvula que não fecha ao ser comandada)	
10	Outros	Outro método de detecção ou uma combinação dos métodos citados acima.	Outros

¹ Para detectores de fogo e gás, sensores de processo e unidades de controle, usar a seguinte interpretação:

Teste funcional	Teste funcional periódico
Observação casual	Observação feita em campo
MC periódica	Estado anormal descoberto na sala de controle sem a geração de alerta
MC contínua	Um alerta ou alarme foi gerado na sala de controle

² Uma monitoração de condição implica no uso de um equipamento específico ou de algoritmos para monitorar a condição de um equipamento, ou seja, são diferentes de “testes” e “inspeções”.

³ No caso de monitorações realizadas no conjunto acionador/carga, (como vibração ou ferrografia), a falha deve ser registrada sempre para a carga acionada.

O modo de falha é o efeito pelo qual uma falha é observada em um item. É normalmente registrado em relação à classe do equipamento, entretanto, pode ser registrado para níveis inferiores como a subunidade ou mesmo o item manutenível. No Quadro 2.11 são apresentados os modos de falha previstos pela ISO 14224 para a classe motor elétrico.

Quadro 2.11 – Modos de falhas de motores elétricos
Fonte: ISO 14224, Tabela B.6

Tipo*	Código	Descrição	Exemplos
1	FTS	Falha na partida sob demanda (<i>Failure to start</i>)	Não funciona quando solicitado.
1	STP	Falha na parada sob demanda (<i>Failure to stop</i>)	Não para quando solicitado.
2	UST	Parada espúria (<i>Spurious stop</i>)	Parada inesperada.
3	BRD	Quebra (<i>Breakdown</i>)	Danos severos (queima, fratura, etc.)
2	HIO	Saída alta (<i>High output</i>)	Sobre velocidade.
2	LOO	Saída baixa (<i>Low output</i>)	Sub velocidade, torque abaixo do requerido.
2	ERO	Saída errática (<i>Erratic output</i>)	Oscilando, instável, etc.
3	ELU	Vazamento externo de fluido de utilidade (<i>External leakage – utility medium</i>)	Lubrificante, água de resfriamento
3	VIB	Vibração (<i>Vibration</i>)	Vibração anormal
3	NOI	Ruído (<i>Noise</i>)	Ruído anormal
3	OHE	Superaquecimento (<i>Overheating</i>)	Na exaustão, na água de refrigeração, em partes do equipamento, etc.
2 (3)	PDE	Desvio de Parâmetro (<i>Parameter deviation</i>)	Parâmetros monitorados acima dos limites permitidos (sobre corrente, etc.)
2 (3)	AIR	Leitura anormal de instrumento (<i>Abnormal instrument reading</i>)	Alarme falso, indicação errada da instrumentação
3	STD	Deficiência estrutural (<i>Structural deficiency</i>)	Danos de material (rachaduras, desgaste, fratura, corrosão, etc.)
3	SER	Pequenos problemas em operação (<i>Minor in-service problems</i>)	Itens frouxos, descoloração, sujeira, etc.
—	OTH	Outros (<i>Other</i>)	Modo de falha diferente dos descritos acima
—	UNK	Desconhecido (<i>Unknown</i>)	Informações disponíveis insuficientes para definir um modo de falha.

* Identifica a que tipo, o modo de falha normalmente se refere:

- 1) A função desejada não é obtida;
- 2) Uma função secundária é perdida ou está fora dos limites operacionais;
- 3) Uma indicação de falha é observada, mas não provoca nenhum efeito imediato ou crítico na função do equipamento. Normalmente está relacionado a falhas incipientes.

O Mecanismo de falha é o processo físico, químico ou outro processo ou combinação de processos que conduziu a uma falha. Em termos práticos, o mecanismo de falha representa um modo de falha, no nível do item manutenível, que descreve a causa aparente da falha. O evento que desencadeou o mecanismo de falha, sempre que identificado, deve ser armazenado no campo causa da falha.

Os mecanismos de falha previstos pela ISO 14224 e apresentados no Quadro 2.12, são agrupados em 6 categorias: mecânicas, materiais, instrumentação, elétricas, influência externa e diversos. A norma recomenda que os mecanismos de falha sejam descritos pelos códigos pertencentes a essas categorias, mas admite que os códigos das categorias também podem ser usados para identificar o mecanismo de falha quando não houver informação suficiente para uma descrição mais precisa. Além disso, quando houver mais de um descritor de mecanismo de falha aplicável, o responsável pela aquisição dos dados deve utilizar aquele que for o mais importante, e sempre tentar evitar os códigos 6.3 e 6.4.

Para não confundir esses conceitos é preciso atentar que o modo de falha descreve como foi percebida a ocorrência de uma falha no equipamento; que o mecanismo de falha descreve os processos (tecnicamente dedutíveis) que degradaram o equipamento; e que a causa desses processos é descrita em outro campo, conhecido como causa da falha, cujo objetivo é revelar a causa subjacente ou raiz de uma falha.

Cabe então ao campo causa da falha identificar o(s) evento(s) iniciador(es) (“causas raízes”) na seqüência que conduziu a uma falha de um equipamento. A ISO 14224 trabalha com cinco categorias de falha: relacionadas ao projeto; relacionadas à fabricação/instalação; relacionadas à operação/manutenção; relacionadas ao gerenciamento; e diversos, as quais estão apresentadas no Quadro 2.13, juntamente com suas subdivisões e códigos relacionados a serem usados nas bases de dados. Assim como para o mecanismo de falha, a causa da falha pode ser registrada em qualquer um desses dois níveis, dependendo da quantidade de informações disponíveis.

Como as causas de falhas não são normalmente conhecidas a fundo quando a falha é observada, para que a causa raiz de uma falha seja descoberta pode ser útil a realização de uma RCA (*Root Cause Analysis* – Análise de Causa Raiz). Isso é particularmente relevante para falhas de uma natureza mais complexa e quando é importante evitar a falha devido a suas graves conseqüências em relação à produção, segurança ou meio ambiente, ou taxas de falhas anormalmente altas em comparação com a média, ou ainda, falhas com um alto custo de reparo.

Quadro 2.12– Mecanismos de falha (continua)

Fonte: adaptado da ISO 14224, Tabela B.2

Mecanismo de falha		Subdivisão do mecanismo de falha		Descrição do mecanismo de falha
Número	Notação	Número	Notação	
1	Falha mecânica	1.0	Geral	Uma falha relacionada a algum defeito mecânico, sem nenhum outro detalhe conhecido
		1.1	Vazamento	Vazamento externo e interno, seja de líquidos ou gases: Quando o modo de falha do equipamento for codificado como “vazamento”, deve-se usar um mecanismo de falha que seja mais relacionado com a causa da falha. Na ausência de maiores detalhes, deve ser escolhido o código 6.4 (desconhecido)
		1.2	Vibração	Vibração anormal: Se o modo de falha no nível do equipamento também for vibração, deve-se usar um mecanismo de falha que seja mais relacionado com a causa da falha. Na ausência de maiores detalhes, deve ser escolhido o código 6.4 (desconhecido)
		1.3	Falha por folga/alinhamento	Falha devido a folgas ou deficiência de alinhamento
		1.4	Deformação	Distorção, dobramento, empenamento, deformação elástica, amassamento, encolhimento, empolamento, deslizamento, etc.
		1.5	Falha na fixação	Desconexão, itens frouxos
		1.6	Agarramento	Agarramento, grimpamento, emperramento por motivos outros que não falhas de deformação ou de folga/alinhamento
2	Falha de material	2.0	Geral	Uma falha relativa a um defeito no material, sem nenhum outro detalhe conhecido
		2.1	Cavitação	Relevante para equipamentos como bombas e válvulas
		2.2	Corrosão	Todos os tipos de corrosão, tanto a molhada (eletroquímica) quanto a seca (química)
		2.3	Erosão	Desgaste erosivo
		2.4	Desgaste	Desgaste abrasivo e adesivo como, por exemplo, riscamento, descamação, desgaste, corrosão por atrito
		2.5	Ruptura	Fratura, brecha, trinca
		2.6	Fadiga	Recomenda-se usar quando for determinado que a causa da ruptura foi fadiga
		2.7	Sobreaquecimento	Danos no material devidos ao sobreaquecimento/queima
		2.8	Ruptura	Item rompido, estourado, explodido, implodido, etc.

Quadro 2.12– Mecanismos de falha (conclusão)

Fonte: adaptado da ISO 14224, Tabela B.2

Mecanismo de falha		Subdivisão do mecanismo de falha		Descrição do mecanismo de falha
Número	Notação	Número	Notação	
3	Falha de instrumentação	3.0	Geral	Falha relativa à instrumentação, sem nenhum outro detalhe conhecido
		3.1	Falha de controle	Controle falho ou desativado
		3.2	Sem sinal/ indicação/alarme	Sem sinal/indicação/alarme quando previsto
		3.3	Sinal/indicação/ alarme com falha	Sinal/indicação/alarme está errado em relação ao processo real. Pode ser espúrio, oscilante, intermitente, arbitrário
		3.4	Desajustado	Erro de calibração, variação de parâmetro
		3.5	Falha de <i>software</i>	Controle/monitoração/operação inexistente ou com falha devido a uma falha de <i>software</i>
		3.6	Falha de causa/ modo comum	Vários itens de instrumentos falharam simultaneamente, como, por exemplo, detectores de incêndio e gás redundantes
4	Falha elétrica	4.0	Geral	Falhas relativas ao fornecimento e transmissão de energia elétrica, sem nenhum outro detalhe conhecido
		4.1	Curto circuito	Curto circuito
		4.2	Circuito aberto	Desligamento, interrupção, fio/cabo partido
		4.3	Sem energia/tensão	Ausência ou insuficiência de alimentação elétrica
		4.4	Falha na tensão/energia	Alimentação elétrica com falha, como, por exemplo, sobretensão
		4.5	Falha no aterramento/ isolamento	Falta de aterramento, baixa resistência elétrica
5	Influência externa	5.0	Geral	Falha causada por algum evento externo ou substâncias de fora da fronteira e sem nenhum outro detalhe conhecido
		5.1	Bloqueio/ entupimento	Fluxo restrito/bloqueado devido a gelo, incrustações, contaminação, hidratos, etc.
		5.2	Contaminação	Fluido, gás ou superfície contaminado, como, por exemplo, óleo lubrificante ou cabeçote do detector de gás contaminado
		5.3	Diversas influências externas	Objetos estranhos, impactos, influência ambiental de sistemas vizinhos
6	Diversos	6.0	Geral	Mecanismo de falha que não se enquadra numa das categorias relacionadas acima
		6.1	Nenhuma causa encontrada	Falha investigada, mas causa não revelada ou muito incerta
		6.2	Causas combinadas	Várias causas: Se houver uma causa predominante convém que a mesma registrada
		6.3	Outros	Nenhum código aplicável: Usar texto livre
		6.4	Desconhecido	Nenhuma informação disponível

Quadro 2.13– Causas da falha
 Fonte: adaptado da ISO 14224, Tabela B.3

Causa da falha		Subdivisão da causa de falha		Descrição da causa da falha
Num.	Notação	Num.	Notação	
1	Causas relacionadas ao projeto	1.0	Geral	Projeto ou configuração de equipamento inadequada (formato, manutenibilidade, operabilidade, tecnologia, tamanho, etc.), mas sem maiores detalhes conhecidos
		1.1	Capacidade inadequada	Dimensionamento/capacidade inadequadas
		1.2	Material inadequado	Seleção de material inadequado
2	Causas relacionadas à fabricação / instalação	2.0	Geral	Falha relativa à fabricação ou instalação, mas sem maiores detalhes conhecidos
		2.1	Erro de fabricação	Falha de processamento ou fabricação
		2.2	Erro de instalação	Falha de instalação ou montagem (não inclui a montagem após a manutenção)
3	Falhas relacionadas à operação / manutenção	3.0	Geral	Falha relacionada à operação/uso ou manutenção do equipamento, mas sem maiores detalhes conhecidos
		3.1	Serviço fora das condições de projeto	Condições de serviços não intencionais ou fora das especificações de projeto, como, por exemplo, pressão maior do que a especificada
		3.2	Erro de operação	Erro, desatenção, negligência, uso indevido, etc. durante a operação
		3.3	Erro de manutenção	Erro, desatenção, negligência, uso indevido, etc. durante a manutenção
		3.4	Desgaste previsto	Falha causada pelo desgaste resultante da operação normal do equipamento
4	Falhas relacionadas ao gerenciamento	4.0	Geral	Falha relativa a questões de gerenciamento, mas sem maiores detalhes conhecidos
		4.1	Erro de documentação	Falha relativa a procedimentos, especificações, desenhos, relatórios, etc.
		4.2	Erro de gerenciamento	Falha relativa ao planejamento, organização, garantia da qualidade, etc.
5	Diversos	5.0	Diversos - geral	Causas que não se enquadram numa das categorias relacionadas acima
		5.1	Nenhuma causa encontrada	Falha investigada, mas nenhuma causa específica encontrada
		5.2	Causa comum	Causa/modo comum
		5.3	Causas combinadas	Várias causas agindo simultaneamente. Se possível, destacar a causa predominante
		5.4	Outros	Nenhum dos códigos acima se aplica. Especificar a causa como texto livre
		5.5	Desconhecido	Nenhuma informação disponível relacionada à causa da falha

Assim como para o mecanismo de falha, quando houver mais de um descritor de causa da falha aplicável, o responsável pela aquisição dos dados deve utilizar aquele que for o mais importante, e sempre tentar evitar os códigos 5.4 e 5.5.

É importante perceber que os parâmetros de confiabilidade e manutenibilidade monitorados referem-se a diferentes níveis taxonômicos, conforme os exemplos expostos no Quadro 2.14.

Quadro 2.14 – Parâmetros de confiabilidade e manutenção para os níveis da taxonomia
Fonte: Tabelas 3 da ISO 14224

Dados de CM registrados	Nível de Hierarquia				
	4 Planta/ Unidade	5 Seção/ Sistem	6 Unidade de Equipamento	7 Subunidade	8 Item Manutenível
Impacto da falha sobre a segurança	X				
Impacto da manutenção sobre a segurança	X				
Impacto da falha sobre operações	X	(X)			
Impacto da manutenção em relação a operações	X	(X)			
Impacto da falha sobre o equipamento			X	(X)	(X)
Modo de falha		(X)	X	(X)	(X)
Mecanismo de falha			(X)	(X)	X
Causa da falha				(X)	X
Método de detecção		(X)	X	(X)	(X)
Subunidade com falha				X	
Falha no item manutenível					X
Tempo de indisponibilidade	(X)	(X)	X		
Tempo de manutenção efetiva			X	(X)	(X)

X = Nível normalmente utilizado (X) = alternativas possíveis

A ISO 14224 destaca que o uso de códigos para registro dos eventos de falha facilita a realização de análises, mas recomenda que estes sejam sempre complementados com campos livres de texto para melhorar a interpretação de eventos individuais, tanto para análise da qualidade dos dados quanto para estudos detalhados desses eventos.

O registro de uma falha deve conter a data e hora da falha, o modo de falha, o método de detecção, a atividade de manutenção executada para sanar a falha, o mecanismo de falha e a causa raiz da falha. A relação completa dos dados solicitados pela ISO 14224 é apresentada no Quadro 2.15.

Quadro 2.15 – Dados solicitados para o registro de uma falha
 Fonte: Tabelas 8 e D.4 da ISO 14224

Categoria	Dados a serem gravados	Descrição
Identificação	Número do registro de falha *	Chave que Identifica um registro de falha
	Identificação/Tag do Equipamento *	Número de tag ou número de série
Dados de falha	Data da Falha *	Data em que a falha foi detectada
	Modo de Falha *	Modo como a falha foi observada no equipamento (ISO 14224, item B.2.6)
	Método de Detecção	Como a falha foi detectada (ISO 14224, tabela B.4)
	Condição de operação quando a falha foi detectada	Operando, aguardando, ligando, em teste, ocioso
	Impacto da falha na segurança da planta ^{1 2}	Nenhum, parcial ou total
	Impacto da falha nas operações da planta ²	Nenhum, parcial ou total
	Impacto da falha na função do equipamento ^{2 *}	Crítica, degradada ou incipiente ³
	Mecanismo de Falha	Processo (físico, químico, etc) que ocasionou a falha (ISO 14224, tabela B.2)
	Causa da Falha ⁴	As circunstâncias durante o projeto, fabricação ou uso que provocaram a falha (ISO 14224, tabela B.3)
	Subunidade em pane	Identificação da subunidade que falhou
Item(s) Manutenível(is) em pane	Identificação dos Itens Manuteníveis que falharam	
Observações	Informações Adicionais (texto livre)	Preencher com mais detalhes sobre as circunstâncias que levaram a ocorrência da falha: falha de unidades redundantes, detalhes sobre as causas da falha, etc.
	Informações Adicionais (outros campos)	Por exemplo, para vazamentos a identificação do tamanho do buraco e volume que vazou é importante para análises como a análise quantitativa de risco Outro exemplo é um campo para estabelecer uma ligação entre um banco de dados de acidentes e o de confiabilidade e manutenibilidade

¹ Pode ser subdividida em pessoal, meio ambiente e ativos

² Classe de severidade

³ Para alguns casos é suficiente identificar se a falha é crítica ou não

⁴ A causa e mecanismo de falha não são conhecidos quando os dados são coletados, pois necessitam de uma análise de causa raiz para serem descobertos. Essa análise deve ser efetuada para falhas de grande consequência, grande tempo de reparo ou de indisponibilidade, ou falhas que estejam ocorrendo mais freqüentemente do que o considerado normal para aquela classe de equipamento (*worst actors*)

* Conjunto mínimo de dados exigidos pela ISO 14224

Enquanto o Quadro 2.5 estabelece a relevância das informações sobre as características técnicas dos equipamentos a serem monitorados, os Quadros 2.6, 2.7, 2.9 e 2.15 apresentam o conjunto mínimo de dados a serem adquiridos e controlados para os objetivos da norma ISO 14224. Entretanto, de acordo com os usos pretendidos, o número de dados considerados como obrigatórios pode ser maior.

Isso pode ser visto no anexo D (*Typical requirements for data*) da ISO 14224, que apresenta diferentes tipos de análises que podem ser desenvolvidas com base nos dados de falha e manutenção coletados, e qual o conjunto mínimo de informações requeridas. Para este trabalho, serão considerados como obrigatórios apenas os dados mínimos solicitados pela ISO 14224.

2.3 A Informação e o Conhecimento no Contexto da Gestão e Engenharia do Conhecimento

A cognição humana decorre da capacidade desenvolvida por homens e mulheres para a criação ou composição de representações mentais e processos imaginativos, partindo da memória de sensações, sentimentos e idéias. Essas criações ou composições são provocadas por perturbações internas que, em parte, decorrem diretamente dos estímulos recebidos do ambiente no qual os seres humanos estão inseridos.

O conhecimento humano é produto das operações mentais realizadas como testes de correspondência, entre as previsões propostas por modelos mentais, como representações e processos imaginativos, e as respostas ou novas percepções recebidas pelos sentidos, diante da observação e da ação de homens e mulheres no mundo.

As diversas formas de expressão e linguagem permitiram aos homens e mulheres o compartilhamento de conhecimentos, compondo um amplo acervo de conhecimentos comuns que é denominado “cultura”. A partir dessa base comum, novos conhecimentos podem e são construídos.

Há milênios, a busca de compreensão sobre as habilidades de comunicação e de construção de conhecimentos tem ocupado muitos estudiosos. Durante esse tempo, diversas abordagens teóricas foram desenvolvidas. Alguns termos como: “sinal”, “signo”, “dados”, “mensagem”, “informação” e “conhecimento” foram aplicados em diferentes contextos e assumiram diversas significações, sendo que essa diversidade não é constantemente percebida e devidamente considerada.

A indefinição apresentada acima é parcialmente justificada, pois, de um ponto de vista estrutural, “signo”, “informação”, “conhecimento”, entre outros, quando expressos através de uma linguagem, podem ser considerados sinônimos, porque são compostos por um elemento percebido pelos sentidos ou pela mente (significante) e associado a uma ou mais idéias ou lembranças (significado). A necessidade de diferenciação, entretanto, foi historicamente determinada por relações de valor e considerações sobre a verdade do conhecimento. Toda associação gera um conhecimento, mas nem todo conhecimento é percebido como verdadeiro ou valoroso. Assim, o termo “conhecimento”, na maior parte das vezes, é indicado para designar uma associação considerada verdadeira, enquanto uma associação improvável ou um conhecimento falso são designados como “mito” ou “falácia”, entre outros termos possíveis.

Essas questões histórico-conceituais são problemáticas quando esses termos são usados para designar novos campos de estudos e atividades, como “Engenharia do Conhecimento”, que

procura modelar processos e comportamentos, e “Gestão do Conhecimento”, que busca soluções sistêmicas para gerir processos de conhecimento.

As imprecisões filosófica, ontológica e epistemológica, acerca do termo “conhecimento” incitam questionamentos sobre a Engenharia e a Gestão do Conhecimento, colocando em dúvida a pertinência dessas denominações e, conseqüentemente, a possibilidade de existência desses campos de estudos e atividades.

Este texto visa apresentar uma proposta conceitual sobre o conhecimento no contexto da Gestão e da Engenharia do Conhecimento, recuperando teorias e conceituações sobre informação, comunicação, engenharia e gestão.

2.3.1 Abordagem hierárquica DIKW

Há autores, sendo alguns desses citados a seguir, que se baseiam no conceito de conhecimento expresso por uma relação hierárquica entre dados (*Data*), informação (*Information*), conhecimento (*Knowledge*) e sabedoria (*Wisdom*), que é conhecida como DIKW, hierarquia da informação, hierarquia do conhecimento ou pirâmide do conhecimento.

A idéia de uma hierarquia entre esses termos já havia sido esboçada nas artes, em um poema de 1934 de Thomas Stearns Eliot (1888—1965) intitulado “*The Rock*” e na música de Frank Zappa (1940-1993), “*Packard Grosse*”, do álbum “*Joe’s Garage: Act II&III*” de 1979, e foi consolidada por meio de trabalhos como “*Notes on the information concept*” (BOULDING, 1955), “*Information as a Resource*” (CLEVELAND, 1982), “*A Guidebook to Learning: for a Lifelong Pursuit of Wisdom*” (ADLER, 1986), “*Management support systems: towards integrated knowledge management*” (ZELENY, 1987), “*Architect or Bee: The Human Price of Technology*” (COOLEY, 1987) e “*From Data to Wisdom*” (ACKOFF, 1989) (ROWLEY, 2007; WALLACE, 2007; FRICKÉ, 2008; SHARMA, 2008).

O artigo de Ackoff (1989), “*From data to wisdom*”, muitas vezes, é citado como a origem da DIKW. Apesar de também ter incluído o conceito de “compreensão” situado entre “conhecimento” e “sabedoria”, outros autores, como Bellinger *et al.* (2004), contestaram sua pertinência como um nível distinto, entendendo tratar-se, na verdade, do processo que permite proceder conversões ou transitar entre os demais níveis. Os conceitos de “sinais” e de “esclarecimento” ou “iluminação”, propostos por outros autores, também não foram incorporados a DIKW. Rowley (2007) apresenta as seguintes conceituações propostas por Ackoff (1989):

- Dados são símbolos que representam propriedades de objetos, eventos e seus ambientes. São produtos de observação e sua utilidade depende da forma em que se encontram. A sua diferença para informação não é estrutural e sim funcional.
- Informação está contida em descrições, respostas a perguntas que começam com palavras como quem, o que, quando e quantos. Sistemas de Informação geram, armazenam, recuperam e processam dados. As informações são inferidas a partir dos dados.
- Conhecimento é Know-how, que torna possível a transformação de informações em instruções. O conhecimento pode ser obtido tanto pela transmissão a partir de quem o possui, pelo ensino, ou extraído com base na experiência.
- Sabedoria é a habilidade de incrementar a eficácia, porque a eficiência é incrementada pela inteligência. Conhecimento adiciona valor, o que demanda a função mental que chamamos de julgamento. Os valores éticos e estéticos em que isso implica são inerentes ao ator e são únicos e pessoais.

O conceito de sabedoria foi pouco desenvolvido em comparação com os demais conceitos aqui apresentados. Além disso, esse termo envolve questões subjetivas e intuitivas que não pertencem ao campo de interesse deste trabalho. Pois, aqui são considerados prioritariamente os seguintes termos: “dado”, “informação” e “conhecimento”.

2.3.2 Deficiências da DIKW

Rowley (2007) fez um estudo sobre como esses termos foram tratados por diversos autores, apresentando os acréscimos e processos de transformação propostos, e sugeriu referenciá-los como hierarquia da sabedoria, cuja representação seria a de um funil ou pirâmide invertida.

Essa visão sugere que a sabedoria é obtida após o processamento de dados, informação e conhecimento, indicando que esse processo começa com os dados. Todavia, a visão de Rowley (2007) é criticada por Frické (2008), considerando os seguintes pontos:

- A DIKW é embasada em uma visão operacionalista ou positivista e não interpretativa. Assim, os dados têm que ser objetivos, positivos e verdadeiros. As inferências possíveis devem ser diretamente associadas aos fatos positivos e passíveis de serem reduzidas aos dados, para serem consideradas informações válidas. Isso invalida todos os conceitos subjetivos que não podem ser objetivamente confirmados, relegando-os, como informações sem sentido ou sem valor. Isso não prejudica as deduções

diretamente lógicas, mas inviabiliza os novos conhecimentos, gerados por inferências indutivas ou abduativas que, partindo de dados válidos, geram teorias hipotéticas, que não são necessariamente verdadeiras.

- A visão operacionalista ou positivista limita o conhecimento possível (*Know-how*) às deduções ou inferências lógicas, que são extraídas diretamente dos dados da realidade. Assim, dados, informações e conhecimentos são diretamente interligados.
- Ackoff (1989) considera o conhecimento como *know-how* expresso por meio de regras. Frické (2008) aponta que, mesmo implicitamente, as regras só podem ser expressas na forma declarativa (*know-that*). Por isso, adotando uma visão interpretativa, Frické assinala que os dados são informações, mas assevera que nem todas as informações são dados ou são obtidas diretamente dos dados, destacando que informação e conhecimento são sinônimos.
- A análise de Frické (2008) aponta que um dos problemas da DIKW é considerar que somente a partir dos dados factuais é possível atingir os níveis mais altos da hierarquia e que os cientistas da informação não querem iniciar uma coleta de dados apenas com a esperança que ao final do processo alguns sejam promovidos a informação.
- Frické (2008) indica a dedução como princípio de um estudo que deveria começar com uma boa “teoria de questões” para identificar as informações requeridas e a partir delas os dados necessários. Ao contrário da indução proposta na DIKW, esse autor reconhece que os dados não podem ser o ponto de partida para o conhecimento. Assim, a descrição dos dados não é isenta ou positiva, porque decorre de uma forma padronizada de representação, que foi construída e utilizada socialmente com base em um conhecimento pré-existente. O conceito fundamental é que o conhecimento produz conhecimento e os dados por si mesmos não geram informação.

2.3.3 Revisão Epistemológica

Para esta revisão, adota-se o pensamento complexo de Morin (1990), que propõe “trabalhar com um hiper-paradigma da complexidade, percebendo e concebendo os fenômenos estudados como complexos, para que não sejam mutilados pela simplificação”. É necessário atentar para as “ilusões” que desviam o espírito do problema do pensamento complexo: crer que a complexidade conduz à eliminação da simplicidade e confundir complexidade com completude, quando o objetivo consiste apenas em obter um conhecimento multidimensional.

Outros referenciais importantes são as idéias de Ludwig Wittgenstein (1889-1951), indicando que os significados das palavras são construídos socialmente, e de Michael Polanyi (1891-1976), reconhecendo que todo o conhecimento consiste de ou tem suas raízes em atos de compreensão, e que as pessoas conhecem mais do que são capazes de explicitar.

De acordo com as idéias de Polanyi (1983), alguns conhecimentos, ou partes do conhecimento, apenas se manifestam quando surge uma situação na qual sejam necessários. Além disso, não são todos os conhecimentos que podem ser expressos por meio de palavras.

Essa característica do conhecimento como algo que capacita à ação, demandando interação para ser reconhecido, é destacada por diversos autores, mas ainda não existe um consenso sobre quais são as fronteiras do conceito de conhecimento em relação a esse processo de entendimento e capacidade de solução de um problema. De qualquer maneira, há duas questões básicas envolvidas. A primeira indica que o conhecimento não tem expressão própria, necessitando de elementos expressivos ou atitudes para ser percebido. A segunda diz respeito ao valor do conhecimento, indicando que um conhecimento valioso é aquele que se apresenta como solução. Mas, como foi dito anteriormente, do ponto de vista estrutural um conhecimento propõe associações e não necessariamente soluções.

Há alguns pontos intuitivos de concordância. Para conhecer, é necessário saber associar, para associar, é preciso lembrar, e para lembrar é necessária alguma forma de representação interior da coisa ausente. As representações permitem o reconhecimento de algo e, também, relaciona suas formas de interação conhecidas.

Segundo Maslow (1908-1970), as interações são motivadas pela busca da satisfação de alguma necessidade, e nosso instrumento de interação é o próprio corpo (MASLOW, 1943). O primeiro aprendizado está relacionado ao corpo e, a partir dele, os demais conhecimentos são construídos. Para satisfazer suas necessidades, a mente experimenta e seleciona ações que lhe pareçam mais eficazes, sendo que este é um dos processos de aprendizagem utilizado para a construção de conhecimentos.

De acordo com esta linha de raciocínio, que encontra apoio tanto na Cibernética quanto no Construtivismo, todo o conhecimento é edificado com base nos conhecimentos já possuídos, e essa idéia de posse induz pensar o conhecimento como um objeto, aproximando-o do conceito de informação. Isso reforça a posição de Frické (2008), que indica o conhecimento como decorrente de deduções a partir de teorias anteriores e iguala informação a conhecimento.

O conhecimento só é perceptivo e eficiente quando é denotado por elementos expressivos e se torna atuante por meio de atitudes ou de ações. Por isso é importante não perder de vista a

característica dinâmica do conhecimento, que capacita à ação ou realização de tarefa, o que indica que ele deve possuir tanto características de objeto quanto de processo.

O conhecimento é algo que está embutido no conjunto das estruturas de representações de um ser. As atitudes do ser permitem que o conhecimento seja reconhecido e se mostre atuante e interagente com elementos do meio ao qual está exposto. A capacidade de trabalhar com essas representações permite aos seres projetar os resultados que podem ser alcançados com o repertório de ações conhecido. Assim, é possível ao ser escolher as ações que serão executadas para a satisfação de uma necessidade ou de uma meta.

No âmbito dos estudos da Engenharia e Gestão do Conhecimento, é importante considerar os processos de construção, disseminação, utilização e armazenamento de conhecimentos. A construção e a utilização de conhecimentos são fundamentais para o entendimento do conhecimento no contexto da Engenharia do Conhecimento, enquanto que a disseminação e o armazenamento de conhecimentos são, muitas vezes, confundidos com o conceito de informação. Pode-se referenciar como conhecimento, o resultado do acúmulo de experiências de interações entre objetos mentais ou representações e objetos reais, bem como entre as próprias representações mentais.

Dessa forma, o conhecimento só pode existir em sujeitos ou entidades que sejam capazes de manipular representações e possuam uma estrutura cognitiva mínima apta a reconhecer estímulos e reagir de forma coerente com sua base de conhecimentos. Doravante, esses seres, sujeitos ou entidades, serão denominados de “entes cognoscentes”.

De acordo com a visão clássica ocidental, a cognição ou o domínio cognitivo é uma das três partes que compõem uma mente. As outras duas partes da mente são a emoção ou o domínio afetivo e a volição ou o domínio conativo. O domínio cognitivo (completo) abrange crenças, aquisição e revisão de crenças, aprendizado, conhecimento e raciocínio.

No livro “*Knowledge Management System*” (MAIER, 2007), o conhecimento é assinalado como algo que abrange todas as expectativas cognitivas, considerando as observações que, de modo significativo, foram organizadas, acumuladas e embutidas em um contexto. As experiências, a comunicação e as inferências são instâncias que um ente cognoscente na posição de ator organizacional usa para interpretar situações e gerar atividades, comportamentos e soluções, não importando se essas expectativas são racionais ou usadas intencionalmente.

Uma das conseqüências dos argumentos apresentados é que o conhecimento só pode existir em um conhecedor, um ente cognoscente humano ou não. Além disso, para que esse

conhecimento aconteça de modo objetivo é necessário que o ente ou ator possua uma capacidade de cognição mínima para aplicar esse conhecimento.

Outra consequência é que a construção de conhecimentos implica na construção de modelos e representações internos ou externos, mas conforme McInerney (2002) adverte, o conhecimento não deve ser confundido com suas representações.

Um dos principais recursos utilizados tanto para a representação quanto para a disseminação do conhecimento são as palavras, construídas socialmente através dos processos de comunicação. Mas a comunicação não se restringe a troca de palavras.

A capacidade de comunicação de um indivíduo aumenta conforme ele desenvolve novas formas de expressão. Para isso são usados diferentes sinais que estimulam outros sentidos, como a visão e o tato. Por exemplo, a ordenação da língua foi associada a sinais corporais para compor a Linguagem Brasileira de Sinais (LIBRAS).

Os sinais sonoros, visuais e táteis, entre outros, quando são associados a idéias, pensamentos ou lembranças, adquirem significação ou significado, passando a ser identificados como “informação”. Uma porção de substância é informada quando expressa um pensamento ou mais. Em seu sentido original, “informar” significa dar forma a alguma coisa. Assim, um pouco da substância barro pode ser organizado ou informado para expressar um vaso. Neste caso, o vaso é barro informado que passa a atuar como informação, porque comunica aos espectadores a idéia de vaso.

A idéia de que o conhecimento é construído por informações, como indica a DIKW, é contrariada na medida em que a informação depende do conhecimento e só pode ser comunicada entre entidades que, antecipadamente, detém conhecimentos sobre o código e, pelo menos, sobre parte do repertório utilizado.

O conceito de informação é ligado à comunicação, indicando que uma mídia é organizada ou informada para comunicar uma idéia, como no exemplo do barro organizado para comunicar a idéia de vaso. O campo semiótico estuda a significação, cujo conceito é necessário à comunicação. A unidade de significação é o signo que, de acordo com Ferdinand de Saussure (1857-1913), reúne um significado (idéia) a um significante (substância) que, no caso da comunicação, necessita ser perceptível (SAUSSURE, 2008). No exemplo adotado, o barro modelado é o significante e a idéia de vaso é o significado. Assim, cada sinal capaz de comunicar um significado é chamado de signo.

Por sua vez, Charles S. Peirce (1839-1914) conceitua a parte representativa do signo como “representamen”, que é a própria expressão do signo, “aquilo que, sob certo aspecto ou modo, representa algo para alguém”. Para esse autor, representar é “estar no lugar de, isto é, estar

numa tal relação com um outro que, para certos propósitos, é considerado por alguma mente como se fosse este outro” (PEIRCE, 1977, p. 26). Para se estabelecer uma comunicação é necessário, portanto, um processo de significação, o qual se estabelece com o uso de um conjunto sinais representativos ou signos, que são organizados de acordo com regras sintáticas e semânticas.

Do ponto de vista técnico, considerando especialmente a codificação e a transmissão de sinais por meio de canais tecnológicos, há o trabalho de Claude E. Shannon (1916-2001) “*A Mathematical Theory of Communication*” (1948), o qual ficou conhecido como “Teoria da Informação”. O modelo proposto por Shannon (Fig. 2.3) descreve cinco elementos essenciais ao processo de comunicação: 1- fonte, 2- transmissor ou codificador, 3- canal, 4- receptor ou decodificador e 5- destinatário.

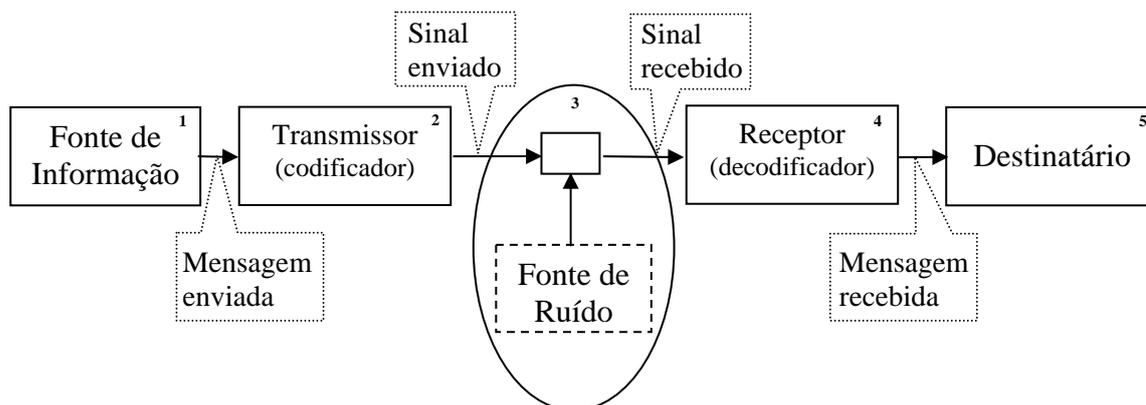


Figura 2.3 – Diagrama esquemático de um sistema de comunicação

Fonte: SHANNON (1948).

O diagrama acima (Fig. 2.3) indica que o conjunto de sinais enviado pelo canal é emitido e recebido como “mensagem”. Nesse sentido, pode-se entender mensagem como informação enviada por algum canal. Do ponto de vista estrutural, segundo o Dicionário Aurélio (FERREIRA, 1999), mensagem é sinônimo de “comunicação, notícia ou recado verbal ou escrito”.

A Teoria da Informação de Shannon prioriza a transmissão de sinais ou mensagens, porque não considera, em princípio, a questão semântica. Essa ênfase nas relações quantitativas desvincula os termos informação e mensagem do conteúdo semântico estabelecido no âmbito do destinatário. Entretanto, a informação ou a mensagem disposta no canal deve ser aceita pelas mídias envolvidas, de acordo com as codificações específicas do sistema. A codificação e o conhecimento previstos no nível técnico dos sistemas são baseados em relações de

quantidade. Por isso, os parâmetros de definição da informação na Teoria de Shannon são quantitativos, mesmo com relação às possibilidades de um repertório de signos pré-determinados.

O problema tratado por Shannon é o grau de sucesso da recuperação de uma mensagem gerada por uma fonte de informação para um destinatário e que foi reestruturada por um codificador, em um formato adequado à transmissão, em função das características do canal utilizado.

É importante destacar que, embora uma mensagem tenha a capacidade de transmitir palavras organizadas em frases e orações a partir do ponto de vista do emissor, o sucesso da comunicação implica na transmissão do significado previsto para as frases e, para que isso ocorra, é necessário que o destinatário da mensagem tenha a capacidade de interpretá-la de acordo com as intenções do emissor.

Do ponto de vista da interpretação do destinatário humano, a informação é a idéia recebida após a decodificação e a decorrente reconstrução do significante de uma mensagem tecnicamente codificada. Depois de decodificada, uma mesma mensagem pode conter informação em três níveis. O primeiro é o nível sintático ou pré-denotativo, tratando da organização físico-espacial da mensagem que interfere na significação ou no nível semântico. O segundo é o nível lingüístico ou denotativo, no qual os sinais ou significantes são percebidos como palavras ou como equivalentes a palavras que denotam coisas. O terceiro é o nível pós-denotativo ou conotativo, no qual as palavras ou significantes equivalentes são relacionados para significar idéias que estão além do significado denotativo.

De um ponto de vista quantitativo, tendo em vista a informação técnica, a correta recuperação ou decodificação de uma mensagem não pode negligenciar a ordenação dos sinais. Por outro lado, de um ponto de vista qualitativo ou interpretativo, a correta recuperação ou decodificação da mesma mensagem irá depender da sensibilidade e do conhecimento do ente cognitivo que a recebe.

Ao aceitar que mensagem é informação disposta em um canal de comunicação, é possível considerar que, para ser mensagem, uma informação deve ser tecnicamente compatível com o canal de transmissão. Por outro lado, uma mensagem para ser informação, deve ser significativa para o ente cognoscente que a recebe. Assim, a elaboração de uma mensagem deve levar em conta as características do canal e como informação enviada, deve ser interpretativamente adequada aos destinatários pretendidos.

De modo prático, o que viabiliza a mensagem são as regras de sintaxe, enquanto a informação é viabilizada pela propriedade semântica. Diante do exposto, as mensagens significativas são

aquelas que apresentam informação coerente e valiosa para o ente cognitivo receptor. O valor da informação varia de acordo com a análise qualitativa que é desenvolvida. Mais do que isso, segundo Capurro e Hjørland (2007), algumas teorias chegam a considerar como informação apenas as que atendem a requisitos, como: valor e utilidade; verdade ou correção; novidade; surpresa; redução da incerteza; etc. Esses autores, em seu artigo “O conceito de informação”, destacam que uma das dificuldades existentes para definir o termo informação advém do fato de ser um conceito polimórfico e fazem um apanhado de diversos significados a ele atribuído de acordo com sua utilização em vários campos diferentes.

Conforme a Teoria do Significado de Wittgenstein (1996), o significado de um termo ou expressão deve ser inferido de acordo com seu uso, ou melhor, com seu valor de uso. Sob essa perspectiva, a informação de valor para Engenharia e Gestão do Conhecimento é aquela que contribui para o bom funcionamento e desenvolvimento de sistemas e organizações. Entretanto, isso não é uma definição do termo e sim uma delimitação do escopo que deve ser refletida por uma qualificação ou classificação apropriada.

Para encontrar a resposta desejada, é importante analisar como a informação se comporta nos processos de armazenamento e de disseminação.

Quanto ao armazenamento, alguns autores consideram que a informação, assim como o conhecimento, só pode existir em entes cognoscentes, o que é contestado por Wittgenstein uma vez que a construção das palavras e de seus significados é um processo coletivo. As palavras, os códigos e, portanto, as significações ou as informações, apesar de serem construídas por entes cognoscentes, têm existência própria no contexto cultural, que é configurado pela mente coletiva do conjunto de entes cognoscentes. As suas características significantes existem e persistem quando registradas em algum meio ou mídia que também são produtos culturais.

Há a permanência semântica na mídia, que considera os aspectos semânticos no contexto da cultura dos emissores e dos destinatários humanos, mas fora das mentes individuais dos entes cognoscentes. Todavia, é interessante considerar também a idéia de uma semântica tecnológica, que se estabelece no próprio contexto de codificação da informação presente na mídia. Isso permite que a mídia, por meio de um processo de interpretação tecnológica, reconheça, armazene e transmita a informação mantendo sua estruturação sintática a qual mantém a ordenação da mensagem para posterior decodificação e interpretação semântica por parte de um sujeito cognoscente humano.

O registro de informações permite que mídias perenes possam exercer parcialmente o papel de fonte e transmissor de informações, de forma que o processo de comunicação possa ocorrer

sem a presença da fonte de informação original. Dessa maneira, a comunicação deixa de ser apenas imediata e passa a ser mediada pelos meios de registro utilizados. De acordo com Maier (2007), isso permite que a comunicação seja uma das vias de construção do conhecimento.

De um ponto de vista estrutural, informação é conhecimento expresso em um sinal ou substância. Por isso, muitos autores usam o termo “conhecimento explícito”, para designar as informações que participam em atividades de conhecimento, capacitando um ente cognoscente a entender ou agir em algum contexto. Neste caso, informação e conhecimento são sinônimos, impedindo a distinção entre as áreas de Gestão da Informação e de Gestão do Conhecimento. Entretanto, é preciso considerar o valor da informação e o valor do conhecimento, para compreender que a Gestão do Conhecimento parte da Gestão da Informação, mas deve ir mais além e propor sistemas mentais, sociais ou materiais capazes de selecionar, organizar, desenvolver e dispor as informações e os conhecimentos de valor de acordo com as demandas de uma organização ou cultura institucional. De outra forma, para as áreas de conhecimento, seu objeto de estudo envolve os processos interativos de construção, disseminação e utilização de conhecimentos, os quais normalmente são embasados em processos de armazenamento e recuperação de informações.

Por fim, para analisar o termo dado é necessário resgatar seu sentido original, de algo entregue ou fornecido. Para Machlup e Mansfield (1983), “dado” é uma coisa fornecida para um analista, investigador ou solucionador de problemas – pode ser números, palavras, sentenças, registros, hipóteses, enfim, qualquer coisa não importando sua forma ou origem. Portanto, o termo “dado” designa os estímulos ou sinais percebidos por um ente cognoscente, humano ou não. Os dados estão no mundo e Ackoff (1989) assinala que a diferença entre dados e informação não é estrutural, mas funcional. Pois, estruturalmente, ambos são sinais significantes, contudo, nem todo dado é considerado como informação de valor.

2.3.4 Implicações Relevantes

Em áreas como Ciência da Informação, Gestão da Informação, Engenharia da Informação ou Gestão do Conhecimento e Engenharia do Conhecimento entre outras, os termos “informação” e “conhecimento”, são cruciais para definição de seus objetos de pesquisa. Porém, apesar da coincidência estrutural entre aquilo que é designado por ambos os termos, a prática nessas áreas distingue a informação do conhecimento. Entretanto, isso muitas vezes é feito utilizando apenas o critério de valor. Toda informação é produto de um processo de

conhecimento registrado ou expresso em uma mídia, mas nem todo conhecimento é valioso do ponto de vista dos entes cognoscentes. Por isso, o conceito de “conhecimento” na Engenharia e Gestão do Conhecimento também assinala o valor do conhecimento, sendo a informação em geral identificada como “dado”. Portanto, do ponto de vista funcional, “dado”, “informação” e “conhecimento” são distintos de acordo com seu valor diante dos entes cognoscentes. Essa distinção entre informação e conhecimento com base apenas no valor foi duramente criticada por Wilson (2002) em seu artigo *“The nonsense of 'knowledge management’”*, no qual identifica que parte do que hoje é chamado de Gestão do Conhecimento, trata-se na verdade do resultado de uma estratégia de *search and replace marketing*, onde o termo informação simplesmente foi substituído por conhecimento.

A necessidade de parâmetros precisos para a distinção desses termos, de acordo com sua função, pode ser percebida em trabalhos como: *“Can Information Modelling be Successful without a Common Perception of the Term ‘Information’?”* (LEHNER; MAIER, 1997), *“Data Is More Than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory”* (TUOMI, 1999a, 1999b), *“When knowledge becomes information: a case of mistaken identity”* (KAY; CECEZ-KECMANOVIC, 2000), *“Knowledge Management: a new idea or a recycled concept?”* (SPIEGLER, 2000), *“Some Observations on the Semantics of ‘Information’”* (NEWMAN, 2001), *“I = 0 : Information has no intrinsic meaning”* (MILLER, 2002), *“Informação e Conhecimento: aspectos filosóficos e informacionais”* (SIRIHAL; LOURENÇO, 2002), *“The nonsense of 'knowledge management’”* (WILSON, 2002), *“It’s tacit knowledge but not as we know it: redirecting the search for knowledge”* (CONNELL *et al.*, 2003), *“Semantic Conceptions of Information”* (FLORIDI, 2005), *“O conceito de informação”* (CAPURRO; HJORLAND, 2007) e *“The Knowledge pyramid: a critique of the DIKW hierarchy”* (FRICKÉ, 2008).

Mais do que a revisão de conceitos, é necessário o entendimento de como trabalhar, aplicando em informações e conhecimentos de valor, tratamentos adequados aos novos paradigmas que orientam o desenvolvimento das áreas que têm a informação e o conhecimento como objeto de interesse.

Esse processo já está em andamento e um de seus frutos é a nova Engenharia do Conhecimento, na qual o velho paradigma de mineração foi substituído pelo novo paradigma de modelagem, cujas mudanças são bem exploradas em Motta (2001).

É necessário, ainda, desenvolver uma nova terminologia para comunicar o escopo, metodologias, ferramentas e resultados que cada área se propõe a alcançar, para que assim seja possível distinguir e compreender os trabalhos elaborados.

Para construir novos termos e significados, não é necessário abandonar a ampla literatura disponível, mas sua utilização demanda um trabalho de re-enquadramento dos termos “informação” e “conhecimento”, através de um processo de contextualização, considerando a hipótese epistemológica, os aspectos e aplicações do conhecimento e porque ele é abordado nas obras consultadas (VENZIN *et al.*, 1999).

Por exemplo, Nonaka e Takeuchi (1997) desenvolveram sua obra reportando a existência de dois tipos de conhecimento. Um dos tipos é o “conhecimento tácito”, que é interiorizado, construído com base na experiência, difícil de ser articulado em palavras e, por consequência, difícil de ser transmitido. O outro tipo é o “conhecimento explícito”, que é fácil de articular e registrar, permitindo sua transmissão a terceiros. A interação dos dois tipos de conhecimento resulta no “Processo SECI”, representando quatro modos de conversão do conhecimento: 1- socialização, 2- externalização, 3- combinação e 4- interiorização.

A “socialização” prevê a transmissão do conhecimento tácito, fazendo a conversão de tácito para tácito, por meio de processos de compartilhamento de experiências entre os indivíduos de um grupo. Frequentemente, essa transmissão ocorre por meio da observação, da imitação e da prática.

A “externalização” estabelece a conversão de tácito para explícito, promovendo o processo de organização do conhecimento tácito em conhecimento explícito, por meio de metáforas, analogias, conceitos, hipóteses e modelos.

A “combinação” caracteriza a conversão de explícito para explícito, possibilitando os processos de conversão com a junção de conjuntos diferentes de conhecimentos, já explicitados, nos quais as pessoas utilizam meios como documentos, telefone, redes de computadores, conversas e reuniões para combinar conhecimentos diferentes.

A “internalização” promove a conversão de explícito para tácito, por meio de processos de incorporação do conhecimento explícito ao conhecimento tácito.

O foco principal da obra de Nonaka e Takeuchi (1997) é a gestão da dinâmica de criação de conhecimentos relevantes ou de valor para as empresas. Assim, o “conhecimento tácito” é um conhecimento relevante, que ainda não foi devidamente codificado e explicitado em mídia convencional, contendo informação implícita em comportamentos e atitudes, cuja prioridade não é a comunicação. Por sua vez, o “conhecimento explícito” é apresentado como informação de valor bem codificada e comunicada em mídia convencional. Parte da dificuldade de explicitação dos conhecimentos tácitos reside na sua pouca informação ou codificação, que é característica da comunicação estética, a qual é vivenciada, mas não é devidamente interpretada. O conhecimento explícito é apresentado como informação

altamente codificada, estabelecendo plena comunicação semântica que, além de ser vivenciada, é plenamente interpretada.

A transmissão de informações e conhecimentos no âmbito das organizações, além de considerar a comunicação tácita e explícita, deve considerar também os processos formais e informais de comunicação e educação.

É importante ressaltar que os processos de conversão de conhecimento descritos por Nonaka e Takeuchi (1997) são mais complexos do que o apresentado em seu trabalho. Bandura (1977) e Wenger (1999) tratam da “socialização” ou “processo de aprendizagem social”, enfatizando o caráter “informal” do compartilhamento de experiências entre indivíduos componentes de um grupo. Isso é efetivado por meio de comunicação informal decorrente da observação, da imitação e da prática conjunta. A “externalização” é indicada como “produção de informações” relevantes ao negócio, a partir do conhecimento tácito informado em metáforas, analogias, conceitos, hipóteses e modelos. A “combinação” também é indicada como “produção de informações”, decorrente do processamento de outras informações, por contextualização e inter-relacionamento. A “internalização” ou “processo de aprendizagem individual” ocorre a partir de informações normalmente chamadas de conhecimento explícito, com extensa base teórica desenvolvida.

A abordagem de Nonaka e Takeuchi (1997) é considerada simplista por ignorar que a correta interpretação da informação demanda um conhecimento tácito. Além disso, desconsidera que a “transmissão” de conhecimentos é realizada por meio de informações. Wallace (2007) analisa as questões epistemológicas inconsistentes na abordagem teórica de Nonaka e Takeuchi (1997).

Os conhecimentos são expressos em informações e, do ponto de vista estrutural, há uma identidade entre o que é percebido como “conhecimento” e como “informação”. Porém, há diferenças com relação à natureza e à função de cada um dos objetos designados pelos termos “conhecimento” e “informação”.

Nonaka e Takeuchi (1997) não apresentam definições formais para os termos “conhecimento tácito” e “conhecimento explícito” e sua abordagem contradiz as idéias de Polanyi¹ (1983), especialmente, quando esse afirma que há uma dimensão do conhecimento que não pode ser expressa, e que conhecimentos que possuem grande parte situada na dimensão tácita não podem ser comunicados diretamente. Portanto, há conhecimentos que não se apresentam como informações comunicáveis, indicando uma diferença entre conhecimento e informação que vai além da questão do valor.

¹ Wallace (2007) destaca que Nonaka e Takeuchi (1997) não incluíram o trabalho de Polanyi na discussão sobre o conhecimento por entenderem tratar-se de um autor menor na filosofia ocidental.

A crítica ao Processo SECI assinala uma tendência de desviar o foco do conhecimento para o que poderia ser informação, simplificando os processos de construção do conhecimento, que envolvem mais que a troca de informações e de recursos (MEIRINHOS, 2006). Para Wenger (1999), a aprendizagem é multidimensional e diversas teorias a abordam por diferentes focos. Por fim, vale analisar como um sistema informatizado é capaz de possuir conhecimento. Para isso, é necessário que ele possua uma cognição mínima que lhe permita manifestar esse conhecimento. Para Lanz e McFarland (2000), isso demanda uma base de conhecimento, com uma memória simbólica que armazena informações, e também demanda um mecanismo de ação que, normalmente, é descrito como *know-how*. A base de conhecimento e o mecanismo de ação são utilizados pelo sistema para atingir alguma meta conhecida.

Lanz e McFarland (2000) destacam que, nesse sentido, ter *know-how* equivale a possuir procedimentos relevantes para a realização de alguma tarefa e que a base de conhecimento pode ser tão simples como um número, resultado de algum processo depositado em uma memória, de modo que esteja disponível para outros processos do sistema.

Os mesmo autores, entretanto, advertem que é difícil para um observador julgar se um sistema é capaz de operar com uma base de conhecimento. Seria como saber diferenciar entre um sistema de atingimento de metas, um sistema perseguidor de metas e um sistema direcionado a metas. No primeiro, a meta só é reconhecida pelo sistema após ser alcançada e, portanto, suas ações não são direcionadas a este fim. No segundo, o sistema persegue uma meta que, no entanto, não está representada no sistema. No terceiro, o sistema possui uma representação da meta a ser atingida e suas ações são direcionadas a este fim. Em todos os casos, o projetista usa seu conhecimento do meio e do sistema para prover um comportamento ao sistema que o conduza à realização de uma meta. Por isso, a diferença entre sistemas é distinta somente para o projetista dos sistemas, apesar de haver a possibilidade de que sejam criados critérios de distinção para um observador externo.

2.3.5 O Conhecimento do “Nível do Conhecimento”

A visão de conhecimento apresentada neste trabalho está alinhada com a que foi expressa por Newell (1981), em seu artigo “*The Knowledge Level*” que, apesar de ter contribuído para o surgimento da segunda geração da Engenharia do Conhecimento, quando houve a mudança do paradigma da extração para a modelagem do conhecimento, (ainda) não foi capaz de provocar o surgimento de uma nova terminologia para esta área.

As idéias que foram expressas por Newell (1981) discorrem sobre o nível do conhecimento, adotando principalmente os conceitos de “agente”, “metas”, “ações”, “comportamento”, “racionalidade”, “conhecimento” e “representação”.

O agente é um ente cognoscente capaz de praticar ações e, nesse contexto, possui metas e conhecimentos a partir dos quais atua em seu ambiente, obedecendo ao princípio da racionalidade. Por este princípio, se o agente tem conhecimento que uma de suas ações conduzirá a uma de suas metas, então irá selecionar aquela ação. Newell (1981) destaca que o conhecimento e a racionalidade estão intimamente entrelaçados. O conhecimento de um agente é definido inteiramente em termos de seu ambiente, que é o objeto de suas metas, e cujas características respondem pelo modo como desenvolve suas ações para atingir essas metas.

As metas são o conjunto de conhecimentos sobre a forma como as coisas devem ficar no ambiente. Sua diferença em relação ao restante do conhecimento reside no fato de influenciar o comportamento do agente de um modo distinto, como aquilo que ele se esforça para realizar. Assim, seu comportamento depende do que sabe, do que quer e dos meios que possui para interagir com o ambiente.

Clancey (1989) considera que o conhecimento, pela visão de um observador, é aquilo que ele atribui a um agente para descrever e explicar as suas interações recorrentes com o seu ambiente. Esse conhecimento deve ser caracterizado de modo completamente funcional, em termos do que permite realizar, e não estruturalmente, em função de objetos físicos com propriedades particulares e relações.

Newell (1981) e outros autores concebem o conhecimento como um potencial para gerar ação, o qual é pressentido ou apenas imaginado como o resultado de processos interpretativos, que atuam sobre expressões simbólicas. Falta a percepção de que, neste contexto, o conhecimento não é simplesmente uma coleção de expressões simbólicas com alguma organização estática, porque também requer processos e estruturas de dados.

Para Clancey (1989), a capacidade para executar uma ação não pode ser reduzida a simples descrições dessa ação ao nível do conhecimento. Por isso, para a Engenharia do Conhecimento, uma representação de conhecimento capaz de executar ações requer estruturas e processos e, assim, o conhecimento permanece sempre como algo abstrato e imaterial, não sendo possível retê-lo em um objeto. É importante notar que mesmo o objeto da modelagem no nível do conhecimento não é o conhecimento em si, mas sim o comportamento, ou seja, a interação observada entre um agente e seu ambiente.

O conhecimento serve como a especificação do que uma estrutura simbólica é capaz de realizar e a representação deste conhecimento deve ser expressa em termos de descrições do modelo e não de instâncias. As representações existem no nível simbólico, como sistemas de estruturas de dados e processos, que realizam um conjunto de conhecimento no nível do conhecimento. Assim, uma representação deve abranger tanto o conhecimento quanto seu acesso. Por isso pode ser descrita como um sistema para prover acesso a um corpo de conhecimento, isto é, para disponibilizar o conhecimento de uma forma que possibilite a sua utilização para efetuar a seleção de ações a serviço das metas.

É importante evidenciar que o conhecimento necessário para resolver um problema é diferente do processamento necessário para resgatar e usar esse conhecimento em tempo real, o que implica na existência de um perfil de custos computacionais para a entrega de partes diferentes do conhecimento total codificado na representação. Isso caracteriza as especificidades das áreas de Engenharia e Gestão do Conhecimento que têm como finalidade o conhecimento e como objetos de estudo os sistemas mentais, sociais ou materiais executados por humanos ou meios digitais, capazes de armazenar, resgatar, processar (utilizar), desenvolver e disseminar o conhecimento.

2.3.6 O Conhecimento dos “Sistemas Baseados em Conhecimento”

Abordar um sistema no nível do conhecimento é tratá-lo como se possuísse alguns conhecimentos e algumas metas e acreditar que ele fará o máximo visando atingir essas metas, escolhendo suas ações de acordo com o princípio da racionalidade.

A análise do conhecimento deve focar o resultado da interação do usuário com o sistema em seu ambiente ao longo do tempo e, por isso, o conhecimento não pode ser reduzido ou mapeado apenas sobre uma estrutura física ou um dispositivo isolado.

Segundo Newell (1981), para possuir conhecimentos um sistema deve possuir também racionalidade, pois é difícil conceber qualquer outra forma na qual possa ser dito que um sistema possui conhecimento. A classe de sistemas que têm conhecimento é a daquelas direcionadas a metas e o nível do conhecimento permite a descrição do comportamento desse sistema acima do nível simbólico, sem considerações sobre o que exatamente seja esse sistema. Enquanto o nível simbólico é orientado para o sistema, o nível do conhecimento é orientado para o domínio.

É relevante evidenciar que os sistemas baseados em conhecimento ou *knowledge-based system* (KBS) operam sem possuir qualquer estrutura física que seja o conhecimento. Para

Clancey (1989), suas “bases de conhecimento” são modelos de mundo, codificados em nível simbólico por estruturas de dados e processos. Os processos extraem das estruturas o conhecimento nelas modelado, como uma característica sistêmica do conjunto programa/ambiente/usuário. Além do exposto, um sistema será considerado KBS quando sua principal interação for com um usuário humano.

O KBS deve auxiliar um usuário humano na execução de processos de conhecimento, como se também fosse um agente humano, no sentido que suas ações devem ser percebidas pelo usuário como respostas guiadas por um conhecimento aplicado de acordo com o princípio da racionalidade.

2.3.7 Diferenciando as áreas de Informação e Conhecimento

A principal diferença entre focar informações ou conhecimentos é que, no primeiro caso, devem ser enfatizadas as informações que circulam ou devem circular nos processos da organização, e no segundo, os conhecimentos que são necessários para a boa execução dos processos ou para promover a inovação.

Há aqui três questões que foram desenvolvidas anteriormente. A primeira diz respeito ao valor da informação e do conhecimento. Estruturalmente, toda informação é expressão do conhecimento, porém, não são todas as informações ou os conhecimentos expressos que são considerados valiosos em uma dada circunstância. A segunda questão considera que, para um observador externo, o conhecimento valioso se manifesta em ações ou atividades também valiosas, de acordo com as circunstâncias. A terceira questão propõe que há conhecimentos valiosos que ainda não foram expressos como informação passível de ser comunicada, mas que é pressentido ou imaginado em atitudes ou atividades.

Uma estratégia de conhecimento¹ deve ser guiada pela compreensão de quais conhecimentos são estratégicos e por que (ZACK, 2002). Para isso, é importante mapear: 1- o que deve ser conhecido; 2- por quem deve ser conhecido; 3- porque deve ser feito; 4- quando deve ser feito; 5- onde deve ser feito; e 6- como deve ser feito. Além disso, o esforço para realizar esse levantamento deve ser ponderado por uma análise de custo e benefício. O conjunto de procedimentos e decisões aqui apresentado é pertinente à Gestão do Conhecimento.

É importante perceber que existe um crescente interesse sobre Gestão do Conhecimento. No livro “Sociedade Pós-Capitalista”, Peter Drucker (1993) afirma que neste momento, o único recurso significativo é o conhecimento que capacita à ação, compondo os meios para se obter

¹ Um aprofundamento dessa questão pode ser encontrado em Zack (1999), que apresenta um framework para descrever e avaliar a estratégia de conhecimento utilizada por uma organização, e em Zack (2002), que propõem diferenças entre os termos estratégia de conhecimento e estratégia de gestão do conhecimento.

resultados sociais e econômicos. O valor do conhecimento está diretamente relacionado ao seu potencial de orientar, de forma econômica, o dispêndio de energia para a realização de uma atividade.

A informação exerce papel fundamental nos processos de conhecimento. Por isso, a recuperação da informação é um problema importante, diante do fenômeno da explosão informacional. A grande quantidade de informações que é produzida e disponibilizada dificulta o acesso eficiente, o trabalho de identificação e os processos de utilização. Para que seja capaz de realizar todo o seu potencial, a informação relevante para um problema ou circunstância precisa estar disponível no tempo certo, para a pessoa certa e na forma adequada (MARCONDES, 2001).

Para lidar com o grande fluxo de informações, não basta disponibilizar diversas tecnologias da informação, é preciso criar um ambiente propício a trabalhar com foco no conhecimento, com estratégias e procedimentos suportados por sistemas adequados aos problemas enfrentados. É recomendável que os sistemas baseados em conhecimento, capazes de interagir com o conhecimento disponível e selecionar conhecimentos de valor ou, ainda, gerar conhecimentos de valor circunstanciado, apenas sejam introduzidos quando seus resultados puderem gerar ganhos significativos para a organização.

É preciso gerar informações de valor decorrentes de processos gestores do conhecimento, porque a capacidade de operar com informações concisas exige um conhecimento sobre o conhecimento. O domínio desse conhecimento aplicado ao conhecimento permite a elaboração de sistemas físico-digitais e gerenciais que permitem ao ente cognoscente, com limitada capacidade para operações simultâneas, a resolução de problemas complexos.

2.4 A Modelagem de Conhecimento no CommonKADS

Diversas abordagens foram desenvolvidas para a modelagem de conhecimento visando seu uso em sistemas informatizados. As três mais conhecidas para a construção de um KBS são, o CommonKADS (SCHREIBER *et al.*, 2000), o PROTÉGÉ (GENNARI *et al.*, 2003), e o MIKE (ANGELE *et al.*, 1998). Em seu artigo *Knowledge Engineering: Principles and methods*, Studer *et al.* (1998) analisa essas três abordagens comparando suas propostas.

De modo geral, enquanto o PROTÉGÉ e o MIKE estão mais preocupados com a questão da implementação de um KBS, o CommonKADS trabalha primeiro com a análise dos elementos organizacionais que podem determinar seu sucesso ou fracasso. Para isso prescreve um conjunto de seis modelos, apresentados na Figura 2.4, capazes de descrever o ambiente no qual um KBS poderá ou não ser inserido.

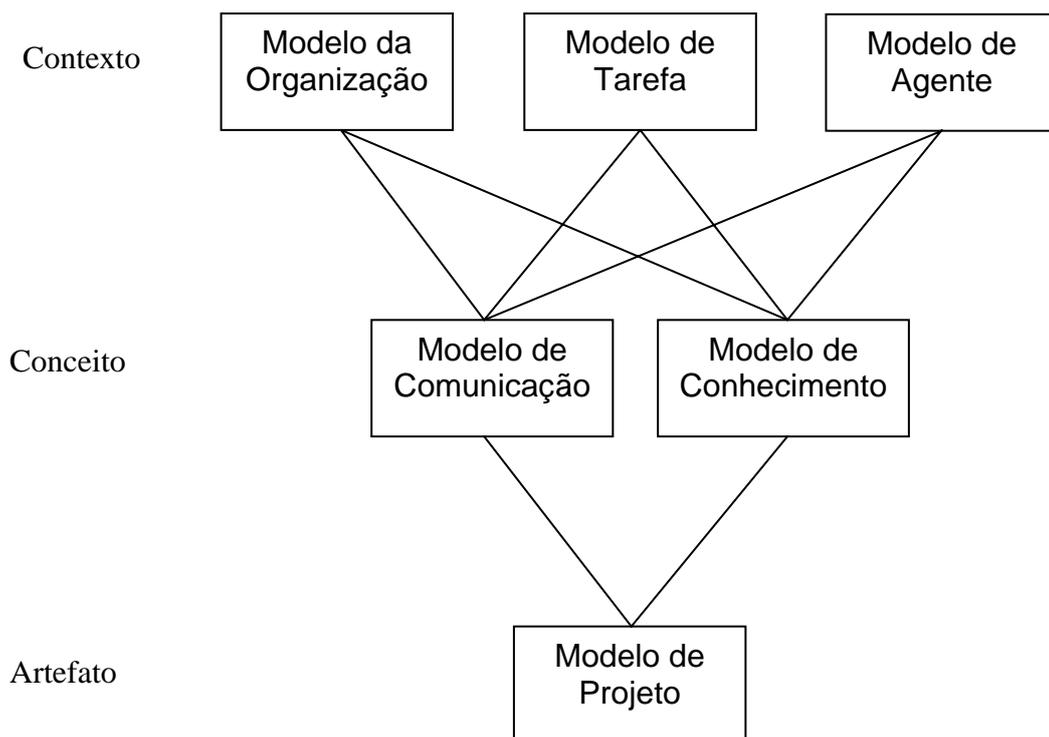


Figura 2.4 – Conjunto de modelos do CommonKADS

Fonte: SCHREIBER *et al.* (2000), p. 18.

Esses modelos estão divididos em 3 grupos, e cada grupo tem como objetivo responder uma pergunta.

1. Porque desenvolver um KBS? Que benefícios, custos e impacto organizacional ele pode provocar? Os modelos da organização, da tarefa e do agente ajudam a entender o ambiente e o contexto organizacional.
2. Qual a natureza e estrutura do conhecimento envolvido e das comunicações a ele relacionadas? A descrição conceitual do conhecimento aplicado nas tarefas é o ponto principal retratado pelos modelos do conhecimento e da comunicação.
3. Como esse conhecimento deve ser implementado em um sistema de computador? Os aspectos técnicos são o foco do modelo de projeto.

Cada modelo se ocupa de um aspecto específico, mas em conjunto podem fornecer uma visão abrangente do problema. Segue agora uma descrição de cada modelo.

O Modelo da Organização descreve as características principais de uma organização visando descobrir as deficiências do processo atual bem como as oportunidades para melhorá-lo com a implementação de um KBS, estabelecendo sua viabilidade e prevendo também os impactos que poderão ser gerados (STUDER *et al.*, 1998; SCHREIBER *et al.*, 2000; MOTTA, 2001).

O Modelo de Tarefa providencia uma descrição hierárquica das tarefas que são executadas na unidade organizacional na qual o KBS será instalado, incluindo a especificação de que agentes são atribuídos a cada tarefa (STUDER *et al.*, 1998). Está diretamente relacionado ao modelo de organização, listando as subpartes do processo empresarial, entradas e saídas, condições e critérios de desempenho, bem como os recursos e competências necessárias. Resumindo, provê uma análise detalhada dos processos de negócio identificados no Modelo da Organização (MOTTA, 2001).

O Modelo de Agente especifica quem realiza as tarefas descritas no Modelo de tarefa (normalmente uma pessoa ou algum sistema de software, como um KBS) e as capacidades de cada agente envolvido na execução dessas tarefas, o que inclui suas características, suas competências e suas autonomias e restrições para agir (STUDER *et al.*, 1998; MOTTA, 2001). Além disso, lista também as ligações que existem entre os agentes para a execução de uma tarefa (SCHREIBER *et al.*, 2000).

O Modelo de Comunicação especifica as várias interações entre os diferentes agentes, especificando o tipo de informação que é trocada e quem inicia a interação (STUDER *et al.*, 1998). O diálogo entre os agentes é estruturado em três níveis; em termos do *plano de comunicação global*, de *transações individuais* e de especificações *de trocas de informações detalhadas*, que abrangem o conteúdo e a forma das mensagens individuais, bem como o tipo de comunicação. Isto torna isto possível especificar a intenção de uma mensagem, em função

de uma lista de tipos predefinidos: solicitação; requisição; encomenda; proposta; oferta; pergunta; resposta; etc. (MOTTA, 2001).

O Modelo de Conhecimento explica em detalhes os tipos e estruturas de conhecimento usadas na execução de uma tarefa, através de uma descrição independente da implementação e inteligível para humanos, sendo assim um importante veículo de comunicação entre o engenheiro do conhecimento e os especialistas e usuários do KBS sobre seus aspectos de solução de problemas (SCHREIBER *et al.*, 2000). Studer *et al.* (1998) descreve como a maior contribuição das abordagens KADS, o uso de sub-modelos distintos para cada grupo de estrutura de conhecimento, chamadas de categorias de conhecimento.

A primeira contém o conhecimento do domínio (visão estática), que possui semelhanças com o modelo de dados ou de objetos da engenharia de software, apresentando o conhecimento específico do domínio, seus atributos, suas relações e heurísticas específicas para o processo de resolução de problemas, além dos tipos de informação usados em uma aplicação (ANGELE *et al.*, 1998; SCHREIBER *et al.*, 2000).

A segunda contém o conhecimento de inferência (visão funcional), que descreve os passos de inferência básicos que são executados com o conhecimento do domínio para a solução de tarefas (MOTTA, 2001). As inferências são os blocos de construção de uma máquina de raciocínio (SCHREIBER *et al.*, 2000).

A terceira contém o conhecimento da tarefa (visão dinâmica), e descreve quais metas guiam uma aplicação e como essas metas podem ser realizadas através de uma decomposição em sub-tarefas no nível de inferências. Para isso possui uma estrutura de controle que descreve o comportamento dinâmico da tarefa, ou seja, especifica as seqüências de inferências necessárias para a resolução de problemas (MOTTA, 2001; ANGELE *et al.*, 1998; SCHREIBER *et al.*, 2000).

O Modelo de Projeto mapeia a análise conceitual explicitada nos modelos de conhecimento e comunicação para uma aplicação específica, e detalha a plataforma designada de hardware e software, os vários módulos de software incluídos no sistema designado, suas especificações funcionais e técnicas e a interligação entre estes módulos e os componentes conceituais identificados durante a fase de análise (MOTTA, 2001).

O fato dos modelos de conhecimento serem apenas outra classe de modelos nesse conjunto de modelos mostra a evolução da metodologia dos primeiros dias na qual o foco era a aquisição de conhecimento, para o cenário de hoje, no qual o CommonKADS provê uma estrutura abrangente que integra as técnicas de Engenharia do Conhecimento no contexto de projetos genéricos de gestão do conhecimento (MOTTA, 2001).

2.4.1 Tarefas Intensivas de Conhecimento

Através da associação entre o conhecimento da tarefa e o conhecimento de inferência são construídos os esquemas de tarefas, um tipo comum de combinação reutilizável de elementos de modelos. Um esquema de tarefa supre o engenheiro do conhecimento com inferências e tarefas típicas para a resolução de problemas de um tipo particular, especificando também o esquema de domínio típico requerido do ponto de vista da tarefa. A apresentação de cada esquema de tarefa no CommonKADS inclui uma descrição geral, o método padrão e sua estrutura de inferência.

Os esquemas de tarefas do CommonKADS, apresentados na Figura 2.5, foram baseados nas tipologias de tarefas intensivas de raciocínio encontradas na psicologia cognitiva, as quais foram adaptadas por vários autores para uso na Engenharia do Conhecimento (SCHREIBER *et al.*, 2000).

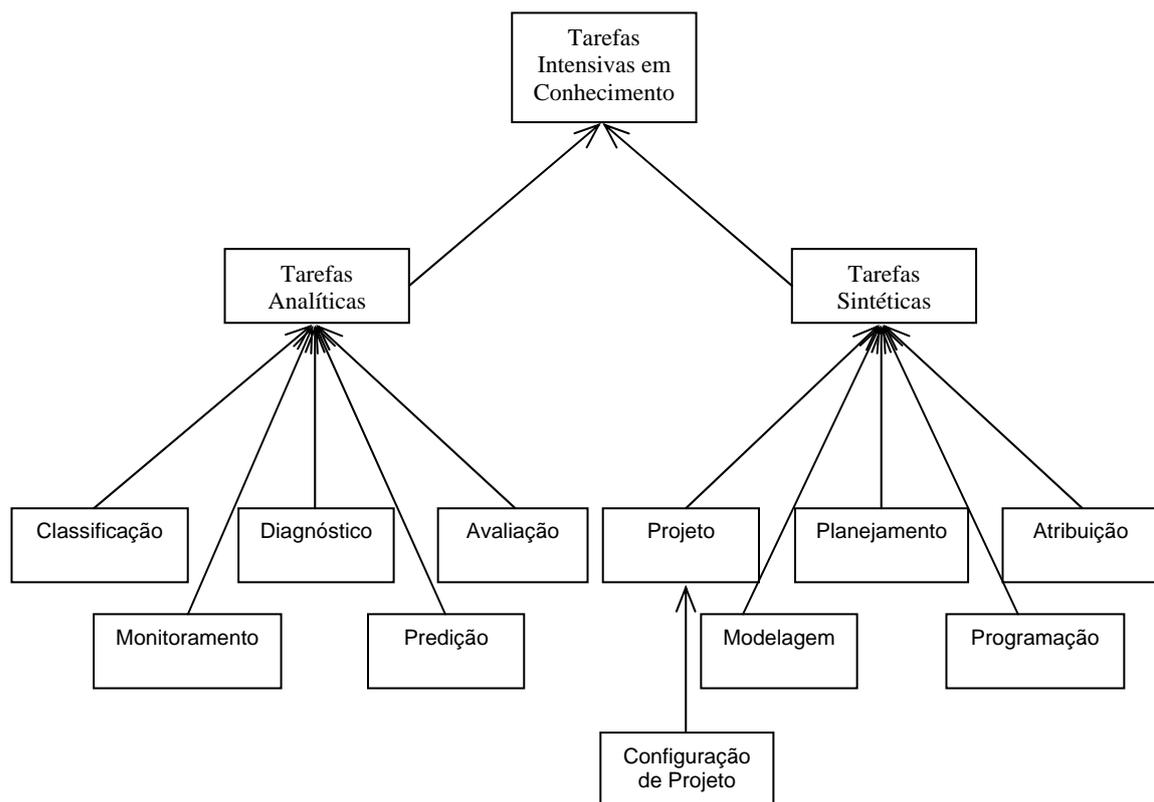


Figura 2.5 – Tipologia das tarefas intensivas em conhecimento
 Fonte: SCHREIBER *et al.* (2000), p. 125.

Esses esquemas de tarefas estão divididos em dois grupos, as tarefas analíticas e as tarefas sintéticas, que se diferenciam pelo sistema ou objeto ao qual são aplicadas. Nas tarefas analíticas, o sistema é pré-existente, embora não seja completamente conhecido, e nas tarefas

sintéticas o sistema ainda não existe, sendo o propósito da tarefa a própria construção do sistema.

As tarefas analíticas têm como entrada observações do sistema e produzem como saídas caracterizações desse sistema, para isso trabalham com conhecimento sobre o modelo do sistema e sua característica principal é fornecer uma descrição do sistema. São elas as tarefas de Classificação, Diagnóstico, Avaliação, Monitoramento e Predição.

- Na Classificação, são avaliadas características do objeto para enquadrá-lo em uma classe de um conjunto pré-definido.
- No Diagnóstico, são avaliados sintomas do objeto para, de acordo com um modelo de comportamento do sistema, determinar as falhas que provocam aqueles sintomas.
- Na Avaliação, o objetivo é caracterizar um caso em termos de classes de decisão, de acordo com um conjunto de normas e critérios pré-definidos.
- No Monitoramento, é avaliado, a partir de um conjunto de dados recebidos, o comportamento dinâmico de um sistema onde o comportamento atual é comparado com os estados anteriores para identificar discrepâncias em relação ao comportamento esperado.
- Na Predição, é analisado o comportamento do sistema até o momento atual para construir uma descrição do estado do sistema em algum ponto no futuro.

As tarefas sintéticas têm como entrada requisitos que devem ser satisfeitos no desenvolvimento de uma solução, normalmente relacionados com a construção de um sistema. São elas as tarefas de Projeto (com a sub-tarefa de Configuração de Projeto), Atribuição, Planejamento, Programação e Modelagem.

- O Projeto tem como objetivo gerar um projeto de construção de um sistema a partir de um conjunto de requisitos e contando com o conhecimento dos materiais, componentes, restrições e preferências aplicáveis. Devido ao seu caráter altamente subjetivo (envolvendo pontos como criatividade) as implementações atuais contemplam apenas a sub-tarefa de Configuração de Projeto, onde os requisitos são atendidos apenas por configurações pré-determinadas.
- A Atribuição tem como objetivo gerar um mapeamento (parcial ou total) entre dois conjuntos de objetos como, por exemplo, atribuir vagas a carros em um estacionamento.
- O Planejamento tem como objetivo tratar as atividades que devem ser executadas no desenvolvimento de um projeto e suas dependências temporais.

- A Programação tem como objetivo controlar a alocação temporal de recursos para a execução de uma seqüência de atividades, normalmente gerada por uma tarefa de Planejamento. Seu resultado é um mapeamento entre atividades e períodos de tempos que obedece às restrições impostas e respeita o máximo possível as preferências apontadas.
- A Modelagem tem como objetivo a construção de uma descrição abstrata de um sistema para explicar ou prever certas propriedades ou fenômenos do sistema e raramente são automatizadas.

2.5 Qualidade da informação

Composta por dois termos cujas definições são ambíguas e pouco precisas, não é de surpreender que a qualidade da informação (QI) também não possua uma definição formal bem estabelecida, ainda que, ao mesmo tempo, quase todos tenham uma noção intuitiva sobre seu significado.

A norma NBR ISO 9000 (Sistemas de gestão da qualidade – Fundamentos e vocabulário) define qualidade como o “grau no qual um conjunto de características inerentes satisfaz a requisitos” e observa que esses requisitos podem ser implícitos. Além disso, sua avaliação depende da percepção do cliente, que pode ser ainda mais importante do que a avaliação de características intrínsecas. Este fator subjetivo da avaliação da qualidade é ressaltado como um complicador em muitas abordagens desenvolvidas.

O gerenciamento da qualidade evoluiu muito ao longo do último século. Segundo Barreto Júnior (1997), os programas de qualidade começaram com inspeções pós-produção centradas na avaliação do produto final, e a seguir vieram as avaliações dos subprodutos das etapas de produção, a avaliação do procedimento de produção como um todo, a avaliação das pessoas envolvidas no processo, a avaliação e otimização por processo, a avaliação do projeto de produção e a avaliação da própria concepção do produto. Montgomery (2004) apresenta em uma tabela a evolução dos métodos de qualidade ao longo do século XX.

Durante essa evolução, um ponto que merece destaque é o entendimento que a qualidade do processo é mais importante que a qualidade do produto. Um bom exemplo disso é um prato de comida, pois é possível dizer mais sobre sua qualidade observando a sua preparação do que analisando o produto final. Afinal, não é possível saber sobre a higiene ou o valor nutricional apenas comendo o prato (BARRETO JÚNIOR, 1997).

De acordo com a NBR ISO 9000, qualquer atividade, ou conjunto de atividades, que usa recursos para transformar insumos (entradas) em produtos (saídas) pode ser considerado como um processo. Dessa forma, um produto é o “resultado de um processo”. A norma identifica quatro categorias genéricas de produto:

- ✓ Serviços (ex.: transporte);
- ✓ Informações (ex.: programa de computador, dicionário);
- ✓ Materiais e equipamentos (ex.: parte mecânica de um motor); e
- ✓ Materiais processados (ex.: lubrificante).

Entretanto essa classificação não é exclusiva, ou seja, muitos produtos abrangem elementos que pertencem a diferentes categorias genéricas de produtos.

Analisar a qualidade da informação normalmente envolve pelo menos duas categorias genéricas de produtos, a de informações e a de serviços. É importante atentar que os produtos dessas categorias são geralmente intangíveis, e que um serviço é o resultado de pelo menos uma atividade desempenhada necessariamente na interface entre o fornecedor e o cliente (NBR ISO 9000). Para o produto informação, essa interface pode ser apenas o meio de solicitação ou abranger também o meio de entrega.

Enquanto a NBR ISO 9000 descreve informação simplesmente como dados significativos, este trabalho, conforme explicado no capítulo 2.3, entende informação como uma expressão de conhecimento materializada em uma substância e que normalmente pode ser registrada em uma mídia ou transmitida através de um canal como mensagem.

Isto permite que, além do processo produtivo, parte do produto informação possa ser analisada antes da entrega, através de suas propriedades conhecidas como características da informação.

Para a NBR ISO 9000, uma característica é uma propriedade diferenciadora que pode ser classificada: quanto a sua origem como inerente ou atribuída; quanto a sua forma de percepção ou avaliação como qualitativa ou quantitativa; e quanto ao seu tipo como físicas, sensoriais, comportamentais, temporais, ergonômicas ou funcionais.

A identificação de quais características são necessárias para a avaliação da qualidade da informação tem sido objeto de estudo em trabalhos de diversas áreas, tendo o interesse por este tema aumentado significativamente a partir da década de 90.

Esses trabalhos normalmente tratam os termos dados e informações ou como sinônimos, ou considerando dados como informações “brutas” que podem ser transformadas no produto informação através de um processo. Neste caso, ressaltam que aquilo que é considerado informação em um processo pode ser visto como dado em outro, ou seja, uma distinção meramente funcional. Então, para simplificar o texto, a partir deste ponto, exceto onde indicado, os estudos de qualidade de dados e qualidade de informações serão tratados como equivalentes, assim como os termos dados e informações.

Batini e Scannapieco (2006) destacam que devido à relevância da qualidade e a variedade de tipos de dados e sistemas de informações, o tratamento da qualidade de dados consiste em uma área de investigação complexa e multidisciplinar que envolve vários problemas de pesquisa e áreas de aplicação, e que os problemas de pesquisa utilizam modelos, técnicas e ferramentas que são permeados pelos tópicos de duas áreas, as dimensões e as metodologias, conforme exibido na Figura 2.6.

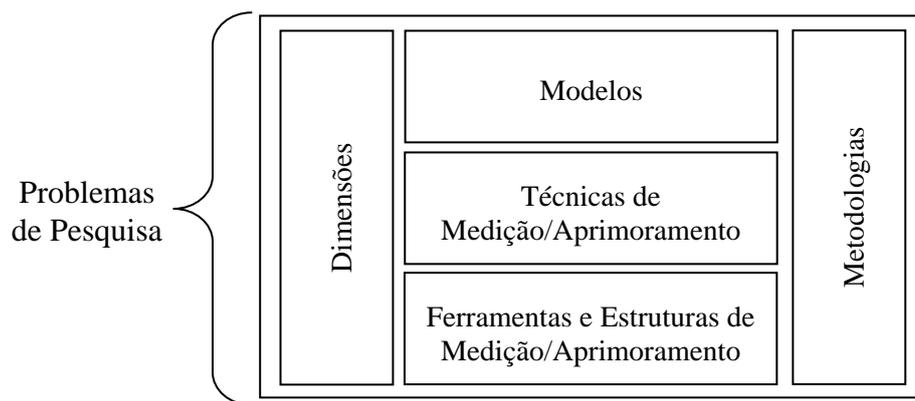


Figura 2.6 – Principais assuntos para a qualidade de dados
 Fonte: adaptado de BATINI e SCANNAPIECO (2006), p. 11.

Segundo Batini e Scannapieco (2006), qualidade de dados é uma nova área de pesquisa, mas diversas áreas de conhecimento têm tratado de problemas a ela relacionados ou sobrepostos, e para isso desenvolveram paradigmas, modelos e metodologias de grande importância para a fundamentação da área de qualidade de dados. Cinco dessas áreas tem suas contribuições destacadas a seguir:

1. Estatística: possui um conjunto de métodos para coletar, analisar, apresentar e interpretar dados, além de um amplo espectro de métodos e modelos desenvolvidos nos últimos dois séculos para realizar previsões e formular decisões em contextos onde as informações disponíveis para o domínio de interesse são incertas e imprecisas. A estatística e a sua metodologia como base da análise de dados estão relacionadas a dois tipos básicos de problemas: (a) resumir, descrever, e explorar dados; e (b) usar dados amostrados para deduzir a natureza do processo que os produziu. Como dados de baixa qualidade são uma representação errônea da realidade, foi desenvolvida uma grande variedade de métodos estatísticos para medir e melhorar a qualidade dos dados.
2. Representação do conhecimento: é o estudo de como o conhecimento sobre um domínio de aplicação pode ser representado e que tipos de raciocínios podem ser implementados a partir dele. Esse conhecimento pode ser representado através de procedimentos embutidos em códigos de programas, implicitamente, como padrões de ativação em uma rede neural, ou mesmo explicitamente, em termos de uma base de conhecimentos, contendo fórmulas ou regras lógicas. Gerar representações ricas do domínio de aplicação que permitam a aplicação de processos de raciocínio está se tornando uma importante forma de alavancagem para muitas técnicas de melhoria da qualidade dos dados.

3. Mineração de dados: é um processo analítico normalmente aplicado na exploração de grandes conjuntos de dados à procura de padrões consistentes ou relações sistemáticas entre atributos ou variáveis. Mineração de dados exploratória é o processo preliminar de descobrir estruturas em um conjunto de dados usando resumos estatísticos, visualizações e outros meios. Neste contexto, obter boa qualidade de dados é um objetivo intrínseco de qualquer atividade de mineração de dados, pois, de outra forma, o processo de descoberta de padrões, relações e estruturas estaria seriamente deteriorado. As técnicas de mineração de dados também podem ser usadas em um amplo espectro de atividades para aprimorar a qualidade dos dados.

4. Sistemas de informações gerenciais: definidos como sistemas que provêm as informações necessárias para administrar uma organização efetivamente. Como dados e conhecimento estão se tornando recursos relevantes nos processos operacionais e decisórios, e dados de baixa qualidade resultam processos de baixa qualidade, está se tornando cada vez mais importante prover os sistemas de informações gerenciais com funcionalidades e serviços que permitam controlar e aprimorar a qualidade dos dados enquanto recursos.

5. Integração de dados: considerada como uma das atividades básicas para aprimorar a qualidade dos dados, tem como meta construir e apresentar uma visão unificada para dados originários de fontes heterogêneas pertencentes a sistemas de informação distribuídos, cooperativos ou *peer-to-peer* (P2P).

Além dessas áreas, Scannapieco *et al.* (2005), destaca a contribuição dos pesquisadores de gestão, que no início dos anos 80 começaram a focar em como controlar sistemas de produção de dados de forma a detectar e eliminar problemas de qualidade de dados, e observa que foi somente no início dos anos 1990 que os cientistas da computação começaram a considerar o problema de definir, medir e aprimorar a qualidade de dados eletrônicos armazenados em bancos de dados, data warehouses e sistemas legados.

Ainda há muito a ser desenvolvido e implementado em relação à qualidade de dados. Turner e Evans (2007), baseados no trabalho de Bitterer (2007), usaram uma pirâmide (Fig. 2.7) para representar cinco possíveis níveis de maturidade de uma organização em relação à qualidade dos dados e identificaram que existe 80% de probabilidade de que até 2010 mais de 75% das organizações não terão avançado ainda além dos níveis 1 e 2 e que até 2012 menos de 10% das organizações terão atingido o nível 5.



Figura 2.7 – Modelo de maturidade da qualidade de dados do Gartner
 Fonte: adaptado de TURNER e EVANS (2007).

Para avançar além dos níveis 1 e 2, a alta gerência precisa compreender que qualidade de dados é mais do que um problema tecnológico e que os dados precisam ser tratados como um ativo corporativo cujo valor depende diretamente da sua qualidade (BATINI e SCANNAPIECO, 2006).

2.5.1 Definições para a qualidade da informação

Eppler (2006) apresenta sete definições para qualidade da informação encontradas na literatura:

- Qualidade da informação pode ser definida como informação que é julgada como adequada para sua utilização pelos consumidores da informação (HUANG *et al.*, 1999).
- Qualidade da informação é a característica da informação de atender ou exceder as expectativas dos clientes (KAHN e STRONG, 1998).
- Informação de qualidade é a informação que está em conformidade com as especificações ou requisitos (KAHN e STRONG, 1998).
- Qualidade da informação pode ser definida como a diferença entre a informação requerida em função de um objetivo e a informação recebida. Em uma situação ideal não existirá nenhuma diferença entre a informação requerida e a obtida. Uma medida qualitativa para a qualidade da informação pode ser expressa pela seguinte tendência: quanto menor a diferença, maior a qualidade da informação (GERKES, 1997).
- Qualidade da informação é a característica da informação de ser de grande valor para os seus usuários (LESCA e LESCO, 1995).

- O grau no qual a informação tem características de conteúdo, forma e tempo que fazem com que ela tenha valor para usuários finais específicos (BRIEN, 1991).
- Qualidade da informação é a característica da informação de satisfazer os requisitos funcionais, técnicos, cognitivos e estéticos dos produtores, administradores, consumidores e especialistas da informação (EPPLER, 1999).

Scannapieco *et al.* (2005), baseados no trabalho de Wang (1998), destaca que a qualidade da informação pode ser simplesmente caracterizada como sua adequação ao uso pretendido. Embora simples de enunciar, a avaliação segundo este conceito é extremamente complexa.

Nehmy e Paim (1998) relatam a forte ligação entre o conceito de qualidade da informação e a questão da valoração da informação, onde os termos qualidade e valor são normalmente tratados como equivalentes devido à insuficiência de elaboração teórica para diferenciá-los da noção oriunda do senso comum. As autoras questionam ainda quando no estudo da qualidade da informação é aplicada o conceito de valor de uso da informação sem que seja observado o rigor exigido pela teoria da economia política.

Concluem, então, que isso reforça a vagueza, imprecisão e ambigüidade da noção de qualidade da informação e destacam que apesar de diversos autores reconhecerem esse fato, grandes esforços são realizados para traduzir a qualidade da informação em atributos imediatos, o que tem contribuído para a elaboração de definições que se concentram no levantamento de aspectos, dimensões, atributos ou características do fenômeno. A principal crítica realizada é em relação à busca da elaboração de uma forma de medição de um conceito que ainda não foi alcançado.

Entretanto, isso não invalida que sejam realizados esforços para o aprimoramento da qualidade da informação, apenas alerta que qualquer abordagem realizada não será capaz de englobar todas as implicações envolvidas nesse conceito enquanto sua abrangência teórica não puder ser determinada.

2.5.2 Conferências sobre qualidade de dados ou de informação

Várias conferências internacionais foram promovidas pelas comunidades de banco de dados e de sistema de informação onde a qualidade de dados foi o tópico principal, como por exemplo: a Conferência Internacional em Qualidade de Informação (ICIQ), iniciada em 1996 e tradicionalmente organizada no Instituto de Tecnologia de Massachusetts (MIT) em Boston; o Workshop Internacional de Qualidade de Informação em Sistemas de informação (IQIS),

mantido em conjunto com a conferência SIGMOD desde 2004; o workshop internacional sobre Qualidade de Dados e Informação (DIQ), mantido em conjunto com a Conferência em Engenharia de Sistemas de informação Avançada (CAiSE) desde 2004; e o workshop internacional em Qualidade de Sistemas de Informação (QoIS), mantido em conjunto com a conferência de Relação de Entidades (ER) desde 2005. Existem também várias conferências nacionais, como as organizadas na França, Alemanha e nos Estados Unidos.

2.5.3 Gestão da qualidade na NBR ISO 9000

A NBR ISO 9000 apresenta oito princípios da gestão da qualidade que podem ser usados pela Alta Gerência para a implantação de um programa de qualidade:

- Foco no cliente: é preciso identificar quais são as necessidades atuais e futuras dos clientes, os seus requisitos e buscar meios de exceder suas expectativas.
- Liderança: é necessário estabelecer uma unidade de propósito e um rumo para a organização e criar e manter um ambiente interno no qual as pessoas estejam comprometidas em alcançar os objetivos da organização.
- Envolvimento de pessoas: é o que possibilita o uso de suas habilidades para o benefício da organização.
- Abordagem de processo: A forma mais eficiente de alcançar os resultados desejados é gerenciar atividades e recursos como um processo.
- Abordagem sistêmica para gestão: identificar, entender e gerenciar processos inter-relacionados como um sistema contribui para atingir os objetivos da organização com eficácia e eficiência.
- Melhoria contínua: deve ser incorporada pela organização como um objetivo permanente.
- Abordagem factual para tomada de decisão: decisões eficazes são baseadas na análise de dados e informações.
- Benefícios mútuos na relação com os fornecedores: uma relação de benefícios mútuos aumenta a habilidade de ambos em agregar valor.

Além desses princípios, a NBR ISO 9000 identifica várias etapas para o desenvolvimento e implantação de um sistema de gestão da qualidade:

- a) determinação das necessidades e expectativas dos clientes;
- b) estabelecimento da política e objetivos de qualidade da organização;

- c) determinação dos processos e responsabilidades necessários para atingir os objetivos de qualidade;
- d) determinação e fornecimento dos recursos necessários para alcançar os objetivos de qualidade;
- e) estabelecimento de métodos para medir a eficácia e eficiência de cada processo;
- f) aplicação dessas medidas para determinar a eficácia e eficiência de cada processo;
- g) determinação dos meios para prevenir não-conformidades e eliminar suas causas; e
- h) estabelecimento e aplicação de um processo para melhoria contínua do sistema de gestão da qualidade.

A Figura 2.8 apresenta um esquema de um sistema de gestão da qualidade onde é possível perceber o inter-relacionamento e influência de diversas etapas com base em uma visão de processos.

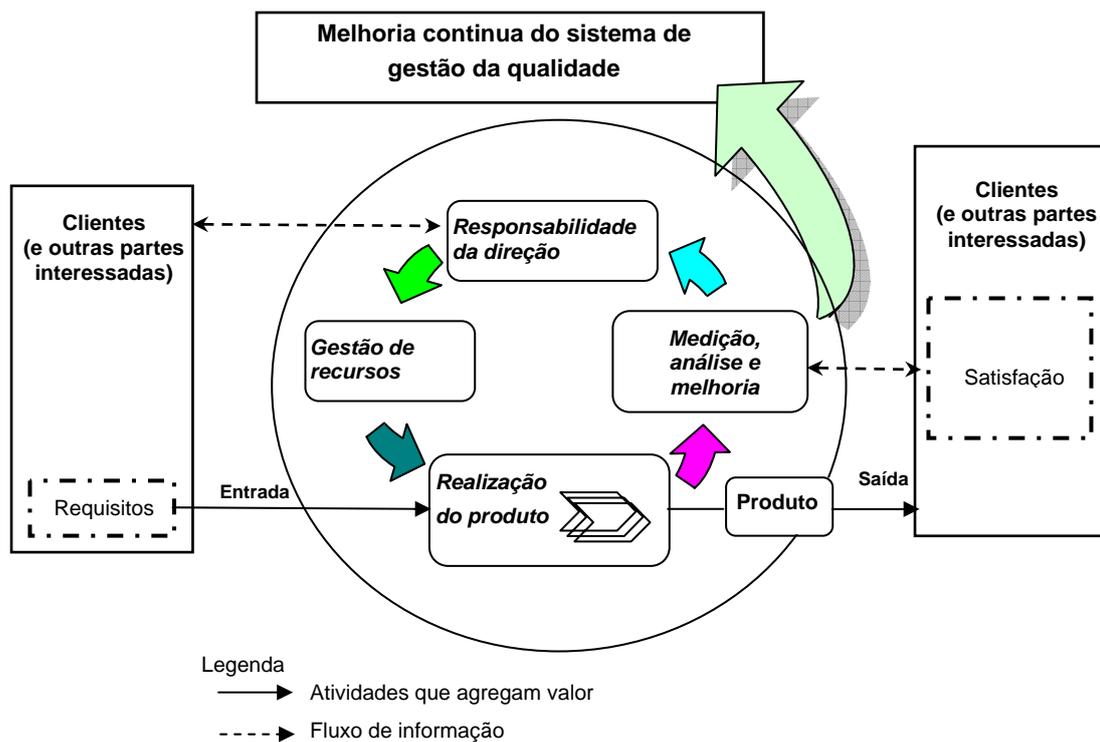


Figura 2.8 – Modelo de um sistema de gestão da qualidade baseado em processo
 Fonte: Baseado na Figura 1 da NBR ISO 9000, p. 3.

Para que um sistema de gestão da qualidade funcione, é primordial o envolvimento da Alta Gerência, que segundo a NBR ISO 9000 deve exercer as seguintes funções:

- ✓ estabelecer e manter a política e os objetivos de qualidade da organização;
- ✓ promover a política e os objetivos de qualidade por toda a organização para aumentar a conscientização, motivação e envolvimento;

- ✓ assegurar o foco nos requisitos do cliente;
- ✓ assegurar que processos apropriados são implementados para possibilitar que requisitos de clientes sejam atendidos e que os objetivos de qualidade sejam alcançados;
- ✓ garantir que um sistema de gestão da qualidade eficaz e eficiente seja estabelecido, implementado e mantido para atingir os objetivos de qualidade;
- ✓ garantir a disponibilidade dos recursos necessários;
- ✓ analisar criticamente o sistema de gestão da qualidade periodicamente;
- ✓ decidir sobre ações a serem adotadas em relação à política e aos objetivos de qualidade; e
- ✓ decidir sobre as ações para melhoria do sistema de gestão da qualidade.

Para a avaliação do sistema de gestão da qualidade é preciso responder a quatro questões básicas em relação a cada um dos processos avaliados:

- ✓ O processo está identificado e apropriadamente definido?
- ✓ As responsabilidades estão atribuídas?
- ✓ Os procedimentos estão implementados e mantidos?
- ✓ O processo é eficaz em alcançar os resultados requeridos?

Vários dos princípios da NBR ISO 9000 destacados neste trabalho estão presentes nas abordagens de qualidade da informação analisadas a seguir.

2.5.4 – Principais abordagens para o tratamento da qualidade da informação

Vários trabalhos têm sido publicados ao longo das últimas duas décadas sobre qualidade da informação. Uma boa relação de livros e artigos sobre esse assunto pode ser obtida no site da *International Association for Information and Data Quality* (IAIDQ). Muitas dessas abordagens foram construídas com base nos trabalhos de gerenciamento e controle da qualidade, com destaque para o ciclo PDCA (*Plan-Do-Check-Act*) popularizado por W. Edward Deming e que teve como precursor o ciclo Shewart (inicialmente como PDSA, onde o “S” significava *Study*). Segundo Eppler (2006), as abordagens baseadas no gerenciamento da qualidade total (*Total Quality Management - TQM*), como as de English (1999), Redman (1996) e Huang *et al.* (1999), oferecem soluções realizáveis para muitas aplicações de bancos de dados e *data warehouse*.

Os trabalhos publicados na área de qualidade da informação normalmente estruturam os problemas a serem tratados de acordo com as dimensões (conteúdo, formato, tempo), visões (informação como produto ou processo/serviço) ou ciclo de vida (produção, armazenagem, uso) adotados em seus escopos, sendo que algumas dessas abordagens procuram estabelecer relações diretas entre problemas e critérios de qualidade da informação (EPPLER, 2006).

A seguir são comentadas algumas contribuições oriundas desses trabalhos.

1985 – Ballou e Pazer (1985) apresentaram quatro atributos de qualidade dos dados: acurácia (*Accuracy*), que reflete a exatidão em relação à realidade; consistência (*Consistency*), que mostra a ausência de conflitos entre os dados; atualidade (*Currency*), que diz quão recente é a informação; e completeza (*Completeness*), que é o grau no qual os valores estão presentes em uma coleção de dados (UMAR *et al.*, 1999).

1986 – Taylor (1986) em seu modelo de valor agregado em sistemas de informação apresenta cinco atributos de qualidade: acurácia (*Accuracy*), abrangência (*Comprehensiveness*), atualidade (*Currency*), confiabilidade (*Reliability*) e validade (*Validity*) (EPPLER, 2006).

1989 – No seminário *Nordic Council for Scientific Information and Research Libraries* (NORDINFO), realizado em 1989, em Copenhagem, Dinamarca, alguns autores como Donald Marchand, Miriam Ginman e Johan Olaisen começaram a discutir e propor dimensões para a qualidade da informação.

Marchand (1989) identificou cinco abordagens na definição do conceito de qualidade da informação: transcendente (valor absoluto da informação); baseada no usuário (com relação às necessidades individuais); baseada no produto (trata a informação enquanto coisa); baseada na produção (adequação aos padrões estabelecidos da necessidade de informação do consumidor); e baseada na qualidade como um dos aspectos do valor (qualidade é um dos atributos do valor da informação) (CALAZANS, 2008; NEHMY e PAIM, 1998). Marchand (1989) identificou também, oito dimensões da qualidade da informação: o valor real que tem para o usuário, as características suplementares da informação como produto ou serviço, a confiabilidade, o significado ao longo do tempo, a relevância, a validade, a estética e o valor percebido (DANIEL, 1993; NEHMY e PAIM, 1998).

Ginman (1989) trabalhou com a necessidade de informação dos CEOs (*chief executive officer*) e identificou que a necessidade da informação e o seu nível de qualidade variam de acordo com a fase do ciclo de vida da organização, o que é refletido por

alterações nas dimensões da gestão da informação ao longo dessas fases (CALAZANS, 2008).

Olaisen (1989) buscou identificar a qualidade da informação obtida em consultas a fontes internas e externas, conforme atributos oriundos de dois grupos filosóficos que se baseiam no julgamento do usuário: o primeiro que trata dos fatores de qualidade de autoridade cognitiva e o segundo dos fatores de amigabilidade técnica (CALAZANS, 2008). Para Nehmy e Paim (1998), Olaisen (1989) privilegiou o caráter qualitativo da avaliação da informação no contexto da tecnologia eletrônica e construiu um modelo com quatro categorias agrupadoras das características da qualidade: qualidade cognitiva, que reflete como a fonte é valorizada pelo usuário ('credibilidade', 'relevância', 'confiança', 'validade' e 'significado no tempo'); qualidade do desenho da informação (fatores referentes à 'forma', 'flexibilidade' e 'seletividade'); fatores referentes ao produto da informação (como 'valor real' e 'abrangência'); e fatores relativos à qualidade da transmissão (definidos pelo critério da 'acessibilidade').

- 1990 – Wang e Madnick (1990) apresentaram *A Polygen Model for Heterogeneous Database Systems: The Source Tagging Perspective* uma pesquisa que daria início ao programa TDQM do MIT. Este trabalho propôs acrescentar campos, como uma extensão do modelo relacional, para identificar a origem e qualidade dos dados em sistemas de bancos de dados heterogêneos com dados oriundos de múltiplas (poli) fontes (gen) (WANG *et al.*, 2002).
- 1990 – Pautke e Redman (1990) elaboraram o artigo *Techniques to Control and Improve Quality of Data in Large Databases*.
- 1991 – Buckland *et al.* (1991) escreveram *Total Quality Management in Information Services*, uma adaptação do TQM para bibliotecas e serviços de informação.
- 1992 – Redman (1992) publicou o livro *Data Quality Management & Technology* no qual apresenta a idéia do ciclo de vida do dado, que foca na seqüência de atividades envolvidas da criação a distribuição dos dados (WANG *et al.*, 1995b).
- 1993 – Wang e Kon (1993) publicam *Towards Total Data Quality Management (TDQM)*.
- 1993 – Wang, Kon e Madnick (1993) editaram a obra *Data Quality Requirements Analysis and Modeling*, onde expõem como estabelecer metadados em bancos de dados para tratar a questão da qualidade da informação. Propõem que os atributos de qualidade são compostos por parâmetros (dimensão subjetiva) e indicadores (dimensão objetiva), onde os indicadores podem ser usados para qualificar os parâmetros.

1994 – Fox, Levitin e Redman (1994) elaboraram em 1992 o artigo *The Notion of Data and its Quality Dimensions* publicado em 1994, onde relatam que apesar da rápida proliferação dos sistemas de informação baseados em computadores evidenciarem a importância da qualidade dos dados, ainda não existe nenhum framework ou mesmo terminologia consolidada para investigar esta questão. O artigo então discute e apresenta uma definição para dados no contexto dos sistemas de informação baseados em computadores e discute as dimensões da qualidade de dados que segundo Redman (1992), são as mais importantes e estão agrupadas em quatro categorias: acurácia; completeza; consistência; e atualidade (Quadro 2.16). O conceito de dados é obtido do ponto de vista de uma atividade de modelagem aplicada a alguma parte do mundo, a partir da qual são gerados atributos (a) para os quais são definidos seus conjuntos de valores (v) e classes de entidades (e) que possuem um conjunto de atributos. Dessa forma, objetos específicos, suas propriedades e seus relacionamentos são modelados como triplas de entidades, atributos e valores, onde cada tripla em um modelo é um dado (*datum*), e dados (*data*) é uma coleção de triplas de dado. Como elementos de um modelo, os dados são abstratos e precisam ser representados de alguma forma. Uma representação de dados (*data representation*) é um conjunto de regras para gravar triplas em alguma mídia e um registro de dados (*data recording*) é uma instância dessa representação. Segundo os autores, essa caracterização dos dados sugere os três maiores focos de interesse da área de qualidade dos dados: a adequação do modelo; a adequação da representação e do registro; e a adequação das triplas no modelo de dados.

Quadro 2.16 – Dimensões da qualidade para valores de dados

Fonte: traduzido de FOX *et al.* (1994)

Dimensões	Descrição alvo (<i>target description</i>)	Medida típica do dado (<i>datum</i>)	Medida típica da base de dados	Noções relacionadas
Acurácia (<i>accuracy</i>)	Preciso ou correto (<i>accurate or correct</i>)	Tamanho do erro	Fração incorreta	Precisão, confiabilidade
Atualidade (<i>currentness</i>)	Atual (<i>current</i>)	Quantidade da defasagem	Fração defasada	Idade, conveniência temporal
Completeza (<i>completeness</i>)	Completo (<i>complete</i>)	S/N	Fração incompleta	Duplicação
Consistência (<i>consistency</i>)	Consistente (<i>consistent</i>)	S/N	Fração inconsistente	Integridade

- 1994 – Ruß-Mohl (1994) escreveu sobre a qualidade da informação no setor de jornalismo descrevendo sete critérios de julgamento; Objetividade (*Objectivity*), abrangência (*Comprehensibility*), relevância (*Relevance*), atualidade (*Currency*), redução da complexidade (*Reduction of Complexity*), transparência / flexibilidade (*Transparency / Reflexivity*) e interatividade (*Interactivity*) (EPPLER, 2006).
- 1995 – Tozer (1995) publica *Information Quality Management* onde adverte sobre os problemas de integração de dados entre sistemas com arquiteturas diferentes.
- 1995 – Wang, Reddy e Kon (1995a) publicam *Toward quality data: An attribute-based approach* onde propõem um modelo baseado em atributos com uma estrutura de dados formalizada com suas álgebras de consulta e regras de integridade que incluem os aspectos de qualidade dos dados no esquema do modelo relacional. Além disso, apresenta uma metodologia para análise dos requerimentos de qualidade de dados baseados no modelo relacional para especificar os tipos de indicadores de qualidade a serem modelados. Apresentam o conjunto de dimensões de qualidade exposto no Quadro 2.17.

Quadro 2.17 – Relações hierárquicas multidimensionais dos conceitos utilizados para avaliação da qualidade dos dados

Fonte: adaptado e traduzido de WANG *et al.* (1995a).

Dimensões 1º Nível	Dimensões 2º Nível	Dimensões 3º Nível
Acessibilidade (<i>accessibility</i>)	Disponibilidade (<i>availability</i>)	-
Interpretabilidade (<i>interpretability</i>)	Sintaxe (<i>syntax</i>)	-
	Semântica (<i>semantics</i>)	-
Utilidade (<i>usefulness</i>)	Relevante (<i>relevant</i>)	-
	Na hora certa (<i>timely</i>)	Atualidade (<i>currency</i>) Volatilidade (<i>volatility</i>)
Credibilidade (<i>believability</i>)	Completeza (<i>completeness</i>)	-
	Consistente (<i>consistent</i>)	-
	Reputação (<i>source is credible</i>)	-
	Acurácia (<i>accuracy</i>)	-

- 1995 – Wang, Storey e Firth (1995b) publicam *A Framework for Analysis of Data Quality Research* onde, através uma analogia entre a fabricação de produtos e a fabricação de dados, é desenvolvido um esquema (*framework*) baseado na ISO 9000, para analisar as pesquisas sobre a qualidade dos dados. Neste artigo é sugerida a existência de duas avenidas principais usadas para a definição das dimensões da qualidade dos dados: a

abordagem cientificamente fundamentada, que separa as dimensões intrínsecas de um sistema de informações daquelas externas ao sistema (ex. ontológica, teoria da informação, e marketing); e a abordagem pragmática, onde a qualidade é definida de modo operacional. Dentre suas conclusões, o artigo expõe que existe uma clara necessidade para o desenvolvimento de técnicas para o gerenciamento da qualidade dos dados, que os custos envolvidos precisam ser avaliados e que existe a necessidade de uma métrica da qualidade global dos dados e de uma forma rigorosa de expressar os requerimentos de qualidade para o projeto do produto dados.

- 1995 – Madnick (1995) publicou *Integration technology: The reinvention of the linkage between information systems and computer science* onde fala sobre o papel da comunidade de Tecnologia de Informação e Sistemas (*Information Technology and Systems*) que, por estar na interface dos campos da Ciência da Computação e dos Sistemas de Informação, pode atuar sobre os problemas que extrapolam os domínios tradicionais dessas duas áreas.
- 1995 – Lesca e Lesca (1995) publica *Gestion de l'information, qualité de l'information et performances de l'entreprise*, onde propõe tratar as dimensões da Qualidade da Informação sob perspectiva de produto – utilidade (*usefulness*), abrangência (*comprehensibility*), relevância (*relevancy*), completeza (*completeness*), representação adequada (*adequate representation*), coerência (*coherence*), clareza (*clarity*) – e processo – confiança (*Trustworthiness*), acessibilidade (*accessibility*), objetividade (*objectivity*), credibilidade (*credibility*), interatividade (*interactivity/feedback*). Este trabalho aconselha tratar os problemas de informação focando não apenas na informação gerada como produto final, mas também nos processos que geram esse produto informação (EPPLER, 2006).
- 1996 – Wang e Strong (1996) escreveram *Beyond Accuracy: what data quality means to data consumers*, onde, após identificar que existem três abordagens utilizadas para estudar a qualidade dos dados (a intuitiva, a teórica e a empírica), apresentaram os resultados de uma extensa pesquisa empírica realizada em duas fases para desenvolver um *framework* hierárquico para organizar as dimensões da qualidade dos dados. Essa abordagem assume implicitamente que os dados podem ser tratados como um produto e, segundo Eppler (2006), está claramente de acordo com a literatura existente sobre a qualidade da informação no contexto da tecnologia da informação. Como resultado final chegaram a 15 dimensões agrupadas em 4 categorias: Qualidade de Dados Intrínseca, que demonstra a qualidade inerente dos dados; Qualidade de Dados

Contextual, que destaca os requisitos de qualidade dos dados para a execução de uma tarefa; e Qualidade de Dados Representacional e de Acessibilidade, que refletem os requisitos de qualidade em relação ao sistema informatizado. O Quadro 2.18 apresenta as dimensões encontradas no estudo.

Quadro 2.18 – Categorias e Dimensões da qualidade dos dados
Fonte: WANG e STRONG (1996).

Categoria	Dimensão
Intrínseca	Acurácia (<i>Accuracy</i>), Objetividade (<i>Objectivity</i>), Credibilidade (<i>Believability</i>), Reputação (<i>Reputation</i>)
Acessibilidade	Acessibilidade (<i>Accessibility</i>), Segurança de acesso (<i>Access Security</i>)
Contextual	Valor agregado (<i>Value-Added</i>), Relevância (<i>Relevancy</i>), Conveniência temporal (<i>Timeliness</i>), Completeza (<i>Completeness</i>), Quantidade apropriada de dados (<i>Appropriate amount of data</i>)
Representacional	Interpretabilidade (<i>Interpretability</i>), Facilidade de Compreensão (<i>Ease of Understanding</i>), Representação Consistente (<i>Representational Consistency</i>), Representação Concisa (<i>Concise Representation</i>)

1996 – Wand e Wang (1996) escreveram *Anchoring Data Quality Dimensions in Ontological Foundations*, onde relatam que a qualidade dos dados normalmente é apresentada na literatura com conceitos multidimensionais, baseados no entendimento intuitivo, na experiência industrial ou na revisão de literatura. Por analogia às idéias do TQM, reconhecem que a qualidade dos dados depende do projeto e processos de produção envolvidos na geração dos dados. Constroem então uma proposta ontológica para as dimensões da qualidade dos dados com base em uma revisão da literatura, onde identificam os termos mais utilizados para qualidade dos dados. Como resultado, obtém quatro dimensões intrínsecas orientadas ao sistema – para ter qualidade os dados devem ser; completos, sem ambigüidade, significativos e corretos. Este modelo, apesar de conter uma definição rigorosa para raciocinar sobre a qualidade dos dados, não provê diretrizes concretas para os projetistas de sistemas.

1996 – English (1996) publica o artigo *Information Quality: Meeting Customer Needs*, onde trabalha a questão da qualidade a partir dos conceitos do TQM. Considera que a qualidade da informação envolve três componentes (definição, conteúdo e apresentação) e que usar o conceito de informação como um produto leva ao conceito da qualidade da informação como um serviço ao cliente. Esse trabalho tem clara preocupação com a implementação de um programa de qualidade, que demanda, entre outros pontos, uma reavaliação dos processos de negócio vigentes. Apresenta uma ótica gerencial para a questão da qualidade dos dados.

1996 – Redman (1996) publicou seu livro *Data quality for the information age*, que segue a literatura de Banco de dados e propõe 27 critérios em 9 dimensões, os quais estão apresentados no Quadro 2.19 (EPPLER, 2006). Uma descrição desses critérios também pode ser obtida em Tyson e McPartland (2000).

Quadro 2.19 – Perspectiva, Dimensões e Critérios da Qualidade dos Dados propostos por Redman

Fonte: Adaptado e traduzido de TYSON e MCPARTLAND (2000) e EPPLER (2006).

Perspectiva	Dimensão	Critérios
Visão Conceitual	Conteúdo	Relevância (<i>relevance</i>), facilidade de obtenção (<i>obtainability</i>), clareza de definição (<i>clarity of definition</i>)
	Escopo	Abrangência (<i>comprehensiveness</i>), essencialidade (<i>essentialness</i>)
	Nível de Detalhes	Granularidade de atributo (<i>attribute granularity</i>), precisão de domínio (<i>precision of domains</i>)
	Composição	Naturalidade (<i>naturalness</i>), de modo identificável (<i>identifiably</i>), homogeneidade (<i>homogeneity</i>), mínima redundância desnecessária (<i>minimum unnecessary redundancy</i>)
	Visão de Consistência	Consistência Semântica (<i>Semantic consistency</i>), Consistência Estrutural (<i>structural consistency</i>)
	Reação a Mudanças	Robustez (<i>robustness</i>), flexibilidade (<i>flexibility</i>)
Valores	–	Acurácia (<i>accuracy</i>), completude (<i>completeness</i>), consistência (<i>consistency</i>), atualidade / tempo de ciclo (<i>currency / cycle time</i>)
Representação	Formatos	Conveniência (<i>Appropriateness</i>), interpretabilidade (<i>interpretability</i>), precisão de formato (<i>format precision</i>), flexibilidade de formato (<i>format flexibility</i>), habilidade de representação de valores nulos (<i>ability to represent null values</i>), uso eficiente do armazenamento (<i>efficient use of storage</i>), portabilidade (<i>portability</i>)
	Instâncias Físicas	Consistência de representação (<i>representation consistency</i>)

1997 – Jarke e Vassiliou (1997) publicam *Data warehouse quality: a review of the DWQ project*, como fruto do *foundations for Data warehouse quality* (DWQ), um projeto cooperativo previsto para 3 anos (1996-1999) integrante do programa ESPRIT da comunidade europeia. Este trabalho defende que as dimensões de qualidade dos dados devem ser consideradas no projeto de um *data warehouse*. Os objetivos de pesquisa do DWQ estão direcionados a três domínios críticos, onde os fatores de qualidade são de importância central para um *data warehousing*: enriquecer a semântica de meta bancos de dados com modelos formais de qualidade da informação para permitir otimizações

adaptativas e quantitativas em projetos de *data warehouse*; enriquecer a semântica dos modelos de recursos de informação para permitir maior propagação de mudanças incrementais e resolução de conflitos; e enriquecer a semântica dos modelos esquemáticos de *data warehouse* para permitir aos projetistas e otimizadores de consultas usufruírem da natureza temporal, espacial e agregada dos dados de um *data warehouse*. Os autores utilizam um conjunto de dimensões da qualidade dos dados baseado no trabalho de Wang *et al.* (1995a) e adaptado para um ambiente de *data warehouse*; definem os conceitos de política de qualidade dos dados (*data quality policy*), gerenciamento de qualidade dos dados (*Data quality management*), sistema de qualidade dos dados (*data quality system*), controle de qualidade dos dados (*Data quality control*) e garantia de qualidade dos dados (*Data quality assurance*); e, por fim, destacam que pretendem desenvolver, ao longo do projeto, uma forma de modelar e medir a qualidade do *data warehouse* e entender como os fatores de qualidade podem ser combinados com a arquitetura de um *data warehouse*. O relacionamento entre os fatores de qualidade propostos e os aspectos de projeto um *data warehouse* são exibidos na Figura 2.9.

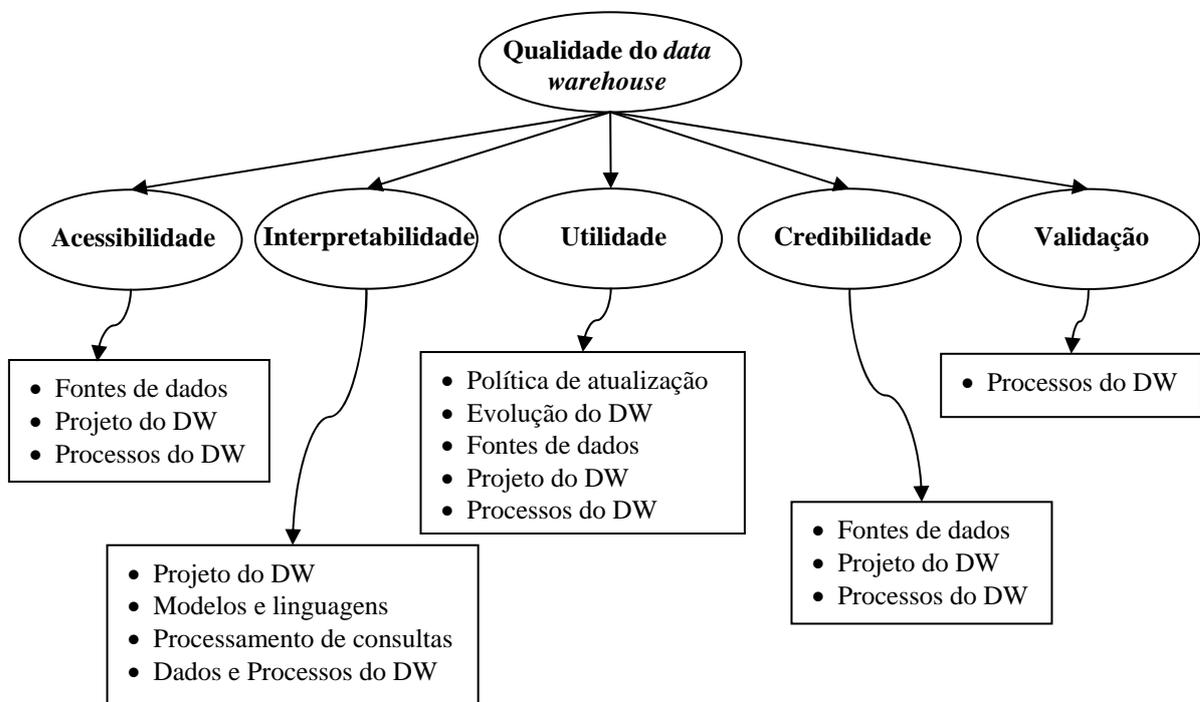


Figura 2.9 – Relacionamento entre os fatores de qualidade e as decisões de projeto de um *data warehouse*.

Fonte: JARKE e VASSILIOU (1997)

1998 – Königer e Reithmayer (1998) publicam *Management unstrukturierter Informationen*, onde apresentam uma abordagem genérica para a qualidade da informação como produto, no contexto de informações não estruturadas, e propõem 18 critérios em 6 dimensões para avaliação da qualidade da informação (Quadro 2.20).

Quadro 2.20 – Dimensões e critérios de qualidade da informação por Königer e Reithmayer
Fonte: EPPLER (2006)

Dimensão	Critério
1. Qualidade intrínseca	Precisão, Objetividade, Confiança
2. Qualidade de Acesso	Acessibilidade, Segurança
3. Qualidade Contextual	Relevância, Valor agregado, Conveniência temporal, Conteúdo de informação
4. Qualidade de Apresentação	Interpretabilidade, Compreensão, Concisão, Consistência
5. Qualidade de MetaInformação	Existência, Adequação
6. Qualidade de Construção	Existência, Adequação, Entendimento

1998 – Kahn e Strong (1998) publicam *Product and Service Performance Model for Information Quality*, onde a partir da visão de qualidade como conformidade com as especificações e qualidade como atender ou exceder as expectativas dos clientes e da visão que a informação possui características de produto e de serviço, produziram um modelo dois por dois, nomeado PSP/IQ. Este artigo foi apresentado pela primeira vez na conferência de Qualidade da Informação de 1997, por Kahn, Strong e Wang, e refinado com base nos comentários recebidos. Aqui é proposto que o uso conjunto do modelo PSP/IQ, da ferramenta IQA (*Information Quality Assessment*) para estimar a qualidade da informação em cada dimensão e de técnicas para a monitoração da QI por quadrantes e regras, provê uma metodologia útil para avaliação, benchmarking e aprimoramento da QI nas organizações. Segundo os autores, esta metodologia, chamada de AIMQ, provê uma base rigorosa e pragmática para monitoração de QI e benchmarks. Uma versão revisada do PSP/IQ foi publicada em Kahn *et al.*(2002) e o AIMQ foi apresentado também com o título “*AIMQ: A Methodology for Information Quality Assessment*” em Wang *et al.* (1998) e novamente em Lee *et al.* (2002). A listagem das dimensões de qualidade avaliadas no PSP/IQ é oriunda do trabalho de Wang e Strong (1996) com a diferença que as dimensões Acurácia (*Accuracy*) e Valor agregado (*Value-Added*) não foram consideradas e foram incluídas as dimensões Facilidade de operação (*Easy of operation*) e Livre-de-Erros (*Free-of-Error*), conforme o Quadro 2.21. Os autores explicam que a dimensão Valor agregado foi

abandonada por ser estatisticamente inconsistente com as outras dimensões do quadrante de Informação utilizável.

Quadro 2.21 – Dimensões da qualidade da Informação no modelo PSP/IQ (1998)
Fonte: KAHN e STRONG (1998).

	Conformidade com as Especificações	Atender ou exceder as expectativas do consumidor
Qualidade de Produto	Informação idônea Dimensões da QI: <ul style="list-style-type: none"> • Livre de erros • Representação Concisa • Completeza • Representação Consistente 	Informação útil Dimensões da QI: <ul style="list-style-type: none"> • Quantidade apropriada • Relevância • Facilidade de Compreensão • Interpretabilidade • Objetividade
Qualidade de Serviço	Informação fidedigna Dimensões da QI: <ul style="list-style-type: none"> • Conveniência temporal • Segurança de acesso 	Informação utilizável Dimensões da QI: <ul style="list-style-type: none"> • Credibilidade • Acessibilidade • De fácil operação • Reputação

1999 – Huang, Lee e Wang (1999) publicam o livro *Quality Information and Knowledge*, que interliga as questões de Qualidade da Informação com a Gestão do Conhecimento. Baseada em conceitos apresentados pelos autores em trabalhos anteriores, esta obra identifica três abordagens diferentes utilizadas para o estudo da qualidade da informação: intuitiva (baseada na experiência ou entendimento intuitivo); de sistema (examina deficiências na informação introduzidas por um sistema de informação); e empírica (baseada nas necessidades apontadas pelos usuários do sistema de informação). Os autores realizam seu estudo em um contexto mais abrangente do que os sistemas de informação, cobrindo os processos e procedimentos organizacionais e a análise dos papéis empregados para coletar, processar, distribuir e utilizar os dados. Dispõem de um método para aprimoramento da qualidade da informação, adaptado do TQM, composto de 5 fases: articular a visão da qualidade da informação em termos de negócios; estabelecer a responsabilidade da qualidade da informação; disseminar essa visão junto a clientes, fornecedores e produtores; ensinar novas habilidades vinculadas a esse contexto; e institucionalizar a melhoria contínua da qualidade da informação (CALAZANS, 2008).

1999 – Larry English (1999) escreve *Improving Data Warehouse and Business Information Quality*, que enfoca a qualidade dos dados (e de *data warehouses*) pela perspectiva do

TQM. Sua metodologia, inicialmente chamada TQdM (*Total Quality data Methodology*) e depois rebatizada como TIQM (*Total Information Quality Methodology*), prevê um ciclo com seis processos, onde cinco são de mensuração e aprimoramento e o último consiste em um processo guarda-chuva para criação de uma cultura de transformação capaz de gerar um ambiente de valor para os clientes da informação. São propostos 15 critérios em duas dimensões para avaliação da qualidade da informação (Quadro 2.22).

Quadro 2.22 – Dimensões e critérios de qualidade da informação na metodologia TIQM
Fonte: ENGLISH (1999).

Dimensão	Crítérios
Qualidade de Informação Inerente	Conformidade de definição
	Completeza (de valores)
	Validade ou conformidade com as regras de negócio
	Acurácia em relação a uma fonte substituta
	Acurácia (em relação à realidade)
	Precisão
	Ausência de duplicações
	Equivalência de dados redundantes ou distribuídos
	Concordância de dados redundantes ou distribuídos
	Acessibilidade
Qualidade de Informação Pragmática	Conveniência temporal
	Claridade contextual
	Integridade de derivação
	Utilidade
	Adequação (ou completeza de fato)

1999 – Naumann e Rolker (1999) publicam *Do metadata models meet IQ requirements?* onde comparam 23 critérios de qualidade da informação oriundos de 6 abordagens diferentes, os quais são divididos em quatro categorias: critérios relacionados ao conteúdo; critérios técnicos; critérios intelectuais; e critérios relacionados a instanciação.

1999 – Bobrowski, Marre e Yankelevich (1999) publicam *A Homogeneous framework to measure data quality*, onde propõe a aplicação da técnica GQM (*Goal, Question, Metrics*) de qualidade de software para medir a qualidade de dados em uma organização. As dimensões de medição são divididas em dois grupos, diretas e indiretas (Quadro 2.23), onde os valores das dimensões indiretas são calculados a partir dos valores encontrados para as dimensões diretas. Para cada dimensão diretamente acessível, é elaborada uma pergunta que a caracteriza e uma métrica para

responder a pergunta, e dessa forma é gerado um questionário para ser respondido com base em amostras do banco de dados.

Quadro 2.23 – Dimensões de qualidade da informação utilizadas com a técnica GQM
Fonte: BOBROWSKI *et al.* (1999).

Dimensões Diretas	Completeza, Relevância, Consistência, Exatidão, Conveniência Temporal e Precisão.
Dimensões Indiretas	Concisão (depende de Relevância), Confiabilidade (depende de Completeza, Concisão, Exatidão, Conveniência Temporal, Precisão e Consistência).

1999 – Alexander e Tate (1999) publicam o livro *Web wisdom: how to evaluate and create information quality on the web*, baseado na literatura de web design e usabilidade, contendo seis critérios para a qualidade da informação (Quadro 2.24).

Quadro 2.24 – Critérios de qualidade da informação para *Web*
Fonte: Alexander e Tate (1999) apud EPPLER (2006).

Critério	Explicação
1. Autoridade	Informação validada, a instituição que contém a informação é conhecida
2. Acurácia	Confiável, Livre de erros
3. Objetividade	Informação é apresentada sem influências pessoais
4. Atualidade	Conteúdo atualizado
5. Orientação a grupo alvo	Audiência alvo nitidamente sinalizada
6. Projeto de Interação & Navegação	Entendimento intuitivo dos elementos e suas funções.

2000 – Jarke, Lenzerini, Vassiliou e Vassiliadis publicam o livro *Fundamentals of Data Warehouses* (JARKE *et al.*, 2000), um dos subprodutos do programa europeu ESPRIT, onde propõe a integração de dimensões de qualidade dos dados no projeto de um *data warehouse*, utilizando uma abordagem quantitativa.

2000 – Naumann e Rolker (2000) publicam *Assessment methods for information quality criteria* onde identificam três classes de critérios para avaliação da qualidade da informação (de sujeito, de objeto e de processo) e nelas distribuem os critérios elencados em seu estudo anterior (NAUMANN e ROLKER, 1999), realizando três alterações na lista de critérios – foram retiradas acessibilidade (*Accessibility*) e desempenhabilidade (*Performability*) e incluído disponibilidade (*Availability*). As classes e critérios resultantes estão apresentados no Quadro 2.25.

Quadro 2.25 – Critérios de qualidade da informação
 Fonte: NAUMANN e ROLKER (2000)

Classe de Avaliação	Critério	Método de avaliação
Critérios de sujeito	Credibilidade (<i>Believability</i>)	Experiência do usuário
	Representação concisa (<i>Concise representation</i>)	Amostragem para usuário
	Interpretabilidade (<i>Interpretability</i>)	Amostragem para usuário
	Relevância (<i>Relevancy</i>)	Avaliação contínua do usuário
	Reputação (<i>Reputation</i>)	Experiência do usuário
	Compreensibilidade (<i>Understandability</i>)	Amostragem para usuário
	Valor agregado (<i>Value-Added</i>)	Avaliação contínua do usuário
Critérios de objeto	Completeza (<i>Completeness</i>)	Análise gramatical (<i>Parsing</i>), Amostragem
	Suporte ao cliente (<i>Customer Support</i>)	Análise gramatical (<i>Parsing</i>), contrato
	Documentação (<i>Documentation</i>)	Análise gramatical (<i>Parsing</i>)
	Objetividade (<i>Objectivity</i>)	Informação de especialista
	Preço (<i>Price</i>)	Contrato
	Confiabilidade (<i>Reliability</i>)	Avaliação contínua
	Segurança (<i>Security</i>)	Análise gramatical (<i>Parsing</i>)
	Conveniência temporal (<i>Timeliness</i>)	Análise gramatical (<i>Parsing</i>)
	Verificabilidade (<i>Verifiability</i>)	Informação de especialista
Critérios de processo	Acurácia (<i>Accuracy</i>)	Amostragem, técnicas de limpeza
	Quantidade de dados (<i>Amount of data</i>)	Avaliação contínua
	Disponibilidade (<i>Availability</i>)	Avaliação contínua
	Representação Consistente (<i>Consistent representation</i>)	Análise gramatical (<i>Parsing</i>)
	Latência (<i>Latency</i>)	Avaliação contínua
	Tempo de resposta (<i>Response time</i>)	Avaliação contínua

2001 – Loshin (2001) publica o livro *Enterprises knowledge management: The data quality approach*, que propõe uma metodologia para definir, medir e aprimorar a qualidade dos dados e gerenciar bases inteligentes de conhecimento ou de inteligência de negócios (*business intelligence*). Apresenta técnicas de análise de dados e assertivas de qualidade de dados e regras de negócio para avaliação da qualidade de dados, utilizando medições estáticas e dinâmicas, abrangendo tanto requisitos quanto metadados. Possui uma lista de dimensões para avaliação da qualidade dos dados,

apresentada no Quadro 2.26, dividida em cinco partes: Modelos de dados; Valores de dados; Domínio de dados; Apresentação de dados; e Política de informação. Entretanto, adverte que as listas de dimensões de qualidade dos dados são extensivas e nunca estão realmente concluídas, pois sempre pode surgir uma nova dimensão relevante.

Quadro 2.26 – Dimensões de qualidade da informação por Loshin
Fonte: LOSHIN (2001).

Modelos de dados	Clareza de definição (<i>clarity of definition</i>), Abrangência (<i>comprehensiveness</i>), Flexibilidade (<i>flexibility</i>), Robustez (<i>robustness</i>), Essencialidade (<i>essentialness</i>), Granularidade do atributo (<i>attribute granularity</i>), Precisão de domínios (<i>precision of domains</i>), Homogeneidade (<i>homogeneity</i>), Naturalidade (<i>naturalness</i>), Identificabilidade (<i>identifiability</i>), Facilidade de obtenção (<i>obtainability</i>), Relevância (<i>relevance</i>), Simplicidade/Complexidade (<i>simplicity/complexity</i>), Consistência semântica (<i>semantic consistency</i>), Consistência estrutural (<i>structural consistency</i>).
Valores de dados	Acurácia (<i>accuracy</i>), Valores nulos (<i>null values</i>), Completeza (<i>completeness</i>), Consistência (<i>consistency</i>), Atualidade/conveniência temporal (<i>Currency/timeliness</i>).
Domínio dos dados	Concordância de uso na empresa (<i>enterprise agreement of usage</i>), Gerenciamento (<i>Stewardship</i>), Ubiquidade (<i>Ubiquity</i>).
Apresentação de dados	Conveniência (<i>Appropriateness</i>), Interpretação correta (<i>correct interpretation</i>), Flexibilidade (<i>flexibility</i>), Precisão do formato (<i>Format precision</i>), Portabilidade (<i>Portability</i>), Consistência de representação (<i>representation consistency</i>), representação de valores nulos (<i>representation of null values</i>), uso de memória (<i>use of storage</i>)
Política de informação	Acessibilidade (<i>accessibility</i>), Metadados (<i>metadata</i>), Privacidade (<i>privacy</i>), Redundância (<i>redundancy</i>), Segurança (<i>security</i>), Custo unitário (<i>unit cost</i>)

2001 – Redman (2001) publica o livro *Data Quality: The Field Guide*, onde identifica três gerações de abordagens para a qualidade dos dados: a primeira consiste em inspeções e retrabalho para encontrar e corrigir os defeitos; a segunda, em prevenir os defeitos na origem; e a terceira, em projetar os sistemas para evitar os defeitos. Defende que os dados são constituídos por dois componentes inter-relacionados, o modelo de dados e os valores de dados, e que dados de alta qualidade são aqueles adequados para seus usos pretendidos em operações, processos decisórios e planejamento. Para isso, os dados devem ser livres de defeitos (acessíveis, acurados, atualizados, completos,

consistentes com outras fontes, etc.) e possuir características desejáveis (relevante, abrangente, com nível adequado de detalhes, fácil de ler, fácil de interpretar, etc.).

- 2001 – Wang, Ziad, e Lee (2001) publicam o livro *Data Quality*, que apresenta a metodologia TDQM e abrange conteúdos publicados em artigos anteriores. Sobre o TDQM, o livro contém: conceitos fundamentais; o *framework* do TDQM; a proposta de extensão do modelo relacional para inclusão dos atributos de qualidade (*polygen model* e *attributed-based model*) e dos requisitos de qualidade; uma forma de julgamento automático de qualidade dos dados; e a álgebra da qualidade dos dados. Além disso, fala sobre o projeto *Context Interchange* do MIT, o projeto de qualidade de *Data Warehouse* (DWQ) da União Européia e o projeto de qualidade dos dados da *Purdue University*.
- 2002 – Piattini, Calero e Genero (2002) publicam o livro *Information and Database Quality*, com contribuições de autores como Monica Bobrowski, Larry P. English, Beverly K. Kahn, David Loshin e Thomas C. Redman, cujo objetivo é apresentar uma visão geral sobre o tratamento da qualidade da informação, de acordo com os níveis indicados na Figura 2.10.

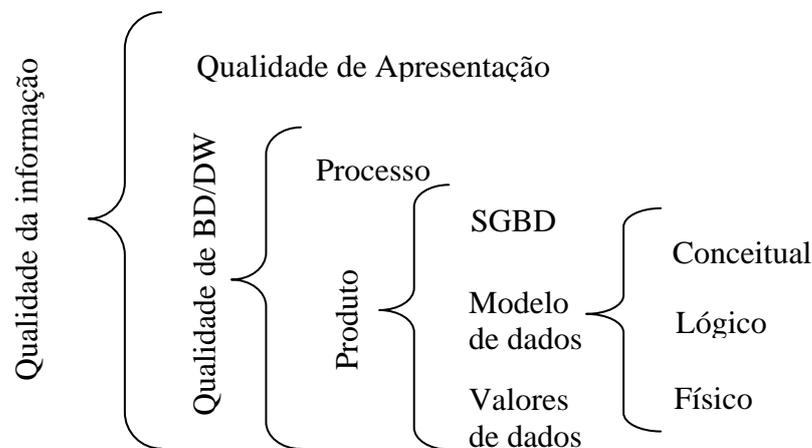


Figura 2.10 – Componentes da qualidade da informação.
Fonte: PIATTINI *et al* (2002)

- 2002 – Kahn, Strong e Wang (2002) publicam *Information Quality Benchmarks: Product and Service Performance*, uma atualização dos artigos divulgados em 1997 e 1998 com as seguintes alterações no conjunto de dimensões de qualidade da informação então utilizados: facilidade de operação (*Ease of operation*) foi renomeada como de facilidade de manipulação (*Ease of manipulation*); e valor agregado (*Value Added*),

excluído em 1998, foi reintegrado ao modelo PSP/IQ. O novo conjunto de dimensões é apresentado no Quadro 2.27.

Quadro 2.27 – Dimensões da qualidade da Informação no modelo PSP/IQ (2002)

Fonte: KAHN *et al.* (2002).

	Conformidade com as Especificações	Atender ou exceder as expectativas do consumidor
Qualidade de Produto	Informação Idônea Dimensões da QI: <ul style="list-style-type: none"> • Livre de erros • Representação Concisa • Completeza • Representação Consistente 	Informação útil Dimensões da QI: <ul style="list-style-type: none"> • Quantidade apropriada • Relevância • Facilidade de Compreensão • Interpretabilidade • Objetividade
Qualidade de Serviço	Informação fidedigna Dimensões da QI: <ul style="list-style-type: none"> • Conveniência temporal • Segurança de acesso 	Informação utilizável Dimensões da QI: <ul style="list-style-type: none"> • Credibilidade • Acessibilidade • De fácil manipulação • Reputação • Valor agregado

2002 – Lee, Strong, Kahn, Wang publicam *AIMQ: a methodology for information quality assessment* (LEE *et al.*, 2002), uma continuação do trabalho de Kahn e Strong (1998), onde comparam sete propostas de dimensões de qualidade da informação existentes na literatura e apresentam a metodologia AIMQ, composta por três componentes: o modelo PSP/IQ; o instrumento IQA; e a técnica de análise *Gap* para qualidade da informação. O instrumento IQA foi desenvolvido e validado para coletar dados sobre o *status* da qualidade da informação na organização, e foi utilizado para avaliar as dimensões de qualidade da informação do modelo PSP/IQ – as mesmas usadas em Kahn e Strong (1998). Por fim, a técnica de análise *Gap* foi aplicada para ponderar a lacuna existente entre a organização estudada e a utilizada como *benchmark*, por possuir as melhores práticas em relação às dimensões de QI analisadas.

2002 – Liu e Chi (2002) publicam *Evolutional Data Quality: A Theory-Specific View* na Conferência Internacional sobre Qualidade da Informação de 2002 do MIT, onde propõem que a qualidade de dados seja definida como um construto evolutivo que começa com a qualidade de coleta, passando pela qualidade de organização, pela qualidade de apresentação e, finalmente, pela qualidade de aplicação. Os autores entendem que os dados só podem existir e ter significado através de uma teoria e

nomeiam esse conceito de qualidade como específico-à-teoria (*theory-specific*). Para cada etapa, os autores elencam um conjunto de requisitos que devem ser observados. Para a qualidade de coleta são importantes a acurácia (*accuracy*), a objetividade (*objectivity*), a probidade do coletor (*trustworthiness of the collector*), a completeza dos dados coletados (*completeness*), a clareza (*clarity*) e os atributos específicos para a teoria de coleta de dados utilizada (*collection theory-specific attributes*). Para a qualidade de organização devem ser avaliadas a qualidade de coleta, a confiabilidade do digitador (*reliability of data clerk*), a consistência (*consistency*), a eficiência de armazenamento (*storage efficiency*), a eficiência de recuperação (*retrieval efficiency*), a navegabilidade (*navigability*) e os requisitos de qualidade de organização específicos para a teoria utilizada (*organization theory-specific quality*). Para a qualidade de apresentação devem ser avaliadas a qualidade de organização, a fidelidade (*faithfulness*), a neutralidade (*neutrality*), a interpretabilidade (*interpretability*), a formalidade (*formality*), a estabilidade semântica (*semantic stability*) e os requisitos de qualidade de apresentação específicos para a teoria utilizada (*presentation theory-specific quality*). Para a qualidade de aplicação devem ser avaliadas a qualidade de apresentação, a facilidade de manipulação (*ease of manipulation*), a conveniência temporal (*timeliness*), a privacidade (*privacy*), a segurança (*security*), a relevância (*relevancy*), se a quantidade de dados é apropriada para a tarefa (*appropriate amount of information*) e os atributos de aplicação específicos para a teoria utilizada (*application theory-specific attributes*).

2003 – Eppler publica *Managing Information Quality: Increasing the Value of Information in Knowledge-intensive Products and Processes*, que recebeu uma segunda edição (EPPLER, 2006), no qual aborda o gerenciamento da qualidade da informação como uma justaposição entre o gerenciamento da qualidade e a gestão do conhecimento. Relaciona vinte trabalhos da área de QI publicados de 1989 a 2000 e, com base na comparação das propostas de sete deles, indica a necessidade de um novo *framework* que enfoque as necessidades pertinentes ao conhecimento, especialmente no contexto de processos intensivos de conhecimento. Seu *framework* é composto de três elementos, um conjunto de quatro categorias de qualidade da informação, populadas por dezesseis critérios de qualidade e quatro princípios gerenciais. As quatro categorias são: informação relevante (*relevant information*), também chamada de visão da comunidade; informação idônea (*sound information*), com as características de produto da informação; processos otimizados (*optimized process*), com as

características de serviço da informação; e infra-estrutura confiável (*reliable infrastructure*), com as características relativas ao sistema de informações utilizado. Os dezesseis critérios estão relacionados no Quadro 2.28, divididos nas quatro categorias relacionadas acima, e as interdependências mapeadas entre esses critérios são apontadas na Figura 2.11.

Quadro 2.28 – Os 16 critérios de qualidade da informação e seus opostos por Eppler
Fonte: EPPLER (2006)

Nível de Qualidade da Informação	Critério de Qualidade da Informação	Critério Oposto
Nível de Comunidade (Relevância)	Abrangência (<i>comprehensiveness</i>)	Incompletude (<i>Incompleteness</i>)
	Acurácia (<i>accuracy</i>)	Deficiência de acurácia (<i>Inaccuracy</i>)
	Clareza (<i>clarity</i>)	Obscuridade (<i>Obscurity</i>)
	Aplicabilidade (<i>Applicability</i>)	Inutilidade (<i>Uselessness</i>)
Nível de Produto (Idoneidade – <i>Soundness</i>)	Concisão (<i>Conciseness</i>)	Prolixidade (<i>Prolixity</i>)
	Consistência (<i>consistency</i>)	Inconsistência (<i>Inconsistency</i>)
	Correção (<i>Correctness</i>)	Falsidade (<i>Falsity</i>)
	Atualidade (<i>Currency</i>)	Obsolescência (<i>Obsolescence</i>)
Nível de Processo	Conveniência (<i>Convenience</i>)	Inconveniência (<i>Inconvenience</i>)
	Conveniência temporal (<i>timeliness</i>)	Atraso (<i>Lateness</i>)
	Rastreabilidade (<i>Traceability</i>)	Indeterminação (<i>Indeterminacy</i>)
	Interatividade (<i>Interactivity</i>)	Rigidez (<i>Rigidity</i>)
Nível de Infra-estrutura	Acessibilidade (<i>accessibility</i>)	Inacessibilidade (<i>Inaccessibility</i>)
	Segurança (<i>security</i>)	Exposição (<i>Exposure</i>)
	Mantenabilidade (<i>Maintainability</i>)	Negligência (<i>Neglect</i>)
	Velocidade (<i>speed</i>)	Lentidão (<i>Slowness</i>)

Os quatro princípios gerenciais utilizados são: o princípio de integração, que sugere que as informações devem ser tão concisas quanto possível e que sejam consolidadas as fontes de informação dispersas; o princípio de validação, que sugere que as informações de alta qualidade devem ser avaliadas em termos de sua acurácia, consistência, conveniência temporal e segurança e que o resultado dessa avaliação seja registrado em meta-informações; o princípio do contexto, pelo qual os produtores de

informação e gerentes forneçam sempre um contexto para as suas informações (sobre o que é, como foi gerada, para o que ela serve e onde não é aplicável, e o que ou quem é necessário para mantê-la); e o princípio da ativação, que sugere que a informação deve ser explicitamente organizada para capacitar a ação, ou seja, visando seu processo de utilização. Eppler destaca também um ciclo de uso das informações (identificação, avaliação, alocação e aplicação) e um ciclo com três etapas, recomendado para o gerenciamento da qualidade da informação (analisar e medir, aprimorar e gerenciar, ensinar e manter). Por fim, ressalta que este é um trabalho destinado a estudantes e gerentes e que, por ter sido desenvolvido para atender vários propósitos e processos intensivos de conhecimento, sua aplicação demanda algumas customizações.

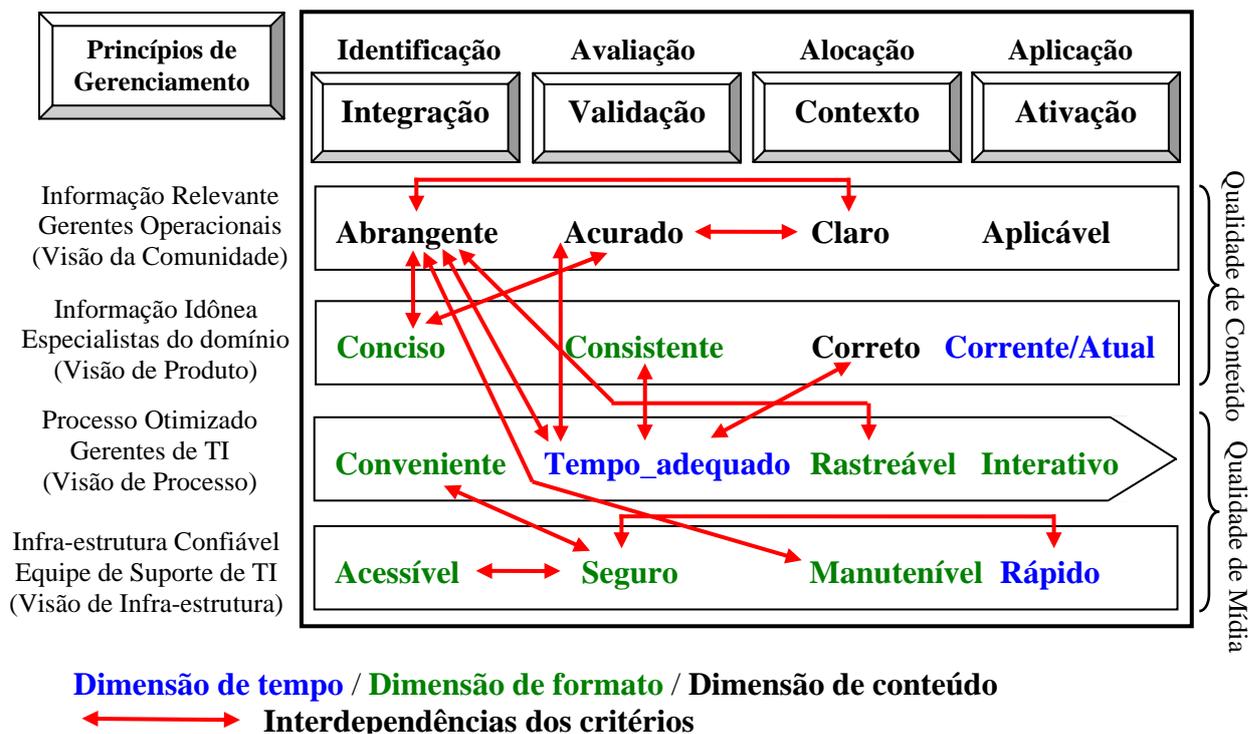


Figura 2.11 – Framework de qualidade da informação.
 Fonte: baseado em EPPLER (2006).

2003 – Olson (2003) publicou “Data Quality: The Accuracy Dimension”, cujos objetivos são demonstrar e caracterizar o problema de qualidade de dados; esboçar as funções de um grupo de garantia da qualidade de dados; e promover o uso de técnicas analíticas intensivas para dados (consideradas como componente fundamental de qualquer programa efetivo de garantia da qualidade dos dados). Sua abordagem é centrada nos dados visando dados estruturados presentes em bancos de dados corporativos e sua

análise é aplicada com um enfoque orientado de dentro-para-fora (estudando os problemas que podem surgir por falta de acurácia nos dados). Olson (2003) entende acurácia de dados como parte de um tópico maior, a qualidade de dados, que é vista da seguinte forma: os dados que têm qualidade são os que satisfazem os requisitos para o uso pretendido e os que carecem de qualidade são os que não satisfazem qualquer parte desses requisitos. Portanto, a qualidade dos dados depende tanto do uso pretendido quanto dos próprios dados, e, para satisfazer o uso pretendido, o dado precisa ser acurado, relevante, completo, entendido, confiável e estar disponível no tempo adequado. Dessas, a dimensão acurácia é considerada a medida fundamental da qualidade de dados, sendo composta por duas características, forma e conteúdo, com destaque para a importância da forma por ser capaz de eliminar ambigüidades sobre o conteúdo. A Figura 2.12 representa o subconjunto de dados com acurácia. Olson (2003) propõe a utilização de *data profiling* (Fig. 2.13), que consiste no emprego de técnicas analíticas para descobrir a verdadeira estrutura, conteúdo e qualidade de dados, e destaca que o conhecimento sobre os dados é a chave para uma bem sucedida avaliação, movimentação e uso dos dados. Para isso descreve um processo completo para formulação de regras de metadados e seu uso para uma avaliação eficiente dos dados e para o aprimoramento da completeza e acurácia dos próprios metadados. Por fim, defende a aplicação de regras para validação dos dados durante sua entrada no sistema, uma vez que eles podem ser inseridos por engano ou deliberadamente sem acurácia, podem ser o resultado de erros do sistema, podem ter sua acurácia deteriorada, podem perder sua acurácia devido a movimentações e reestruturações, ou mesmo podem tornar-se informações erradas por serem gravados ou usados de forma inapropriada. É através do entendimento de todos esses fatores que os profissionais de garantia de qualidade dos dados tornam-se mais especializados para analisar dados e processos na busca por dados sem acurácia. Os principais fatores para avaliação do processo de captura dos dados são: tempo decorrido entre o evento e seu registro; distância entre o evento e seu registro; número de transmissões informais da informação antes do seu registro; disponibilidade de todos os fatos durante o registro; habilidade para ratificar a informação durante o registro; motivação da pessoa que faz o registro; habilidade, treinamento e experiência da pessoa que faz o registro; retorno providenciado para a pessoa que faz o registro; suporte para o entendimento correto dos valores de dados; auto-ajuda no processo de registro; e verificação de erros no processo de registro.

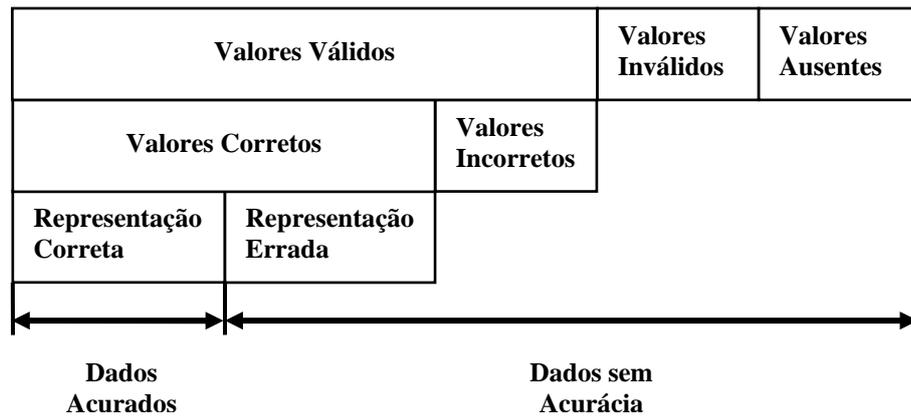


Figura 2.12 – Subconjunto de dados com acurácia.

Fonte: OLSON (2003), p. 33.

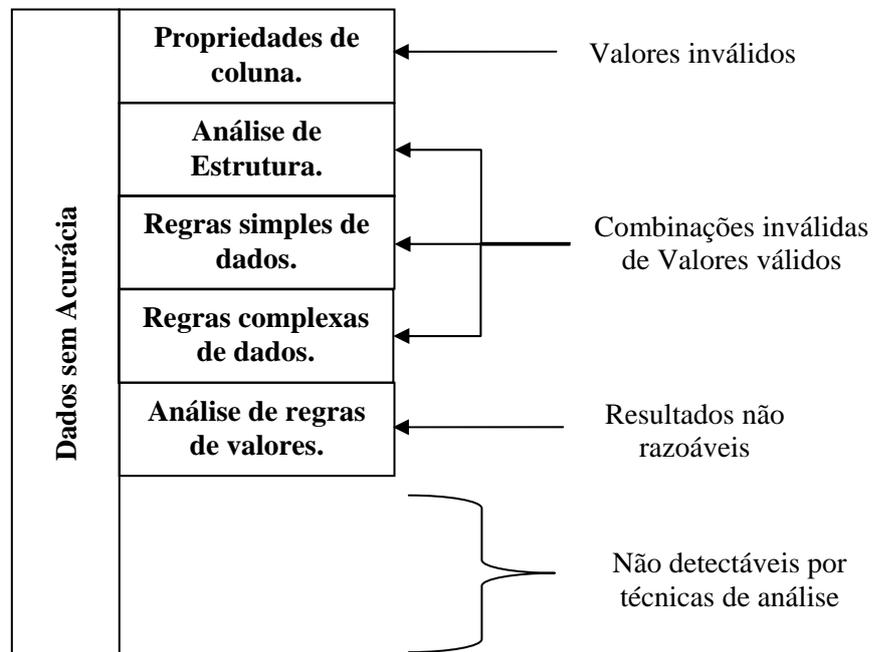


Figura 2.13 – Etapas de *data profiling*.

Fonte: OLSON (2003), p. 131.

2003 – Jarke, Lenzerini, Vassiliou e Vassiliadis publicam a segunda edição, revisada e estendida, de *Fundamentals of Data Warehouses* (JARKE *et al.*, 2003), onde oferecem um *framework* conceitual para avaliação e aprimoramento da arquitetura e da qualidade de um *data warehouse*, baseado em um gerenciamento enriquecido de metadados combinado com técnicas avançadas de bancos de dados, modelagem de negócio e inteligência artificial. Destacam que o conjunto de dimensões de qualidade varia de acordo com o ponto de vista utilizado e identificam as dimensões de qualidade aplicáveis para: o projetista e administrador (Evolução dos Metadados, qualidade do esquema, exatidão – *correctness*, completudeza – *completeness*,

minimalismo – *minimality*, rastreabilidade – *traceability* – e interpretabilidade – *interpretability*); o programador, obtidas da norma ISO/IEC 9126-1, que avalia a funcionalidade (adequação - *suitability*, acurácia - *accuracy*, interoperabilidade - *interoperability*, conformidade - *compliance*, e segurança de acesso – *security*), confiabilidade (maturidade - *maturity*, tolerância a falhas - *fault tolerance*, e recuperabilidade - *recoverability*), usabilidade (inteligibilidade - *understandability*, apreensibilidade - *learnability*, operacionalidade – *operability*), eficiência do programa (comportamento em relação ao tempo – *time behavior*, utilização de recursos – *resource behavior*), manutenibilidade (analisabilidade - *analyzability*, modificabilidade - *changeability*, estabilidade - *stability*, testabilidade - *testability*) e portabilidade (adaptabilidade - *adaptability*, capacidade para ser instalado - *installability*, coexistência – *Co-existence metrics*, capacidade para substituir - *replaceability*); e o usuário que irá tomar decisões a partir do *data warehouse*, divididas em acessibilidade – *accessibility* (disponibilidade do sistema – *system availability*, disponibilidade de transação – *transactional availability* e segurança – *security*) e utilidade – *usefulness* (interpretabilidade – *interpretability*, correspondência – *responsiveness* e conveniência temporal – *timeliness*, composta por atualidade – *currency* e volatilidade – *volatility*).

- 2004 – Perry, Signori e Boon (2004) publicam *Exploring the Information Superiority: A Methodology for Measuring the Quality of Information and its Impact on Shared Awareness*, como fruto de uma pesquisa conduzida no Centro de Política de Aquisições e Tecnologia do Instituto de Pesquisa de Defesa Nacional (*National Defense Research Institute – NDRI*) da Corporação RAND. O NDRI é um centro de pesquisa e desenvolvimento financiado em nível federal que presta serviços e é patrocinado pelo gabinete do secretário de Defesa, pelo Estado Maior Conjunto e pelas Agências de Defesa dos Estados Unidos. Este livro trata a questão da qualidade da informação aplicada ao *framework* de Guerra Centrada em Redes (*Network-Centric Warfare – NCW*) visando atingir melhores resultados nos combate em função da superioridade de informação. Para isso propõe um *framework* matemático que pode facilitar o desenvolvimento de medidas alternativas de desempenho e métricas associadas que avaliam a contribuição da qualidade da informação e da colaboração da equipe para a obtenção de uma consciência compartilhada da situação a ser enfrentada. Sua construção baseia-se no *framework* NCW e no modelo de Referência de Superioridade de Informação para avaliar os potenciais ganhos nos resultados de

combates que podem ser obtidos com uma arquitetura C⁴ISR (*Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance*), onde são considerados os fatores de Comando, Controle, Comunicações, Computadores, Inteligência, Vigilância e Reconhecimento. Os problemas de tratamento das informações são discutidos em três domínios: o de verdades básicas (o domínio físico); o de informações percebidas (o domínio de informação); e aquele no qual as consciências individuais da situação, a consciência compartilhada da situação e a colaboração e tomada de decisão acontecem (o domínio cognitivo). Este trabalho aborda apenas uma aplicação específica da arquitetura C⁴ISR, na qual uma matriz de sensores transmite dados sobre o campo de batalha a partir dos quais é gerado e distribuído o quadro operacional comum relevante (Common Relevant Operating Picture – CROP) resultante para os membros da força militar, que mentalmente interpretam o CROP recebido e colaboram com os outros membros da equipe para aprimorar a consciência de todos. O C⁴ISR é visto como um processo de extração de dados de verdades básicas e processamento no domínio da informação para produzir um CROP. A qualidade do CROP combina com a qualidade de colaboração da equipe para incrementar (ou degradar) a consciência compartilhada da situação no domínio cognitivo. As transformações são processos que incluem a coleta de dados, o processamento, a fusão de dados e a difusão de informação. É, ainda, gerada e analisada uma cadeia de valor para representar tanto os processos associados (com a cobrança, fusão e difusão da informação) como também o grau no qual a informação contribui para a consciência compartilhada da situação e, por fim, para a tomada de decisão, focando o valor que estes processos adicionam (ou subtraem) a cada passo. A qualidade da informação no *framework* NCW é descrita por oito atributos divididos em dois grupos de quatro, conforme o Quadro 2.29. A informação também é entendida como sendo orgânica, individual ou compartilhada. Para o escopo deste estudo, foram usadas apenas três medidas de qualidade da informação (completeza, correção e atualidade) com algumas adaptações em suas definições para a aplicação pretendida.

Quadro 2.29 – Atributos de qualidade da informação do *framework* NCW
 Fonte: PERRY *et al.* (2004)

Medida	Definição
Medidas Objetivas	São independentes do contexto
Correção (<i>Correctness</i>)	A extensão na qual a informação é consistente com as verdades básicas.
Consistência (<i>Consistency</i>)	A extensão na qual a informação está de acordo com as informações anteriores ou relacionadas.
Atualidade (<i>Currency</i>)	A idade da informação.
Precisão (<i>Precision</i>)	O nível de detalhes da medição em um item de informação.
Medidas de Adequação	A adequação ao uso é dependente do contexto
Completeza (<i>Completeness</i>)	A extensão na qual a informação relevante para verdades básicas é coletada (a relevância das verdades básicas é dependente do cenário)
Acurácia (<i>Accuracy</i>)	O quão apropriado é a precisão da informação para um uso particular
Relevância (<i>Relevance</i>)	A proporção da informação coletada que está relacionada à tarefa em mãos.
Conveniência Temporal (<i>Timeliness</i>)	A extensão na qual a atualidade da informação é apropriada para o seu uso.

2004 – Gackowski (2004) publica *Logical Interdependence of Data/Information Quality Dimensions - A Purpose-Focused View on IQ*, onde propõe um *framework* baseado em uma visão “orientada a propósito”, o qual identifica as dimensões de QI em uma taxonomia orientada a resultados. Após constatar que não existe ainda qualquer concordância entre os autores da área sobre o nível de importância, a seqüência de avaliação ou a completeza da lista de atributos relativos a QI, Gackowski propõe que as dimensões de QI possuem um inter-relacionamento e uma hierarquia que não devem ser desconsiderados para a avaliação da QI, e apresenta uma lista para ajudar a direcionar a atenção do analista sobre qual ordem de avaliação deve ser seguida, considerando o ponto de vista do propósito de negócio. Destaca, ainda, que mais importante do que os esforços gastos no desenvolvimento de métricas mais acuradas é a busca por um *framework* qualitativo, mais genérico e com mais força, para avaliação da QI. O *framework* proposto visa atender situações de decisão de negócios e sua aplicação requer quatro passos: 1– criar um diagrama de causa/efeito, qualitativo e relativamente completo, para identificar os principais fatores que impactam os resultados de negócio esperados, o que envolve atender ao menos algumas das expectativas e preferências dos clientes; 2– conduzir uma análise de impactos e força

dos fatores identificados; 3– desenvolver um modelo das informações da situação de decisão estudada que abranja tanto as informações conhecidas (*data component*) como as desconhecidas (*informational component*), as quais devem ser ranqueadas em relação ao seu impacto no resultado operacional usando qualquer medida estabelecida/combinada; 4 – com base no resultado dos passos anteriores, avaliar cada item de dados ou informação em relação ao seu uso prático. Gackowski destaca que sua proposta é orientada a propósito, diferente das demais existentes que são orientadas a produtos, serviços, preferências de usuários ou especificação de requisitos. Por fim, ao elencar as dimensões de QI consideradas relevantes, Gackowski as classifica como atributos primários, mandatórios ou desejáveis, e atributos secundários, as quais estão listadas no Quadro 2.30.

Quadro 2.30 – Exemplo de taxonomia orientada a resultados de atributos de qualidade de dados ou informação em uma seqüência econômica de avaliação
Fonte: GACKOWSKI (2004)

			Exemplos de atributos de qualidade de dados/informação		
			Atributos diretos	Atributos indiretos	
Seqüência de avaliação	 Atributos Primários	Atributos mandatórios	Interpretável	Legível, usuário treinado, não treinado, educado	
			Significativamente relevante	Conciso, atual, admissível, seguro, quantidade apropriada	
			Criticidade da conveniência temporal	Obtenível, acessível, estilo e modo de tomada de decisão, individual ou coletivo	
		Atributos desejáveis	Criticamente acreditável	Acreditável, respeitável, características do tomador de decisões: aversão a risco, passivo, hesitante, cauteloso, prudente, motivado, nervoso	
			Aceitavelmente completo (totalidade de fatores)		
		 Atributos Secundários	-	Economicamente	Na hora certa
	Imparcial				Amostragem, pontos de observação
	Acurado (livre de erros)				Mapeamento (completo, sem ambigüidades, significativo e correto), granularidade, idade
	Preciso				Número de dígitos significativos, pontos
	Irrelevante			Fácil de usar	Resumido de que forma, detalhado, texto, gráfico, diagrama, figura, mídia, clareza, ordem, consistência, homogêneo, inteligível, naturalidade, eficiência de codificação

- 2005 – Gackowski (2005a) publica *Informing Systems in Business Environments: A Purpose-Focused View*, onde analisa os problemas encontrados em sistemas informativos em ambientes empresariais sob o ponto de vista dos disseminadores de informação e com base em uma breve revisão da literatura de Sistemas de Informações Gerencial. Segundo Gackowski, este artigo também traz uma versão aprimorada do *framework* orientado ao propósito apresentado em Gackowski (2004).
- 2005 – Gackowski (2005b) publica *Operations Quality of Data and Information: Teleological Operations Research-Based Approach, Call for Discussion* como uma tentativa teórica de descobrir no vasto universo das dimensões de qualidade dos dados e informações, um conjunto que seja obrigatório e princípios que tenham validade mais duradoura. O artigo conclui que, independente de qualquer situação, há cinco requisitos universais obrigatórios de qualidade de informação referentes a necessidades primárias, que avaliam o quanto: é interpretável durante sua aquisição (*acquisition interpretability*); é de impacto significativo (*of significant impact*); está disponível em um tempo conveniente e oportuno para a operação (*operationally timely availability*); possui um nível aceitável de credibilidade (*actionable credibility*); e o conjunto de dados ou informações é considerado completo para a tarefa a ser executada (*task-specific effective operational completeness of sets of usable data/information values*). Qualquer deficiência nesses atributos muda a situação de decisão conforme definido na ciência de decisão. O artigo considera também como um requisito primário mandatório de qualidade, a avaliação do quão interpretável é a apresentação dos dados ou informação (*presentation interpretability*), mas o trata em separado, pois entende que sua aplicabilidade limita-se a um aspecto informativo indireto. Por fim, quatro relações custo/benefício são identificadas como requerimentos diretos e secundários de qualidade: de ser interpretável durante a aquisição (*Economically interpretable during acquisition*); de estar disponível em um tempo conveniente e oportuno para a operação (*Economically operationally timely available*); de possuir um nível aceitável de credibilidade (*Economically actionably credible*); e de possuir um conjunto de dados ou informações considerado completo para a tarefa a ser executada (*Task-specific economically operationally complete*). São requerimentos de natureza econômica que influem quantitativamente no resultado das operações, mas que podem ser mandatórios quando o fator econômico for uma preocupação primordial.
- 2005 – Wang, Pierce, Madnick, e Fisher publicam *Information Quality* (WANG *et al.*, 2005), livro que abre a série *Advances in Management Information Systems* (AMIS) cujo

objetivo é codificar o campo de gerenciamento de sistemas de informações como uma disciplina de pesquisa e prática. Esta obra apresenta um artigo introdutório e mais 14 outros artigos sobre: medição de qualidade de dados; modelagem e desenvolvimento de processos de informação para qualidade da informação; casos de estudo sobre aprimoramento da qualidade de dados e informação; problemas organizacionais de qualidade da informação; e educação e construção de capacitação em qualidade da informação.

- 2005 – Fisher, Lauría, Chengalur-Smith, e Wang publicam *Introduction to Information Quality* (FISHER *et al.*, 2005) – reimpresso mais 3 vezes, sendo a 4^a impressão em 2008 – que pretende despertar as pessoas sobre os problemas críticos de qualidade de dados e informação que têm infestado os sistemas de informação por muitos anos. Este livro foi construído sobre duas fontes primárias: um questionário sobre a importância de conhecimentos e habilidades em qualidade de dados respondido por 110 pesquisadores e praticantes de qualidade de dados na Conferência Internacional em Qualidade de Informação organizada pelo MIT; e as pesquisas de qualidade de dados e informação conduzidas pelos autores do livro (cuja soma dos artigos publicados já passa de 100). Desse modo foram identificadas as habilidades consideradas mais críticas para a qualidade de informação, as quais compõem os tópicos primários deste livro. Embora escrito visando estudantes de áreas como sistemas de informação (SI), tecnologia de informação (TI) e análise de sistemas e negócios, este livro pode ser utilizado também pelas áreas de marketing, economia, contabilidade e gestão de empresas. Os seus autores defendem que a qualidade de dados e informação deve ser reconhecida e tratada como uma disciplina organizada e independente. Os primeiros quatro capítulos do livro provêm uma base abrangente para o entendimento dos conceitos e filosofia de qualidade de dados e informação. Os demais capítulos apóiam-se sobre estes conceitos para apresentar ferramentas e técnicas essenciais para que um analista de qualidade de dados possa realizar aprimoramentos de qualidade.
- 2006 – Lee, Yang, Pipino, Funk e Wang publicam *Journey to Data Quality* (LEE *et al.*, 2006), abordando vários tópicos sobre a qualidade de dados. Nele relatam que todas as organizações enfrentam hoje problemas de qualidade de dados, tanto sistêmicos quanto estruturais, e que nem as abordagens *ad hoc* e nem os ajustes no sistema podem resolver o problema básico de práticas ruins de qualidade de dados. O livro objetiva servir de guia para o planejamento e implantação de programas viáveis de qualidade de dados e informação, podendo ser usado por praticantes, executivos e estudantes.

Para isso discorrem sobre análises de custo/benefício, avaliação da qualidade dos dados, amostragem para garantia da qualidade dos dados, a anatomia dos problemas e padrões da qualidade dos dados, identificação de causa raiz dos problemas de qualidade dos dados (a partir de um estudo de caso de uma organização de saúde), gerenciamento da informação como produto, desenvolvimento de mapas de produto informação, iniciativas da qualidade de dados a partir de um estudo de caso de um hospital escola, políticas de qualidade de dados e futuros desafios.

2006 – Batini e Scannapieco (2006) publicam *Data Quality: Concepts, Methodologies and Techniques (Data-Centric Systems and Applications)*, tendo como público alvo pesquisadores das áreas de bancos de dados e sistemas de informação interessados em investigar as propriedades de dados e informação que têm impactos na qualidade de processos e na vida real. A meta deste livro é prover uma descrição sistemática e comparativa do vasto número de questões de pesquisa relacionadas à qualidade de dados e, com base nesse estudo, tentar responder as seguintes perguntas: O que é, em essência, qualidade de dados? Quais técnicas, metodologias e questões de qualidade de dados já estão consolidadas? Quais são as abordagens conhecidas e confiáveis? E quais questões continuam em aberto? Os autores ressaltam que existe uma lacuna entre as abordagens com enfoque mais prático, que não se preocupam muito com a questão conceitual, e as abordagens mais formais, orientadas a questões de pesquisa, mas que não tratam tópicos ou perspectivas específicas. Em função disso, os resultados encontrados na literatura ainda são fragmentados e escassos, sendo notória a ausência de uma visão sistemática da área. Assim, este livro se estabelece como um complemento aos demais já publicados, provendo não apenas uma visão comparativa geral e um *framework* explanatório das propostas existentes, como também traz soluções originais que combinam a solidez de abordagens práticas com a sonoridade de formalismos teóricos e está estruturado da seguinte forma: os dois primeiros capítulos são preparatórios para os demais, com o primeiro apresentando conceitos básicos e estabelecendo coordenadas para explorar a área de qualidade de dados e o segundo abordando as dimensões que permitem a medição da qualidade de valores e esquemas de dados. No capítulo 3, são investigados modelos para expressar a qualidade de dados em bancos de dados e sistemas de informação. O capítulo 4 descreve as atividades principais para medição e aprimoramento da qualidade de dados, introduzido e descrevendo algumas atividades, como localização e correção de erro. Os dois capítulos seguintes são dedicados às atividades e áreas de pesquisas mais

importantes relacionadas ao aprimoramento da qualidade de dados, especialmente a identificação de objeto (capítulo 5) e integração de dados (capítulo 6), que são investigadas extensivamente a partir das perspectivas de paradigmas de pesquisa relevantes e técnicas disponíveis. No capítulo 7 são examinadas e comparadas as metodologias existentes para a medição e aprimoramento da qualidade de dados, abrangendo dimensões, modelos, atividades e técnicas, e então é proposta uma metodologia original e abrangente e é apresentado um caso de estudo com sua aplicação. O capítulo 8 descreve algumas ferramentas, *frameworks* e caixas de ferramentas (*toolboxes*) existentes para apoiar as metodologias propostas na literatura e o capítulo 9 finaliza o livro colocando em perspectiva todas as idéias discutidas nos capítulos anteriores e especulando sobre os problemas em aberto e as possíveis evoluções da área.

2007 – Al-Hakim (2007) publica *Information Quality Management: Theory And Applications*, onde reúne trabalhos de diversos autores e opina que a qualidade de informação finalmente está se firmando rapidamente como uma disciplina importante e distinta. Endossa a seguinte visão sobre dados e informações: dados são itens numéricos, alfanuméricos, figuras, sons ou imagens, sobre coisas, eventos atividades e transações, registrados, classificados e armazenados, mas não organizados com o objetivo de transmitir qualquer significado específico; e informações são dados organizados de forma a prover significado para o receptor, que pode ser algo já conhecido ou ter um valor surpreendente por revelar algo não conhecido. Essa visão é consistente com conceito de informação como o produto de um sistema de produção de informações, onde na entrada são recebidos os dados necessários para a produção da informação desejada. É destacada que apesar das similaridades usadas para comparar a produção de informações com a de produtos tangíveis, existem diferenças que podem ser classificadas em cinco tópicos: intangibilidade; entrada; usuários; consumo; e manuseio (Quadro 2.31). Ainda, considerando que uma informação pode precisar de várias etapas para ser gerada, o termo informação pode ser usado para referenciar tanto dados como informações. Entretanto, Al-Hakim (2007) estabelece uma diferença entre qualidade de dados e qualidade de informação, argumentando que para obter qualidade de informação é preciso conseguir qualidade de dados, mas que apenas ter qualidade de dados não necessariamente garante a qualidade de informações, pois podem ocorrer erros no processo de transformar dados em informações. Neste livro, quando os autores usam o termo qualidade de dados é porque estão focando apenas nos

Quadro 2.31 – Principais diferenças entre fabricação de produtos e produção de informações
 Fonte: AL-HAKIM (2007)

Item	Diferença
Intangibilidade	Um sistema de produção industrial fabrica produtos tangíveis, visíveis ou físicos enquanto que a informação é intangível. A qualidade de produto pode ser avaliada por medidas físicas como as especificações de projeto. As medidas para qualidade de informação são subjetivas e baseadas principalmente na opinião e expectativa do usuário.
Entradas	Um processo de produto requer matéria-prima, experiência/conhecimento e tecnologia; enquanto um processo de informação requer quatro entradas: dados, experiência, tecnologia e tempo.
Usuário Final	O usuário de um sistema de informação faz parte desse sistema, enquanto os produtos são produzidos longe dos seus usuários.
Consumo	Dados atuam como matéria-prima para fabricação de informações que podem ser consumidas por mais de um consumidor sem depleção, diferente das matérias-primas em sistemas de manufatura que se esgotam à medida que são usadas em produtos físicos. Além disso, a informação pode ser produzida e consumida simultaneamente, enquanto os produtos precisam ser produzidos antes de serem consumidos.
Manuseio	Diferente dos produtos, os mesmos dados e informações podem ser transportados simultaneamente a um número indefinido de consumidores através de um transporte físico, como um disco, ou por uma mídia intangível, por exemplo, e-mail. Porém, tanto a informação quanto os produtos podem ser armazenados e inspecionados antes da entrega aos clientes. Isto torna a qualidade de informação semelhante à qualidade do produto, mas diferente de qualidade de serviço que não pode ser armazenada e inspecionada antes da entrega.

problemas de dados enquanto matéria bruta (por exemplo, a qualidade dos dados de alimentação de um *data warehousing*) e quando se referem a qualidade de informação é porque estão tratando também dos processos de transformação dos dados. A QI é multidimensional, por isso são necessárias múltiplas medições para avaliar a qualidade da informação ou mesmo dos dados. Este livro lida tanto com os aspectos teóricos da QI quanto com suas aplicações, sendo útil para profissionais e pesquisadores que trabalham com gerenciamento da informação e do conhecimento e para gerentes e praticantes da indústria e serviços preocupados com o gerenciamento da informação.

Possui doze capítulos organizados em cinco sessões cobrindo vários problemas teóricos e práticos. A 1ª sessão possui quatro capítulos que tratam de problemas de processamento em qualidade da informação, incluindo a criação de métricas de qualidade da informação, processamento de consultas e atributos de representação simbólica. A 2ª sessão, com dois capítulos, trata sobre avaliação e aprimoramento da qualidade da informação. A 3ª sessão tem dois capítulos sobre a importância do mapeamento de processos para obtenção de qualidade da informação. A 4ª sessão possui dois capítulos que versam sobre aplicações de qualidade da informação na indústria (relativa ao projeto e fabricação de um produto) e na gestão (gerenciamento de ativos de engenharia). A 5ª sessão, também com dois capítulos, trata sobre aplicações na China e Malásia, dois países em desenvolvimento.

2008 – Loshin (2008) publica *Master Data Management*, que visa lidar com o problema dos diversos dados incongruentes espalhados em uma organização. A intenção do programa de gerenciamento de dados mestres (*Master Data Management – MDM*) é criar um repositório único de dados mestres de alta qualidade para alimentar as aplicações ao longo da organização com uma visão consistente e sincronizada dos dados da empresa. O autor destaca que MDM é uma Técnica e não uma ferramenta, pois envolve Tecnologia, Processos e Pessoas. Um programa de MDM incorpora aplicações de negócio, métodos de gerenciamento da informação e ferramentas de gerenciamento de dados para implantar as políticas, procedimentos e infra-estrutura que suportam a captura, integração e subsequente uso compartilhado de dados mestres com qualidade. O autor destaca que MDM envolve governança (políticas, procedimentos, infra-estrutura), ação (captura, integração, compartilhamento) e qualidade (acurácia, conveniência temporal, completeza) e enfatiza que apenas o uso de tecnologia não é suficiente para o seu êxito. Para que um programa de MDM tenha sucesso, é fundamental obter o reconhecimento de sua necessidade pelos altos níveis da companhia e então gerar políticas e governança que suportem o programa. É preciso que a governança proveja uma clara atribuição de responsabilidades quanto às informações e identifique seu relacionamento com os objetivos do negócio. Este livro está organizado da seguinte forma; o primeiro capítulo é sobre dados mestres e gerenciamento de dados mestres (*Master Data Management – MDM*), onde é demonstrado como a evolução do uso das Tecnologias de Informação levou a necessidade da existência de um programa de MDM. O segundo capítulo fala sobre *stakeholders*, requisitos e planejamento para o sucesso de um programa de MDM. O

terceiro capítulo fala sobre os componentes de MDM e o modelo de maturidade de forma a identificar o nível de maturidade atual da organização, o nível pretendido e traçar um mapa atingi-lo. O quarto capítulo explana sobre governança de dados para MDM. O quinto capítulo trata sobre qualidade de dados e MDM, abrangendo as seguintes dimensões: unicidade (*uniqueness*); acurácia (*accuracy*); consistência (*consistency*); completeza (*completeness*); conveniência temporal (*timeliness*); atualidade (*currency*); conformidade com formato (*format compliance*); e integridade referencial (*referential integrity*). O sexto capítulo versa sobre gerenciamento de metadados para MDM. O sétimo capítulo aborda a identificação dos dados mestres e seus metadados. O oitavo capítulo fala sobre modelagem de dados para MDM. O nono capítulo tece considerações sobre paradigmas de arquiteturas para MDM. O décimo capítulo explana os processos de integração e consolidação de dados. O décimo primeiro capítulo trata sobre sincronização de dados mestres. O décimo segundo capítulo é sobre MDM e a camada de serviços funcionais que visa integrar algumas funcionalidades entre as diversas aplicações que usam os dados mestres. O décimo terceiro capítulo conclui o livro com algumas orientações gerenciais para MDM.

2008 – McGilvray (2008) publica *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information (TM)* onde apresenta o seu processo de dez passos para a qualidade de dados. O livro é estruturado da seguinte forma. O capítulo 1 fala sobre o Impacto da Informação e Qualidade dos Dados, sobre a Metodologia, sobre as abordagens para a Qualidade dos Dados em Projetos e sobre Gestão de Engajamento. O Capítulo 2 é sobre os conceitos chaves e apresenta um framework para a Qualidade da Informação (*Framework for Information Quality – FIQ*), o Ciclo de vida da informação, as dimensões da qualidade dos dados, as técnicas de impacto nos negócios, as categoria dos dados, a especificação dos dados, o gerenciamento e governança de dados, o ciclo de aprimoramento da qualidade de dados e informação, o processo de dez passos (*The Ten Steps Process*), e as melhores práticas e diretrizes. No capítulo 3 é detalhado o processo de dez passos, que consiste em: 1- Definir a necessidade do negócio e a abordagem – onde devem ser priorizados os problemas de negócio e planejado o projeto; 2- Analisar o ambiente de informação – onde devem ser entendidos os pontos relevantes quanto a requisitos, especificações e dados, tecnologias, processos, organizações e pessoas, para então ser delimitado o ciclo de vida da informação e projetada a captura dos dados e o plano de avaliação; 3- Avaliar a qualidade de dados – incluindo a especificação e os fundamentos de integridade dos

dados e sua avaliação quanto a duplicação, acurácia, consistência e sincronização, disponibilidade e conveniência temporal, facilidade de uso e manutenibilidade, cobertura (de dados), qualidade de apresentação, percepção, relevância, confiança, decaimento de qualidade (dos dados) e o grau no qual o dado produz a saída (transação de negócio) desejada (*Transactability*); 4- Avaliar o impacto nos negócios – através de anedotas, identificação de usos, utilização dos cinco porquês e da matriz de custos X benefícios para a classificação e priorização, considerando os impactos identificados no processo, o custo da baixa qualidade de dados e a análise de custo-benefício; 5- Identificar as Causas Raiz – com o uso dos cinco porquês, da técnica de rastrear e planejar (*Track and Trace*) e do diagrama de espinha de peixe para análises de causa e efeitos; 6- Desenvolver planos de aprimoramento; 7- Prevenir futuros erros de dados; 8- Corrigir os erros de dados atuais; 9- Implementar controles; 10- Comunicar ações e Resultados. O capítulo 4 versa sobre a estruturação do projeto, pontos de atenção em relação ao processo de dez passos (*The Ten Steps*), as regras de projetos de qualidade de dados, e o timing do projeto. O capítulo 5 comenta sobre outras técnicas e ferramentas, englobando as abordagens de ciclo de vida das informações, a captura de dados, a análise e documentação de resultados, métricas, ferramentas de qualidade de dados e o relacionamento entre o processo de dez passos (*The Ten Steps*) e o seis sigma. E o capítulo 6 conclui o livro com outras considerações sobre o assunto.

2008 – O'Brien (2008) publica *Implementing an Effective Data Quality Strategy*, onde compara algumas metodologias da literatura e discute questões relevantes para a implementação de um programa de qualidade da informação. O capítulo 1 introduz a questão da qualidade de dados. O capítulo 2 apresenta o cenário atual no contexto da Era da Informação e os condutores principais (*key drivers*). O capítulo 3 evidencia as estruturas (*frameworks*) padrões ISO 27002:2005 (vulgo ISO 17799:2005) e COBIT, sob a perspectiva de qualidade de dados, e os Padrões ISO emergentes para qualidade de dados. O capítulo 4 apresenta a visão geral de algumas metodologias, como a de Tom Redman, a de Larry P. English, e a de Danette McGilvray, além de temas comuns e percepções (*insights*). O capítulo 5 trata sobre a definição de estratégias efetivas para o sucesso de qualidade de dados, abrangendo tópicos como: mudança de gestão, descontentamento com o *status quo*, escolha dos condutores corretos a serem enfocados, fontes de descontentamento, quantificação de descontentamento, visão de futuro alternativo, primeiro passos, governança e maturidade, melhoramento contínuo,

custo da não-qualidade e comparação da informação com outros ativos. O capítulo 6 propõe uma estratégia para qualidade de dados, desmembrada em introdução, implementação, evolução, alinhamento e conclusão. O capítulo 7 discorre sobre a importância da linguagem, com tópicos sobre Programação Neurolingüística, sub-modalidades gerenciais e construção de entendimento gerencial. O capítulo 8 trata sobre alguns obstáculos e barreiras que precisam ser superados. O capítulo 9 traz uma análise final sobre o assunto de qualidade de dados e no capítulo 10 são apresentados dois cenários de estudos de caso, um da empresa BetaCorp e outro da AlphaTech.

2008 – Redman (2008) publica *Data Driven: Profiting from Your Most Important Business Asset*, que visa despertar os gerentes sobre a necessidade de lidar adequadamente com os dados e informações nas empresas, pois, apesar de sua alardeada importância, eles ainda não são tratados como um ativo. Com termos simples, esse livro convida a pensar sobre qualidade, focando na prevenção dos erros nas suas origens, e sugere que, na implantação de um programa de qualidade, seja incorporado ao menos um dos dez hábitos listados a seguir: 1– concentre-se nas necessidades mais importantes dos clientes mais importantes; 2– aplique rígida atenção ao processo; 3– gerencie todas as fontes críticas de dados, incluindo fornecedores; 4– meça a qualidade na origem e em termos de negócio; 5– empregue controles em todos os níveis para estancar erros simples e estabelecer uma base para seguir adiante; 6– desenvolva um “jeito” para o melhoramento contínuo; 7– estabeleça e atinja alvos agressivos para o aprimoramento; 8– formalize as responsabilidades gerenciais pelos dados; 9– lidere o esforço usando o grupo sênior da empresa; 10– reconheça que os problemas mais difíceis de qualidade de dados são ligados a parte *soft* da gestão da qualidade (pessoas, política, cultura e organização) e gerencie ativamente as mudanças culturais necessárias. Ainda, como a implantação de um programa de qualidade implica em mudança cultural, propõe que esta seja pensada usando um modelo simples, que consiste em entender e aplicar os quatro elementos a seguir: 1– senso de urgência (porque isso merece meu tempo e atenção, uma vez que já estou trabalhando duro?); 2– visão compartilhada nítida (se eu fizer essa mudança, como será esse novo mundo que vai surgir?); 3– primeiro passo atuável (então o que você quer que eu faça diferente a partir de amanhã de manhã?); 4– capacidade para mudança (deve ser providenciado tempo, treinamento, dinheiro, e qualquer outro recurso necessário para que as pessoas realmente possam contribuir).

2.5.5 Algumas lições sobre medição de qualidade da informação

Conforme destacado anteriormente, quando se busca medir um conceito que ainda não foi alcançado, qualquer abordagem realizada não será capaz de englobar todas as implicações envolvidas, uma vez que a abrangência teórica do conceito ainda não foi determinada (NEHMY e PAIM, 1998). Não obstante, é possível evidenciar alguns requisitos básicos que devem ser observados:

- ✓ A QI é multidimensional, por isso são necessárias múltiplas medições para avaliar a qualidade da informação ou dos dados (AL-HAKIM, 2007).
- ✓ As dimensões ou atributos de QI possuem um inter-relacionamento e uma hierarquia que não devem ser desconsiderados para a avaliação da QI (GACKOWSKI, 2004).
- ✓ Como os dados são normalmente estruturados nos sistemas usando uma tripla composta por classes de entidades (*e*), atributos (*a*) e valores (*v*), a avaliação da sua qualidade deve examinar: a adequação do modelo; a adequação da representação e do registro; e a adequação das triplas (entidade, atributo, valor) presentes no modelo de dados (FOX *et al.*, 1994).
- ✓ A avaliação da QI deve analisar não apenas a informação como produto final, mas também os processos que geram esse produto informação (LESCA e LESCOA apud EPPLER, 2006).
- ✓ Os processos de obtenção, processamento, distribuição, armazenamento, recuperação e uso de dados e informações devem ser definidos com precisão e observados por todos os envolvidos.
- ✓ É importante que os processos a serem avaliados estejam suficientemente desenvolvidos ao ponto que as seguintes perguntas da norma NBR ISO 9000 possam ser respondidas afirmativamente; “O processo está identificado e apropriadamente definido?”, “As responsabilidades estão atribuídas?”, “Os procedimentos estão implementados e mantidos?” e “O processo é eficaz em alcançar os resultados requeridos?”.
- ✓ É importante poder identificar a origem e qualidade dos dados em sistemas de bancos de dados heterogêneos com dados oriundos de múltiplas fontes (WANG e MADNICK, 1990).
- ✓ Devem ser conhecidas as premissas utilizadas na atividade de modelagem (aplicada a alguma parte do mundo) que originou os dados e informações presentes em um sistema (FOX *et al.*, 1994).

- ✓ Quando possível, é importante trabalhar com um repositório único de dados mestres de alta qualidade responsável por alimentar as aplicações ao longo da organização com uma visão consistente e sincronizada dos dados da empresa (LOSHIN, 2008).
- ✓ A avaliação da QI deve considerar tanto o ciclo de vida da informação quanto o ciclo de vida da organização, pois eles alteram a importância dos requisitos considerados.
- ✓ Para definição dos atributos que serão monitorados, é importante adotar a visão orientada a propósito de GACKOWSKI (2004) e classificá-los em três grupos; os atributos primários mandatórios, os atributos primários desejáveis e os atributos secundários.

Para atender esses requisitos, um indicador de QI deve evoluir para se adequar às mudanças que ocorrem em uma organização, que são conseqüências de sua evolução e amadurecimento. No início, esse indicador pode até mesmo abranger apenas uma única dimensão, conforme proposto por Olson (2003), sendo as demais incluídas conforme a situação monitorada se estabilize em um patamar aceitável.

Por fim, é preciso entender que tratar as tendências apontadas pelo indicador é mais importante do que perseguir um valor como meta, e que o indicador não é um fim em si mesmo, mas um meio para aprimorar o controle e a confiabilidade dos dados da organização, minimizando assim os riscos de realizar os enormes custos potenciais decorrentes das deficiências na qualidade dos dados e informações.

3 CONSTRUÇÃO EVOLUTIVA DE UM INDICADOR DE QUALIDADE DE INFORMAÇÃO APLICADO NA COLETA DE DADOS DE INTERESSE PARA BANCOS DE DADOS DE CONFIABILIDADE E MANUTENABILIDADE

3.1 Indicadores

O processo de construção de qualquer indicador é, em essência, uma atividade de modelagem, e, como consequência, o seu resultado é uma representação aproximada e sintetizada de algum aspecto da realidade. Assim, um indicador é apenas um meio para o acompanhamento indireto da realidade.

No contexto de uma organização, Caribe (2009) define indicador como:

“... um elemento informacional de síntese, obtido a partir de um modelo de relacionamento de variáveis, cujo produto, seja ele numérico ou simbólico, frente a uma escala ou parâmetros de referência, mostra a situação relativa de um determinado problema/oportunidade ou aspecto da realidade considerado importante ou relevante, em função do que lhe é possível estabelecer e orientar decisões, ações e atividades.”

A construção de um indicador demanda um conhecimento específico da realidade a ser modelada e para que os resultados por ele apresentados possam ser corretamente interpretados é necessário entender tanto os aspectos abrangidos pelo indicador quanto o seu comportamento esperado em relação a alterações na realidade monitorada. Além disso, o mais importante não é o valor apresentado por um indicador e sim a projeção da realidade que é obtida pelo seu comportamento histórico, valor atual e tendência.

Quanto ao modo de utilização, Caribe (2009) entende que o indicador é um:

“... dado informacional que, correlacionado com um referencial de análise, traduz uma informação de valor em uma apreciação relativa, ‘para mais’ ou ‘para menos’, na qual pode interpretar-se e traduzir-se um ou mais aspectos de uma realidade e decidir (ou não) ações ao seu respeito.”

Para que seja estabelecido um referencial de análise é preciso entender a inter-relação entre o indicador e o modelo utilizado para sua construção mas, conforme adverte Caribe (2009), isso nem sempre é devidamente considerado quando são importados indicadores de outras organizações ou mesmo de normas internacionais.

Todo indicador é construído a partir de dados coletados sobre algum aspecto da realidade. É o registro sistemático de dados sobre um evento observado ou controlado que fornece subsídios para entender o seu comportamento esperado e fatores de influência. Entretanto, entre a coleta

e a utilização dos dados, são vários os fatores influem na qualidade das conclusões que podem ser obtidas.

O uso de um indicador de qualidade da informação, nesse contexto, tem como objetivo monitorar alguns desses fatores para que seja possível acompanhar e aprimorar tanto a qualidade dos dados coletados quanto das transformações a que são submetidos para a geração das informações desejadas.

O foco deste trabalho é analisar a qualidade dos dados armazenados e a influência de aspectos que atuam nos processos de observação, coleta e armazenamento. Para isso, é fundamental a definição do processo de registro das ocorrências de manutenção sofridas pelos equipamentos, dos fatores de influência desse processo e dos atores envolvidos.

A construção de um indicador de qualidade das informações coletadas, para bancos de dados de confiabilidade e manutenibilidade, deve levar em conta as particularidades do processo de observação e registro das informações relevantes. Por isso, o ponto de partida deve ser o entendimento da execução e controle da manutenção e do ambiente no qual tudo acontece.

3.2 Entendendo o ambiente de Manutenção

Segundo Monchy (1989), a manutenção dos equipamentos de produção é um elemento chave tanto para a produtividade de uma planta industrial quanto para a qualidade de seus produtos. Sua atuação começa desde a concepção dos equipamentos e vai até o final de sua vida útil e, de certa forma, pode ser considerada como a medicina das máquinas.

Esse nível de acompanhamento demanda o registro detalhado de diversas informações ao longo desse ciclo. Como isso costuma gerar custos consideráveis, o nível de acompanhamento efetivamente executado depende, principalmente, das perdas potenciais associadas às falhas de cada equipamento.

Essa visão que desloca o foco da análise do equipamento para o sistema estudado, e da falha para a perda de uma função, é um dos princípios do processo de manutenção centrada em confiabilidade e ajuda a identificar as funções críticas para o sistema de produção, para as quais é justificável um investimento diferenciado.

Quando o acompanhamento é realizado apenas por função, é dito tratar-se de um acompanhamento do Tag mas, em alguns casos, é necessário acompanhar também o equipamento, como por exemplo, quando isso é previsto por lei ou quando envolve altos custos ou longos prazos para sua substituição.

Definida a questão de quais equipamentos devem ser acompanhados, é importante destacar a dinâmica dos eventos de manutenção. A principal atuação da equipe responsável pela execução da manutenção ocorre em duas situações bem específicas, para realizar intervenções periódicas, que visam manter a confiabilidade e disponibilidade dos equipamentos, e para realizar intervenções corretivas que visam restaurar uma função perdida devido à falha de um equipamento.

Enquanto o controle das intervenções é realizado pela manutenção, cabe a operação reportar as falhas detectadas. É importante destacar que a equipe de manutenção trabalha atendendo os equipamentos, ao passo que a equipe de produção está preocupada com a meta de produção a ser atingida ou superada, e quando a intervenção de manutenção necessita reduzir ou parar a produção, isso costuma gerar atrito entre essas equipes.

Existe muita pressão tanto para conseguir a liberação dos equipamentos quanto para o seu retorno a produção o mais rápido possível. Por isso, muitas vezes informações importantes deixam de ser registradas, ou são registradas de maneira incorreta ou incompleta, e essas lacunas podem comprometer a qualidade de futuros estudos que tomarem como base apenas os dados registrados.

Essa deficiência não costuma ser notada pela equipe de operação, pela equipe de manutenção ou mesmo pelo corpo gerencial, e quando é percebida, nem sempre são encontrados meios eficientes para tratá-la.

Contribuem para esse cenário: o fato do conteúdo dos registros das falhas de equipamentos e atividades de manutenção não ter uma relação direta com a linguagem gerencial; o fato das deficiências nos registros normalmente não produzirem um impacto direto nos resultados do negócio; e a ausência de um mecanismo de acompanhamento adequado. Por isso, a introdução de um indicador de qualidade de informação nesse cenário é capaz de prover o elemento de ligação que está ausente.

Para prover uma resposta adequada, a construção desse indicador demanda alguns pré-requisitos:

- ✓ Conforme descrito na norma NBR ISO 9000, o processo de manutenção precisa estar identificado e apropriadamente definido, com as responsabilidades atribuídas e os procedimentos implementados e mantidos.
- ✓ Deve existir um sistema de controle de demanda e execução das atividades de manutenção.
- ✓ Deve ser possível identificar os usuários que registraram os dados no sistema de manutenção.

Para a definição e modelagem do processo de manutenção, existem algumas metodologias de modelagem de processos de negócio que pode ser empregadas, como o ARIS – *Architecture of integrated information systems* (SCHEER, 2000) e o IDEF3 – *Integration Definition methods 3 Process Description Capture Method* (MAYER *et al.*, 2005).

De forma simplificada, existem dois eventos cujos fluxos precisam ser identificados: quando a operação detecta uma falha e solicita a intervenção da manutenção e quando o sistema de gestão de manutenção gera uma tarefa de manutenção periódica (Fig. 3.1).

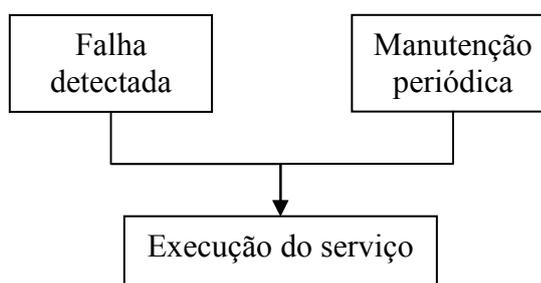


Figura 3.1 – Eventos que demandam a execução de um serviço de manutenção.

Quando se trata de uma falha, a solicitação de serviço pode iniciar com o próprio documento de registro da falha, que deve possuir os campos descritos no Quadro 2.15. Parte desse registro cabe à operação e parte à manutenção. O sistema gera um número para cada falha registrada e é esperado que a operação indique o equipamento que falhou, a data em que a falha foi encontrada, o modo de falha, o método de detecção, a condição de operação quando a falha foi detectada, o impacto da falha na segurança da planta, o impacto da falha nas operações da planta, o impacto da falha na função do equipamento, e uma breve descrição sobre a falha observada e o serviço solicitado. Cabe a manutenção completar o registro identificando o mecanismo de falha, a causa da falha e as subunidades, itens manuteníveis ou partes em pane, além de registrar também os serviços executados conforme o Quadro 2.9.

Quando o serviço é gerado pelo sistema de gestão da manutenção, é esperado da equipe de manutenção apenas o registro dos campos descritos no Quadro 2.9. Por meio do número de registro da tarefa de manutenção deve ser possível identificar o equipamento, o número do registro de falha (quando aplicável), as datas da manutenção (desejada, programada, executada e registrada), a categoria da manutenção, a prioridade, a periodicidade, a atividade, o impacto da manutenção na operação da planta, a subunidade ou item manutenível atendido, peças trocadas ou reparadas, disponibilidade do sobressalente, homem-hora utilizado na tarefa por disciplina e total, ferramentas utilizadas, tempo efetivamente gasto na realização da tarefa,

tempo de indisponibilidade do equipamento em função da tarefa ou da falha, tempo de espera para disponibilização dos materiais utilizados, tempo que a solicitação de manutenção levou para ser concluída, identificação das causas que contribuíram para atrasos na tarefa e em um campo de texto livre, quaisquer informações adicionais relevantes.

Outras informações, mesmo que pertinentes para o controle desses processos, não são tratadas nesse trabalho.

A primeira dificuldade para o controle da qualidade dos registros desses processos é que as informações presentes em um mesmo registro de falha ou manutenção são alimentadas no sistema de gestão por pessoas diferentes, pertencentes a gerências diferentes (às vezes até mesmo a empresas diferentes) e em momentos diferentes. Por isso o sentimento de responsabilidade sobre as informações presentes em cada registro fica diluído.

Então, a primeira etapa deve ser a definição clara de responsabilidades pelas informações que devem ser armazenadas ao longo do processo de registro tanto das falhas quanto das manutenções executadas.

3.3 Delegando propriedade e responsabilidade sobre os dados

O fluxo de execução da manutenção para resolver uma falha reportada é o que envolve a atuação de um maior número de pessoas e gerências. Nesse fluxo acontecem, pelo menos, os seguintes eventos: a detecção da falha; a verificação se a falha já foi reportada e, em caso negativo, o registro da falha; a solicitação de serviço; o planejamento do serviço; a execução do serviço; o retorno sobre o serviço executado; e o retorno sobre os problemas que causaram a falha.

Apesar do claro envolvimento da gerência de operação, ainda é muito presente a idéia que a função manutenção é executada apenas pela área de manutenção. Isso contribui para a falta de interesse na qualidade dos registros de falha abertos pela operação, uma vez que são entendidos apenas como burocracia necessária para solicitação de serviço. Contudo, essa percepção está mudando, principalmente nas empresas onde os operadores estão envolvidos em atividades de RCM ou TPM.

Como a detecção de falha pode ser feita tanto pelo operador quanto pela equipe de manutenção, a primeira questão a ser respondida é; a quem caberá a responsabilidade pelo registro das falhas detectadas? É preciso considerar também que a digitação do registro de falha no sistema ainda pode ser realizada por uma terceira pessoa. E por fim, ainda existe a decisão sobre uma solicitação do serviço ou não, em função da falha reportada.

Cada ator que participa nessa etapa deve ter sua responsabilidade bem definida e no mínimo deve ser solicitada a todos uma verificação básica de coerência sobre as informações recebidas. Fundamentalmente, é necessário relacionar todos os dados sob responsabilidade de cada ator identificado pelo sistema. Isso é válido para todas as etapas do fluxo.

Na etapa seguinte, existe um ator que recebe as solicitações e programa os serviços, ao qual cabe a responsabilidade por verificar quaisquer problemas nas solicitações recebidas e realizar os ajustes necessários ou comunicar para que o solicitante os faça. Para as solicitações corretas, irá providenciar os recursos necessários e programar a execução do serviço.

Agora começa a participação da equipe de execução da manutenção, que recebe e cumpre as tarefas programadas e também deve relatar as informações sobre os serviços realizados. Essas informações, além de completarem o registro sobre o serviço, servirão de base para a conclusão da análise de falha. Dependendo da situação, essa análise pode ser executada por um grupo multidisciplinar e cabe a cada organização identificar como será atribuída a responsabilidade sobre as conclusões obtidas em relação à causa básica do evento de falha.

3.4 Traduzindo a percepção de qualidade em um conjunto de atributos

Com o processo definido e as responsabilidades atribuídas, é possível então analisar a qualidade dos registros armazenados. Essa é uma tarefa intensiva de conhecimento, normalmente executada de maneira *ad hoc*, visando à utilização dos registros em algum estudo específico.

Como a natureza dos problemas de qualidade da informação é bastante diversa, e nem sempre é possível detectá-los de um modo economicamente viável, a melhor abordagem é a prevenção. Mas estabelecer um ambiente propício a uma cultura de qualidade demanda, entre outras coisas, algum modo de acompanhamento que permita verificar ou pelo menos inferir a efetividade das ações empregadas.

A medição de qualidade normalmente se baseia não na presença, e sim na ausência de um ou mais atributos requeridos, e, conforme autores como Olson (2003) e Gackowski (2004), existem alguns desses atributos que tem um impacto maior que os demais na avaliação da qualidade. Além disso, conforme Loshin (2001), as lista de dimensões de qualidade dos dados são extensivas e nunca estão realmente concluídas, pois sempre pode surgir uma nova dimensão relevante.

Isso não invalida os esforços para a gestão da qualidade da informação, apenas alerta que a medição da qualidade da informação pode variar com o amadurecimento da área de QI. Essa

visão de que os requisitos usados para avaliar a qualidade da informação não são fixos é também compartilhada por diversos autores que identificam outros fatores que influenciam a forma de avaliação, como, por exemplo, o ciclo de vida da organização (GINMAN, 1989) e o ciclo de vida do dado (REDMAN, 1992).

Liu e Chi (2002), em afinidade com essa idéia de variação dos requisitos de acordo com o ciclo de vida do dado, propõem que a medição de qualidade da informação seja realizada como um construto evolutivo com incremento dos requisitos de qualidade de acordo com a fase de vida do dado, identificando claramente os seguintes estágios: coleta, organização, apresentação e aplicação. E para a fase de coleta, propõem que os dados sejam avaliados quanto à acurácia, à objetividade, à probidade do coletor, à completeza dos dados coletados, à clareza e aos atributos específicos para a teoria de coleta de dados utilizada.

Ao adaptar e expandir as idéias acima, este trabalho propõe que a avaliação de qualidade dos dados alimentados em um sistema de gestão também seja realizada de forma evolutiva. Sugere-se começar com um conjunto básico de atributos de qualidade que sejam significativos e de fácil obtenção a partir do sistema utilizado para o gerenciamento do processo. Então, quando os fatores representados por esse indicador inicial estiverem bem controlados e apresentando um bom resultado, deve ser realizada uma nova análise dos dados para verificar os problemas de qualidade ainda existentes e quais atributos adicionais, economicamente viáveis, podem ser acrescentados no processo de medição.

Esse modo de implantação ajuda a amenizar o choque provocado pela mudança de cultura, uma vez que foca em pequenos conjuntos de atributos os quais evoluem com a participação de todos os envolvidos no processo, e está em conformidade com a abordagem utilizada em atividades de modelagem, que são conduzidas de modo cíclico uma vez que novas observações podem levar a um refinamento, modificando ou complementando um modelo já construído. Ainda, essas aquisições adicionais de conhecimento podem ser guiadas pelo próprio modelo.

Em relação à construção de um indicador, Caribe (2009) destaca que se deve levar em conta requisitos e características como: especificidade, mensurabilidade, exequibilidade, orientação a resultados, variação no tempo, facilidade, simplicidade, baixo custo de medição, relevância, representatividade, rastreabilidade e disponibilidade da informação.

Para um indicador de QI, é necessário definir: a sigla, o nome, as dimensões ou atributos monitorados (características de qualidade empregadas na construção do indicador), as variáveis (utilizadas para avaliação das características), a equação (relacionamento entre as variáveis que produz um valor) e a frequência de sistematização (intervalo de tempo no qual

os valores são analisados e o indicador é gerado). Também, para cada variável utilizada deve ser definido: tipo de medida (baseado em dados, estimativas ou opinião), fonte (onde a informação é buscada, obtida ou encontrada), unidade de medida (definida para expressar o fenômeno, as categorias, escalas e parâmetros utilizados na coleta e sistematização dos dados), unidade de análise (base sobre a qual são construídas e comparadas as medidas, por exemplo; campos, registros, equipamentos, classes, etc.), instrumentos de coleta e registro (aplicados para obter os valores das variáveis) e frequência de observação ou coleta (intervalo de tempo entre as coletas de valores).

Tendo em mente os comentários acima, é preciso estabelecer um conjunto básico de atributos de qualidade da informação. Conforme visto no tópico 2.5 deste trabalho, ainda não existe um conjunto padronizado de dimensões ou atributos, nem mesmo uma forma padronizada de classificá-las. Mais do que isso, mesmo atributos com os mesmos nomes recebem definições diferentes entre os trabalhos analisados.

Batini e Scannapieco (2006) atribuem isso a relativa imaturidade da área de qualidade de dados, para a qual nenhuma organização internacional elaborou qualquer padrão sobre classificações ou definições de dimensões e métricas de dados. Os padrões atuais que são mais próximos tratam sobre sistemas de gestão da qualidade (NBR ISO 9000) e troca de dados mestres (ISO/TS 8000-110)¹.

Segundo o ISO/TS 8000-110 (*Data quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification*), a qualidade de dados é obtida através do planejamento da qualidade no processo que cria, mantém, entrega e apresenta esses dados a pessoas de modo a prover suporte a processos de decisão. Então, os atributos escolhidos devem ser capazes apresentar indícios sobre a qualidade dos dados gerados e relacioná-las as suas fontes.

A ISO 14224 aborda a necessidade de considerar, monitorar e controlar a qualidade dos dados em seu item sete, onde destaca que dados de alta qualidade devem: ser completos em relação à sua especificação; estar em conformidade com as definições dos parâmetros de confiabilidade, tipos de dados e formatos; ter sua inserção, transferência, manuseio e armazenamento realizados de modo acurado; abranger uma população suficiente e um período de observação adequado para proporcionar confiança estatística; e ser relevantes às necessidades dos usuários. Além disso, destaca que: a origem dos dados deve ser documentada e rastreável, as informações devem ser consistentes, os operadores e mantenedores devem ser consultados para validar os dados e deve ser estabelecido um nível

¹ Liberado em maio de 2008, com status de especificação técnica.

de prioridade para os dados (por exemplo, alta – 100% adequados, média – 75% adequados, e baixa – 50% adequados).

Considerando apenas dados coletados em sistemas informatizados de gestão da manutenção, parte desses atributos são requisitos do sistema, parte são pertinentes aos estágios posteriores a coleta (como organização e processamento), parte dizem respeito ao planejamento do que deve ser coletado (definições dos parâmetros de confiabilidade, tipos de dados e formatos) e para o processo de coleta em si, sobram apenas Completeza, Consistência, Relevância e Acurácia.

Conforme o capítulo 2.5 deste trabalho, esses atributos não são capazes de esgotar a questão da qualidade de dados, mesmo em relação apenas ao processo de coleta, mas são bem próximos do conjunto de atributos identificados por Batini e Scannapieco (2006) como pertinentes a valores de dados (Acurácia, Completeza, Atualidade e Consistência). Vale ressaltar que alguns trabalhos, como Ballou e Pazer (1985) e Fox *et al.* (1994), são baseados apenas nestes quatro atributos.

Combinando esses atributos com os identificados por Liu e Chi (2002) para a fase de coleta, é obtido o seguinte conjunto de atributos: Acurácia, Objetividade, Clareza, Completeza, Probidade do coletor, Atualidade, Consistência e Relevância. Além desses, outros atributos bastante citados na literatura e importantes para avaliar a qualidade da fase de coleta são Reputação, Controle de Duplicação e Credibilidade.

Ao analisar o propósito dos atributos propostos por Liu e Chi (2002), conforme o Quadro 3.1, pode-se perceber que a análise de clareza demanda um conhecimento mais complexo e por isso será excluída do conjunto inicial de atributos a serem monitorados.

Quadro 3.1 – Atributos relativos à qualidade de coleta segundo Liu e Chi (2002)

Fonte: Liu e Chi (2002)

Estágio	Atributo	Definição
Qualidade de coleta	Acurácia	A extensão na qual os dados coletados são livres de erros de medição.
	Completeza	Todos os valores, cuja coleta era prevista pela teoria de coleta, foram coletados.
	Probidade do coletor	A extensão na qual o coletor tem integridade de não cometer falsificação.
	Objetividade	A extensão na qual a amostra selecionada para observação é representativa da população.
	Clareza	A extensão na qual os dados não contêm observações vagas ou ambíguas.

A análise de Objetividade representa o quanto a amostra selecionada para análise é representativa da população. Portanto, somente é aplicável para ponderar as avaliações de atributos que sejam realizadas com base em amostras da população.

Quanto à Acurácia, sua avaliação demanda a realização de auditorias ou entrevistas para que possam ser confrontados os valores registrados e a situação que realmente ocorreu. Isso torna inviável que sua avaliação ocorra com a mesma frequência que a dos demais atributos, mas devido a sua importância, é primordial que seja incluída. Como a avaliação de Acurácia demanda tempo e recursos consideráveis, pode ser executada por amostragem, a qual um julgamento de Objetividade pode ser aplicado para validar a representatividade dos registros escolhidos em relação à população total coletada.

Para avaliar Completeza é preciso verificar a presença de valores, tuplas, atributos e relações. Desses, os valores e relações podem ser verificados a partir dos dados cadastrados, mas a Completeza de tuplas, assim como a avaliação de Acurácia, somente pode ser realizada através de auditorias. A Completeza de atributos não faz parte da avaliação, uma vez que não sofre influência do processo de coleta.

A probidade do coletor, da forma descrita no Quadro 3.1, não é muito significativa para o ambiente de manutenção, mas um atributo semelhante a esse que pode ser aplicado é a Reputação, que segundo Wang e Strong (1996) é descrita como a extensão na qual a informação é altamente considerada em termos de sua fonte. Para este trabalho, a Reputação é construída em função da avaliação da qualidade dos dados gerados por cada fonte, e este valor é utilizado para estimar possíveis perdas de qualidade do banco em função da qualidade das fontes. As fontes são os usuários responsáveis pela entrada de dados, os quais são identificados pelo sistema.

A Consistência, assim como a Completeza, pode ser avaliada em diversas formas: entre os valores de um mesmo registro, entre registros relacionados, e em relação ao comportamento esperado do grupo de equipamentos monitorado. Como esta última análise demanda um conhecimento mais complexo, será realizada inicialmente apenas através de auditorias, como um aspecto da Acurácia (o que está de acordo com o entendimento de Olson, 2003).

O Controle de Duplicação é fundamental para verificar se há mais de um evento registrado para o mesmo equipamento no mesmo período e, caso positivo, se eles são conflitantes. A análise de conflito pode ser estabelecida em uma versão mais simples e evoluir com acréscimos de conhecimento dos especialistas da área sobre o que deve ou não ser considerado como um conflito. A duplicação ou o conflito, a despeito de poder ser avaliado

como um problema de Consistência, deve ser controlado em separado devido sua causa raiz ser um pouco diferente dos demais problemas de Consistência.

A Atualidade e seus atributos afins são incluídos para ponderar o impacto gerado pela diferença entre aquilo que está registrado no sistema e o que está acontecendo em tempo real. Apesar deste trabalho analisar a qualidade dos dados usados para o controle da atividade de manutenção, sua preocupação principal não está direcionada às eventuais dificuldades de gestão, devido à diferença entre o momento que ocorre um evento e o momento em que ele é registrado, e sim para a perda potencial de detalhes no registro ou mesmo ao aumento da probabilidade de distorções entre o que foi registrado e o que realmente ocorreu.

Os últimos dois atributos são: a Relevância, que pondera os problemas registrados em relação à importância das informações afetadas, e a Credibilidade, que é medida por amostragem, em auditorias, e visa avaliar até que ponto a informação é considerada verdadeira e acreditável.

3.5 Mapeando variáveis para os atributos

Alguns dados utilizados para avaliação dos atributos de qualidade devem ser gerados pelo sistema de gestão da manutenção, como a data de entrada e de alteração dos registros e a identificação dos responsáveis pelas alterações. É aconselhável, também, que o sistema de gestão da manutenção possua um mecanismo de exclusão ou marcação para que registros criados erroneamente, quando identificados, possam ser desconsiderados.

Sugere-se que todos os atributos passíveis de avaliação automática, com base nos dados registrados no sistema de gestão de manutenção, sejam acompanhados no mínimo mensalmente, e que, além dos seus valores atuais, sejam acompanhados também os valores mensais dos últimos 12 meses e as médias dos valores acumulados nos últimos 12 e 36 meses. Para as avaliações que dependem de auditoria, não se recomenda a adoção de intervalos de avaliação maiores que um ano.

São especificadas a seguir as variáveis de cada atributo selecionado para a primeira geração do indicador de qualidade.

3.5.1 Relevância

O atributo Relevância tem dupla função, permite tanto estimar a relevância dos dados registrados quanto diferenciar o impacto que os problemas de qualidade encontrados podem

trazer para o negócio, outorgando pesos maiores às informações mais críticas quando é calculada a qualidade da informação referente à fase de coleta.

Para um sistema de gestão de manutenção, as principais falhas e serviços executados devem estar corretamente registrados. A Relevância de cada registro deve ser classificada como pertencente a um dentre os seguintes níveis: crítica, alta, média e baixa.

A Relevância dos registros de falha está diretamente relacionada às suas conseqüências. Segundo a tabela C.1 da ISO 14224, cada falha deve ser avaliada em relação à segurança, ao meio ambiente, à produção e aos custos de manutenção, em quatro graduações: catastrófica, severa, moderada e menor. É importante adicionar uma quinta opção para identificar quando a avaliação de conseqüência não se aplica ou quando a falha não gerou conseqüências. O Quadro 3.2 apresenta uma proposta de fatores para identificar cada graduação.

Quadro 3.2 – Matriz quantitativa de conseqüência para os registros de falha

Segurança	Produção	Meio Ambiente	Custo	Conseqüência
10^4	10^4	10^4	10^4	Catastrófico
10^3	10^3	10^3	10^3	Severo
10^2	10^2	10^2	10^2	Moderado
10^1	10^1	10^1	10^1	Menor
0	0	0	0	Não aplicável ou Sem conseqüências

A ISO 14224 (em sua Tabela 6) requer que seja identificado também, o impacto da falha em relação à função do equipamento, em três graduações: crítica, degradada e incipiente (para as quais são atribuídos os fatores 3, 2 e 1, respectivamente).

O valor total da conseqüência em relação à falha é obtido pela soma dos fatores de cada critério. Esse resultado é útil para identificar quais as falhas mais relevantes dentre as registradas. Entretanto, para o propósito de identificar o grau ou valor de qualidade da informação esperado para o registro, é suficiente utilizar uma escala qualitativa, conforme exibido no Quadro 3.3. A Relevância mínima do registro é determinada pelo maior valor atribuído a qualquer um dos quatro critérios anteriormente mencionados.

Para avaliar a Relevância em relação ao serviço deve ser considerado: o impacto da manutenção na operação da planta (total, médio, baixo, nenhum), a prioridade da manutenção (urgente, alta, média, baixa), o homem-hora total empregado (muito alto, alto, médio, baixo) e o tempo efetivo de manutenção (muito alto, alto, médio, baixo). Para auxiliar a classificação dos valores de homem-hora total empregado e o tempo efetivo de manutenção em faixas, pode ser usado o quartil do conjunto analisado ou, caso existam dados ou conhecimento

suficiente, definidas as faixas aplicáveis. A prioridade da manutenção é utilizada para determinar a programação dos serviços e deve ser estabelecida em função da importância do serviço planejado para a operação da planta, considerando tanto as tarefas previstas quanto o Tag e equipamentos atendidos.

Quadro 3.3 – Matriz de Relevância dos registros de falha

Conseqüência	Segurança	Produção	Meio Ambiente	Custo	Relevância
Catastrófico	Crítica	Crítica	Crítica	Crítica	Crítica
Severo	Alta	Alta	Alta	Alta	Alta
Moderado	Média	Média	Média	Média	Média
Menor	Baixa	Baixa	Baixa	Baixa	Baixa
Não aplicável ou Sem conseqüências	–	–	–	–	

O valor total da Relevância em relação à manutenção é obtido pela soma dos valores relativos a cada critério, os quais estão indicados no Quadro 3.4. Esse resultado é útil para identificar quais as manutenções mais relevantes registradas, e a Relevância mínima do registro é determinada pelo maior valor atribuído em relação a qualquer um dos quatro critérios.

Quadro 3.4 – Matriz de Relevância dos registros de manutenção

Pesos para cada nível	Impacto na Operação	Prioridade da manutenção	Homem-hora Total	Tempo efetivo de Manutenção	Relevância
10^4	Total	Urgente	Muito Alto	Muito Alto	Crítica
10^3	Médio	Alta	Alto	Alto	Alta
10^2	Baixo	Média	Médio	Médio	Média
10^1	Nenhum	Baixa	Baixo	Baixo	Baixa

A classificação da Relevância dos registros de falha e manutenção deve considerar também as Relevâncias do Tag e do equipamento para os quais foram criados, que atuam como modificadores para as Relevâncias dos registros de manutenção e de falha.

A Relevância do Tag deve classificar a importância de sua função para a planta, em quatro níveis: Crítica (a falha provoca agressão ao meio-ambiente ou risco a pessoas), Alta (a falha provoca parada ou grande perda de produção), Média (a falha provoca significativa perda de produção ou afeta a qualidade/especificação do produto) e Baixa (a falha não afeta significativamente a produção, o meio, ambiente e nem gera risco a pessoas, mas pode provocar perda de redundâncias).

A Relevância em relação ao equipamento identifica quais são os que possuem os registros de falha ou de serviço mais relevantes, pois, devido ao compreensível interesse em seu histórico, a Relevância dos demais registros desses equipamentos pode ser maior do que a esperada apenas pelos seus conteúdos. Uma forma simples de determinar a importância do equipamento é apresentada no Quadro 3.5.

Quadro 3.5 – Relevância do Equipamento

Falha mais relevante	Manutenção mais relevante	Equipamento
Crítica	Crítica	Crítica
Alta	Crítica	Crítica
Crítica	Alta	Crítica
Crítica	Média	Alta
Crítica	Baixa	Alta
Média	Crítica	Alta
Baixa	Crítica	Alta
Alta	Alta	Alta
Alta	Média	Alta
Média	Alta	Alta
Alta	Baixa	Média
Média	Média	Média
Baixa	Alta	Média
Média	Baixa	Baixa
Baixa	Média	Baixa
Baixa	Baixa	Baixa

A graduação final de Relevância dos registros de falha está exibida no Quadro 3.6, e foi determinada com base nas seguintes premissas:

1. Falhas com Relevância críticas e altas: todos os campos devem ser corretamente preenchidos, inclusive as análises de mecanismo de falha e causa básica, sem depender das Relevâncias do Tag e do equipamento, por isso suas Relevâncias são consideradas críticas.
2. Para as falhas com Relevância média e baixa podem ter sua Relevância aumentada em função das Relevâncias do Tag e o do equipamento:
 - a. Se ambas forem críticas, a Relevância será crítica.
 - b. Se apenas uma for crítica, a Relevância será alta.
 - c. Se a Relevância do Tag for alta, a Relevância mínima será alta.
 - d. Se a Relevância do Tag for média, a Relevância mínima será média.
 - e. Se a Relevância do Equipamento for alta, a Relevância mínima será média.

Quadro 3.6 – Matriz de Relevância dos registros de falha

Falha	TAG	Equipamento	Relevância
Crítica ou Alta	*	*	Crítica
Média ou Baixa	Crítica	Crítica	Crítica
Média ou Baixa	Alta, Média ou Baixa	Crítica	Alta
Média ou Baixa	Crítica	Alta, Média ou Baixa	Alta
Média ou Baixa	Alta	Alta, Média ou Baixa	Alta
Média ou Baixa	Média ou Baixa	Alta	Média
Média	Média ou Baixa	Média ou Baixa	Média
Baixa	Média	Média ou Baixa	Média
Baixa	Baixa	Média ou Baixa	Baixa

* O valor não afeta a classificação final da Relevância.

A graduação final de Relevância dos registros de manutenção está exibida no Quadro 3.7, e foi determinada com base nas seguintes premissas:

1. Manutenções com Relevância crítica: todos os campos devem ser corretamente preenchidos, sem depender das Relevâncias do Tag e do equipamento.
2. Manutenções com Relevância alta: são reajustadas como crítica, apenas aquelas que atendem um equipamento crítico instalado em um Tag crítico, nas demais situações sua Relevância permanece como alta.
3. Manutenções com Relevância média: são reajustadas como alta, apenas aquelas que atendem um equipamento crítico ou um Tag crítico, nas demais situações sua Relevância permanece como média.
4. Manutenções com Relevância baixa: são reajustadas como média, apenas aquelas que atendem um equipamento crítico ou um Tag crítico, nas demais situações sua Relevância permanece como baixa.

Uma vez determinada a Relevância de cada registro, é preciso estabelecer para cada graduação qual será o percentual mínimo de atendimento das especificações. São sugeridas as seguintes metas: 100% para os registros com Relevância crítica, 90% para os com Relevância alta, 75% para os com Relevância média e 50% para os com Relevância baixa.

Além disso, o cálculo do valor total de qualidade da informação deve ser elaborado considerando o peso de cada nível. São sugeridos os seguintes valores de pesos: Relevância crítica – 6, Relevância alta – 3, Relevância média – 2, Relevância baixa – 1.

Quadro 3.7 – Matriz de Relevância dos registros de manutenção

Manutenção	TAG	Equipamento	Relevância
Crítica	*	*	Crítica
Alta	Crítica	Crítica	Crítica
Alta	Crítica	Alta, Média ou Baixa	Alta
Alta	Alta, Média ou Baixa	Crítica	Alta
Alta	Alta, Média ou Baixa	Alta, Média ou Baixa	Alta
Média	Crítica	*	Alta
Média	*	Crítica	Alta
Média	Alta, Média ou Baixa	Alta, Média ou Baixa	Média
Baixa	Crítica	*	Média
Baixa	*	Crítica	Média
Baixa	Alta, Média ou Baixa	Alta, Média ou Baixa	Baixa

* A graduação não afeta a classificação final da Relevância.

A avaliação da Relevância dos dados coletados, calculada de acordo com o Quadro 3.8, é uma informação importante para estimar a qualidade das conclusões que podem ser obtidas, mas não tem utilidade para avaliar a qualidade da fase de coleta, e por isso não é utilizada na composição do indicador de qualidade da informação para a fase de coleta de dados.

Quadro 3.8 – Ficha do atributo Relevância (continua)

Sigla	Relev
Nome	Relevância
Variáveis	As conseqüências da falha em relação à segurança, ao meio ambiente, à produção e aos custos de manutenção, o impacto da falha em relação à função do equipamento, o impacto da manutenção na operação da planta, a prioridade da manutenção, o homem-hora total empregado, o tempo efetivo de manutenção, a Relevância do Tag, a identificação do Tag, a identificação do equipamento, o número de registro da falha, a data e hora de início da falha, a data e hora de fim da falha, o número de registro da manutenção, a data e hora de início da manutenção e data e hora de fim da manutenção.
Instrumentos de coleta	Ficha de registro de falhas detectadas Ficha de registro de manutenções efetuadas
Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção
Tipo de medida	Baseado nos registros do sistema de gestão da manutenção
Unidade de análise	Registros de falha e manutenção

Quadro 3.8 – Ficha do atributo Relevância (conclusão)

Variáveis calculadas	Rv_Equip*	→ Relevância do equipamento
	Quartil_HH*	→ Posição do registro em relação aos quartis de homem-hora total empregado
	Quartil_TEM*	→ Posição do registro em relação aos quartis de tempo efetivo de manutenção empregado
	qtd_fal_Rv_C*	→ Quantidade de registros de falha com Relevância crítica
	qtd_fal_Rv_A*	→ Quantidade de registros de falha com Relevância alta
	qtd_fal_Rv_M*	→ Quantidade de registros de falha com Relevância média
	qtd_fal_Rv_B*	→ Quantidade de registros de falha com Relevância baixa
	qtd_man_Rv_C*	→ Quantidade de registros de manutenção com Relevância crítica
	qtd_man_Rv_A*	→ Quantidade de registros de manutenção com Relevância alta
	qtd_man_Rv_M*	→ Quantidade de registros de manutenção com Relevância média
qtd_man_Rv_B*	→ Quantidade de registros de manutenção com Relevância baixa	
* determinados de acordo com o item 3.5.1 deste trabalho.		
Fatores utilizados	peso_Rv_C	→ Fator para registros com Relevância crítica
	peso_Rv_A	→ Fator para registros com Relevância alta
	peso_Rv_M	→ Fator para registros com Relevância média
	peso_Rv_B	→ Fator para registros com Relevância baixa
	peso_fal	→ Fator aplicado para ponderação dos registros de falhas
	peso_man	→ Fator aplicado para ponderação dos registros de manutenção
Saídas	<p>Relevância (Relev)</p> $\text{Relev} = (\text{peso_Rv_C} * (\text{qtd_fal_Rv_C} * \text{peso_fal} + \text{qtd_man_Rv_C} * \text{peso_man}) + \text{peso_Rv_A} * (\text{qtd_fal_Rv_A} * \text{peso_fal} + \text{qtd_man_Rv_A} * \text{peso_man}) + \text{peso_Rv_M} * (\text{qtd_fal_Rv_M} * \text{peso_fal} + \text{qtd_man_Rv_M} * \text{peso_man}) + \text{peso_Rv_B} * (\text{qtd_fal_Rv_B} * \text{peso_fal} + \text{qtd_man_Rv_B} * \text{peso_man})) / (\text{peso_fal} * (\text{qtd_fal_Rv_C} + \text{qtd_fal_Rv_A} + \text{qtd_fal_Rv_M} + \text{qtd_fal_Rv_B}) + \text{peso_man} * (\text{qtd_man_Rv_C} + \text{qtd_man_Rv_A} + \text{qtd_man_Rv_M} + \text{qtd_man_Rv_B}))$	
Unidade de medida	Valor percentual (varia de peso_Rv_B a 100%)	
Frequência de medição	Mensal (gerado no dia 4 de cada mês)	
Fonte	Sistema de gestão da manutenção	

3.5.2 Controle de Duplicação

O atributo Controle de Duplicação deve analisar os campos de identificação e datas da manutenção e de identificação, modo e datas da falha para garantir que um mesmo evento não seja registrado mais de uma vez. A ficha completa para a medição do atributo Controle de Duplicação pode ser consultada no Quadro 3.9.

É considerada como duplicação de registro de falha quando um equipamento apresenta interferência entre os períodos de ocorrência de falhas (definidos pelas datas de início e fim de falha) em registros com o mesmo tipo de impacto de falha (crítica, degradada ou incipiente) em relação à sua função.

Sua medição é baseada nos dados registrados no sistema de controle das atividades de manutenção. A análise de duplicação é feita pela comparação dos campos, mas a duplicação é contada por registros. Seu acompanhamento deve considerar tanto o número absoluto quanto o percentual de duplicações encontradas.

É considerada como duplicação de registro de serviço quando um equipamento apresenta interferência entre os períodos de execução de serviços (definidos pelas datas de início e fim real de execução do serviço) em registros com a mesma categoria de manutenção (preventiva ou corretiva).

Como os tempos de ocorrência e de registro dos eventos são diferentes, deve ser estipulado um prazo para que as falhas encontradas e manutenções realizadas sejam registradas, como por exemplo, até três dias após a ocorrência dos eventos. Dessa forma, o valor do Controle de Duplicação pode ser gerado todo o dia 4 de cada mês, com base nos registros dos eventos ocorridos no mês anterior.

Quadro 3.9 – Ficha do atributo Controle de Duplicação (continua)

Sigla	Ctrl_Dup
Nome	Controle de Duplicação
Variáveis	Relação de falhas de Relevância crítica, relação de falhas de Relevância alta, relação de falhas de Relevância média, relação de falhas de Relevância baixa, relação de manutenções de Relevância crítica, relação de manutenções de Relevância alta, relação de manutenções de Relevância média, relação de manutenções de Relevância baixa. Identificação/Tag do Equipamento, número do registro de manutenção, categoria de manutenção, data e hora de início da manutenção, data e hora de fim da manutenção, número do registro de falha, tipo de impacto de falha no equipamento, data e hora de início da falha, data e hora de fim da falha.

Quadro 3.9 – Ficha do atributo Controle de Duplicação (continua)

Variáveis calculadas	qtd_reg_fal_C	→ Quantidade, no período, de registros de falhas vigentes com Relevância crítica
	qtd_reg_fal_A	→ Quantidade, no período, de registros de falhas vigentes com Relevância alta
	qtd_reg_fal_M	→ Quantidade, no período, de registros de falhas vigentes com Relevância média
	qtd_reg_fal_B	→ Quantidade, no período, de registros de falhas vigentes com Relevância baixa
	qtd_dup_fal_C	→ Quantidade, no período, de registros de falhas duplicados com Relevância crítica
	qtd_dup_fal_A	→ Quantidade, no período, de registros de falhas duplicados com Relevância alta
	qtd_dup_fal_M	→ Quantidade, no período, de registros de falhas duplicados com Relevância média
	qtd_dup_fal_B	→ Quantidade, no período, de registros de falhas duplicados com Relevância baixa
	qtd_reg_man_C	→ Quantidade, no período, de registros de manutenção vigentes com Relevância crítica
	qtd_reg_man_A	→ Quantidade, no período, de registros de manutenção vigentes com Relevância alta
	qtd_reg_man_M	→ Quantidade, no período, de registros de manutenção vigentes com Relevância média
	qtd_reg_man_B	→ Quantidade, no período, de registros de manutenção vigentes com Relevância baixa
	qtd_dup_man_C	→ Quantidade, no período, de registros de manutenção duplicados com Relevância crítica
	qtd_dup_man_A	→ Quantidade, no período, de registros de manutenção duplicados com Relevância alta
	qtd_dup_man_M	→ Quantidade, no período, de registros de manutenção duplicados com Relevância média
	qtd_dup_man_B	→ Quantidade, no período, de registros de manutenção duplicados com Relevância baixa
Fatores utilizados	peso_Rv_C	→ Fator para registros com Relevância crítica (100%).
	peso_Rv_A	→ Fator para registros com Relevância alta (90%)
	peso_Rv_M	→ Fator para registros com Relevância média (75%)
	peso_Rv_B	→ Fator para registros com Relevância baixa (50%)
	peso_fal	→ Fator aplicado para ponderação dos registros de falhas.
	peso_man	→ Fator aplicado para ponderação dos registros de manutenção.
Saídas	Controle de Duplicação de Registros de Falhas (Ctrl_Dup_Fal)	
	$\text{Ctrl_Dup_Falha} = 1 - (\text{peso_Rv_C} * \text{qtd_dup_fal_C} / \text{qtd_reg_fal_C} + \text{peso_Rv_A} * \text{qtd_dup_fal_A} / \text{qtd_reg_fal_A} + \text{peso_Rv_M} * \text{qtd_dup_fal_M} / \text{qtd_reg_fal_M} + \text{peso_Rv_B} * \text{qtd_dup_fal_B} / \text{qtd_reg_fal_B}) / (\text{peso_Rv_C} + \text{peso_Rv_A} + \text{peso_Rv_M} + \text{peso_Rv_B})$	

Quadro 3.9 – Ficha do atributo Controle de Duplicação (conclusão)

Saídas	Controle de Duplicação de Registros de Manutenção (Ctrl_Dup_Man) $\text{Ctrl_Dup_Manut} = 1 - \left(\frac{\text{peso_Rv_C} \cdot \text{qtd_dup_man_C}}{\text{qtd_reg_man_C}} + \frac{\text{peso_Rv_A} \cdot \text{qtd_dup_man_A}}{\text{qtd_reg_man_A}} + \frac{\text{peso_Rv_M} \cdot \text{qtd_dup_man_M}}{\text{qtd_reg_man_M}} + \frac{\text{peso_Rv_B} \cdot \text{qtd_dup_man_B}}{\text{qtd_reg_man_B}} \right) / (\text{peso_Rv_C} + \text{peso_Rv_A} + \text{peso_Rv_M} + \text{peso_Rv_B})$
	Controle de Duplicação de Registros (Ctrl_Dup) $\text{Ctrl_Dup} = \frac{(\text{Ctrl_Dup_Fal} \cdot \text{peso_dup_fal} + \text{Ctrl_Dup_Man} \cdot \text{peso_dup_man})}{(\text{peso_dup_fal} + \text{peso_dup_man})}$
Instrumentos de coleta	Ficha de registro de falhas detectadas Ficha de registro de manutenções efetuadas
Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção
Tipo de medida	Dados do sistema de gestão da manutenção
Unidade de análise	Registros de falha e manutenção
Unidade de medida	Valor percentual (varia de 0% a 100%)
Frequência de medição	Mensal (gerado no dia 4 de cada mês)
Fonte	Sistema de gestão da manutenção

3.5.3 Completeza

O atributo Completeza precisa analisar todos os campos dos registros de falha e de serviços de manutenção, definidos nos Quadros 2.15 e 2.9, respectivamente. A ausência de qualquer campo definido como obrigatório pelo ISO 14224 compromete integralmente o registro, enquanto que a ausência de informações adicionais degrada sua Completeza. Seu acompanhamento deve considerar, além do percentual de Completeza, o valor de todas as variáveis calculadas.

Toda a falha detectada em um equipamento deve receber algum tratamento, ou a programação de um serviço ou o acompanhamento da situação encontrada, que deverá estar identificado no sistema de gestão de manutenção. Assim como todo o serviço realizado em resposta a uma falha deve ter, no sistema de gestão de manutenção, o registro de falha correspondente.

Além disso, é importante avaliar a Completeza em relação ao registro de todas as ocorrências de manutenção através de auditorias periódicas. Sugere-se que estas auditorias sejam

realizadas semestralmente com base, por exemplo, em eventos significativos registrados em outros sistemas da empresa ou pela análise de Consistência entre os registros de participação dos funcionários em serviços de manutenção e os registros de frequência no trabalho.

Como frutos dessa auditoria, devem ser gerados os valores estimados da Completeza em relação aos eventos de falha (est_comp_fal) e de manutenção (est_comp_man) presentes no sistema de gestão da manutenção e deve ser providenciada a inserção de todos os eventos não-registrados identificados. Ressalta-se que esses registros, inseridos como frutos de um trabalho de auditoria, penalizam a qualidade da coleta tanto em relação à Completeza quanto à Atualidade. A ficha completa para a medição do atributo Completeza pode ser consultada no Quadro 3.10.

Quadro 3.10 – Ficha do atributo Completeza (continua)

Sigla	Comp
Nome	Completeza
Variáveis	<p>Relação de falhas de Relevância crítica, relação de falhas de Relevância alta, relação de falhas de Relevância média, relação de falhas de Relevância baixa, relação de manutenções de Relevância crítica, relação de manutenções de Relevância alta, relação de manutenções de Relevância média, relação de manutenções de Relevância baixa.</p> <p>Dados obtidos no sistema de gestão da manutenção:</p> <p>Dados obrigatórios de manutenção: Número do registro de Manutenção, Identificação/Tag do Equipamento, Número do registro de falha, Datas da manutenção, Categoria de manutenção, <i>Down time</i>, Tempo de manutenção ativa.</p> <p>Dados adicionais de manutenção: Prioridade da manutenção, Periodicidade (manutenção oriunda de plano), Atividade de Manutenção, Impacto da manutenção na operação da planta, Subunidade atendida pela manutenção, Item(s) manutenível(is) atendido(s) pela manutenção, Sobressalente utilizado, Localização do sobressalente, Homem-hora de manutenção por disciplina, Homem-hora total de manutenção, Equipamentos de manutenção utilizados, <i>Lead Time</i>, <i>Turn Around Time</i> (TAT), Atrasos da manutenção, Informações Adicionais necessárias para a gestão da manutenção.</p> <p>Dados obrigatórios de falha: Número do registro de falha, Identificação/Tag do Equipamento, Data da Falha, Modo de Falha, Método de Detecção, Impacto da falha na função do equipamento.</p> <p>Dados adicionais de falha: Condição de operação quando a falha foi detectada, Impacto da falha na segurança da planta, Impacto da falha nas operações da planta, Mecanismo de Falha, Causa da Falha, Subunidade em pane, Item(s) Manutenível(is) em pane, Informações Adicionais (descrição do evento em texto livre), Outras informações adicionais necessárias.</p>

Quadro 3.10 – Ficha do atributo Completeza (continua)

	Dados obtidos em relatórios de auditoria:
Variáveis	est_inc_fal_C → Estimativa do percentual de falhas com Relevância crítica não registradas por mês
	est_inc_fal_A → Estimativa do percentual de falhas com Relevância alta não registradas por mês
	est_inc_fal_M → Estimativa do percentual de falhas com Relevância média não registradas por mês
	est_inc_fal_B → Estimativa do percentual de falhas com Relevância baixa não registradas por mês
	est_inc_man_C → Estimativa do percentual de manutenções com Relevância crítica não registradas por mês
	est_inc_man_A → Estimativa do percentual de manutenções com Relevância alta não registradas por mês
	est_inc_man_M → Estimativa do percentual de manutenções com Relevância média não registradas por mês
	est_inc_man_B → Estimativa do percentual de manutenções com Relevância baixa não registradas por mês
Instrumentos de coleta	Ficha de registro de falhas detectadas; ficha de registro de manutenções efetuadas e relatórios de auditoria.
Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção Auditorias realizadas semestralmente
Tipo de medida	Dados do sistema de gestão da manutenção Estimativa do percentual de eventos não registrados
Variáveis calculadas	qtd_conc_fal_C → Quantidade, no período, de registros de falha concluídos com Relevância crítica
	qtd_conc_fal_A → Quantidade, no período, de registros de falha concluídos com Relevância alta
	qtd_conc_fal_M → Quantidade, no período, de registros de falha concluídos com Relevância média
	qtd_conc_fal_B → Quantidade, no período, de registros de falha concluídos com Relevância baixa
	qtd_incp_o_fal_C → Quantidade, no período, de registros de falhas com Relevância crítica concluídos sem alguma informação obrigatória
	qtd_incp_o_fal_A → Quantidade, no período, de registros de falhas com Relevância alta concluídos sem alguma informação obrigatória
	qtd_incp_o_fal_M → Quantidade, no período, de registros de falhas com Relevância média concluídos sem alguma informação obrigatória
	qtd_incp_o_fal_B → Quantidade, no período, de registros de falhas com Relevância baixa concluídos sem alguma informação obrigatória
	qtd_incp_a_fal_C → Quantidade, no período, de registros de falhas com Relevância crítica concluídos sem alguma informação adicional
	qtd_incp_a_fal_A → Quantidade, no período, de registros de falhas com Relevância alta concluídos sem alguma informação adicional
	qtd_incp_a_fal_M → Quantidade, no período, de registros de falhas com Relevância média concluídos sem alguma informação adicional
	qtd_incp_a_fal_B → Quantidade, no período, de registros de falhas com Relevância baixa concluídos sem alguma informação adicional

Quadro 3.10 – Ficha do atributo Completeza (continua)

Unidade de análise	Avaliação automática: Campos dos registros de falha e manutenção Auditoria: Eventos não registrados
Variáveis calculadas	<p>qtd_conc_man_C → Quantidade, no período, de registros de manutenção concluídos com Relevância crítica</p> <p>qtd_conc_man_A → Quantidade, no período, de registros de manutenção concluídos com Relevância alta</p> <p>qtd_conc_man_M → Quantidade, no período, de registros de manutenção concluídos com Relevância média</p> <p>qtd_conc_man_B → Quantidade, no período, de registros de manutenção concluídos com Relevância baixa</p> <p>qtd_incp_o_man_C → Quantidade, no período, de registros de manutenção concluídos com Relevância crítica sem alguma informação obrigatória</p> <p>qtd_incp_o_man_A → Quantidade, no período, de registros de manutenção concluídos com Relevância alta sem alguma informação obrigatória</p> <p>qtd_incp_o_man_M → Quantidade, no período, de registros de manutenção concluídos com Relevância média sem alguma informação obrigatória</p> <p>qtd_incp_o_man_B → Quantidade, no período, de registros de manutenção concluídos com Relevância baixa sem alguma informação obrigatória</p> <p>qtd_incp_a_man_C → Quantidade, no período, de registros de manutenção concluídos com Relevância crítica sem alguma informação adicional</p> <p>qtd_incp_a_man_A → Quantidade, no período, de registros de manutenção concluídos com Relevância alta sem alguma informação adicional</p> <p>qtd_incp_a_man_M → Quantidade, no período, de registros de manutenção concluídos com Relevância média sem alguma informação adicional</p> <p>qtd_incp_a_man_B → Quantidade, no período, de registros de manutenção concluídos com Relevância baixa sem alguma informação adicional</p>
Fatores utilizados	<p>peso_Rv_C → Fator para registros com Relevância crítica (100%).</p> <p>peso_Rv_A → Fator para registros com Relevância alta (90%)</p> <p>peso_Rv_M → Fator para registros com Relevância média (75%)</p> <p>peso_Rv_B → Fator para registros com Relevância baixa (50%)</p> <p>peso_fal → Fator aplicado para ponderação dos registros de falhas.</p> <p>peso_man → Fator aplicado para ponderação dos registros de manutenção.</p> <p>peso_info_a → fator de degradação por ausência de uma ou mais informações adicionais</p>
Saídas	<p>Completeza de Registros de Falhas (Comp_Fal)</p> $\text{Comp_Fal} = \left(\left(\text{peso_Rv_C} * (1 - (\text{qtd_incp_o_fal_C} + \text{qtd_incp_a_fal_C} * \text{peso_info_a}) / \text{qtd_conc_fal_C}) \right) * (1 - \text{est_inc_fal_C}) + \left(\text{peso_Rv_A} * (1 - (\text{qtd_incp_o_fal_A} + \text{qtd_incp_a_fal_A} * \text{peso_info_a}) / \text{qtd_conc_fal_A}) \right) * (1 - \text{est_inc_fal_A}) + \left(\text{peso_Rv_M} * (1 - (\text{qtd_incp_o_fal_M} + \text{qtd_incp_a_fal_M} * \text{peso_info_a}) / \text{qtd_conc_fal_M}) \right) * (1 - \text{est_inc_fal_M}) + \left(\text{peso_Rv_B} * (1 - (\text{qtd_incp_o_fal_B} + \text{qtd_incp_a_fal_B} * \text{peso_info_a}) / \text{qtd_conc_fal_B}) \right) * (1 - \text{est_inc_fal_B}) \right) / (\text{peso_Rv_C} + \text{peso_Rv_A} + \text{peso_Rv_M} + \text{peso_Rv_B})$

Quadro 3.10 – Ficha do atributo Completeza (conclusão)

Saídas	<p>Completeza de Registros de Manutenção (Comp_Man)</p> $\text{Comp_Man} = \left(\left(\text{peso_Rv_C} * (1 - (\text{qtd_incp_o_man_C} + \text{qtd_incp_a_man_C} * \text{peso_info_a}) / \text{qtd_conc_man_C}) \right) * (1 - \text{est_inc_man_C}) + \left(\text{peso_Rv_A} * (1 - (\text{qtd_incp_o_man_A} + \text{qtd_incp_a_man_A} * \text{peso_info_a}) / \text{qtd_conc_man_A}) \right) * (1 - \text{est_inc_man_A}) + \left(\text{peso_Rv_M} * (1 - (\text{qtd_incp_o_man_M} + \text{qtd_incp_a_man_M} * \text{peso_info_a}) / \text{qtd_conc_man_M}) \right) * (1 - \text{est_inc_man_M}) + \left(\text{peso_Rv_B} * (1 - (\text{qtd_incp_o_man_B} + \text{qtd_incp_a_man_B} * \text{peso_info_a}) / \text{qtd_conc_man_B}) \right) * (1 - \text{est_inc_man_B}) \right) / (\text{peso_Rv_C} + \text{peso_Rv_A} + \text{peso_Rv_M} + \text{peso_Rv_B})$ <p>Completeza (Comp)</p> $\text{Comp} = (\text{Comp_Fal} * \text{peso_fal} + \text{Comp_Man} * \text{peso_man}) / (\text{peso_fal} + \text{peso_man})$
Unidade de medida	Valor percentual (varia de 0% a 100%)
Frequência de medição	Mensal (gerado no dia 4 de cada mês) Fatores atualizados semestralmente.
Fonte	Sistema de gestão da manutenção Relatório de auditoria

3.5.4 Atualidade

A Atualidade influencia a qualidade do processo de coleta de duas formas diferentes: conforme aumenta o tempo entre uma ocorrência e seu registro, aumenta também a quantidade de detalhes perdida; e quanto maior a defasagem entre o que ocorre na planta mantida e o que está registrado no sistema de gestão da manutenção, menos confiáveis são as informações do sistema de gestão da manutenção para tomada de decisões. Registros que gerarem valores negativos em relação à Atualidade são desconsiderados para a avaliação desse atributo, mas penalizam o processo de coleta em relação à Consistência, por terem sido registrados antes que pudessem ter ocorrido.

Para a avaliação da Atualidade de um registro de falha, são computadas as diferenças entre a data e hora de início da falha e a data e hora do seu cadastro e entre a data e hora de fim da falha e a data e hora do seu cadastro. O maior desses intervalos é considerado como o tempo de desatualização do registro. Para retratar a desatualização do conjunto de registros estudado, deve ser apresentado o valor médio e o desvio padrão. Quanto menor for o desvio padrão, menor será a oscilação entre os tempos de registro das falhas.

Para a avaliação da Atualidade de um registro de manutenção são computadas as diferenças entre a data e hora de fim dos serviços e a data e hora de seus cadastros. Uma vez que um registro pode ter vários serviços, ou um serviço pode ser executado em mais de um período, o maior intervalo entre o fim de um período de serviço e sua inserção no sistema é considerado como o tempo de desatualização do registro. Para retratar a desatualização do conjunto de registros estudado deve ser apresentado o valor médio e o desvio padrão. Quanto menor for o desvio padrão, menor será a oscilação entre os tempos de registro dos serviços.

Tanto a Atualidade de um registro de falha quanto a de um registro de manutenção são avaliadas em termos do percentual de qualidade da informação que comprometem. Para isso é preciso definir quais são os tempos aceitáveis para seus desvios padrão e médias em relação aos requisitos do processo gerido.

Por exemplo, vamos considerar que até 12h de defasagem média ou de desvio padrão não comprometem a qualidade, mas que valores maiores degradam a qualidade em 2% a cada 12 horas para a média e para o desvio padrão e quando a média atinge o limite de 72h sua taxa de degradação passa a ser 4% a cada 12 horas.

Assim tanto a defasagem, medida pela média, quanto a variação dos tempos de registro, medida pelo desvio padrão, impactam na avaliação desse atributo. A ficha completa para a medição do atributo Atualidade pode ser consultada no Quadro 3.11.

Quadro 3.11 – Ficha do atributo Atualidade (continua)

Sigla	Atual
Nome	Atualidade
Variáveis	Relação de falhas de Relevância crítica, relação de falhas de Relevância alta, relação de falhas de Relevância média, relação de falhas de Relevância baixa, relação de manutenções de Relevância crítica, relação de manutenções de Relevância alta, relação de manutenções de Relevância média, relação de manutenções de Relevância baixa. Data e hora de início da falha (dthr_ini_fal), data e hora do cadastro do início da falha (dthr_ini_sistema), data e hora de fim da falha (dthr_fim_fal) e a data e hora do cadastro do fim da falha (dthr_fim_sistema), data e hora de fim de um período de serviço (dthr_fim_serv) e a data e hora de cadastro do fim de um período de serviço (dthr_reg_serv).
Instrumentos de coleta	Ficha de registro de falhas detectadas Ficha de registro de manutenções efetuadas
Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção
Unidade de análise	Registros de falha e manutenção.

Quadro 3.11 – Ficha do atributo Atualidade (continua)

	<p>Atualidade de Registros de Falhas (Atual_Fal_A)</p> $\mu_{\text{Atual_Fal_A}} = (1/\text{qtd_reg_fal_A}) * \sum \text{Max}(\text{dthr_ini_sistema} - \text{dthr_ini_fal}; \text{dthr_fim_sistema} - \text{dthr_fim_fal})$ $\sigma_{\text{Atual_Fal_A}} = ((1/\text{qtd_reg_fal_A}) * \sum (\text{Max}(\text{dthr_ini_sistema} - \text{dthr_ini_fal}; \text{dthr_fim_sistema} - \text{dthr_fim_fal}) - \mu_{\text{Atual_Fal_A}})^2)^{1/2}$ <p>Atual_Fal_A:= Se $\mu_{\text{Atual_Fal_A}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Fal_A}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Fal_A}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se Atual_Fal_A = 1 - (Peso_med*($\mu_{\text{Atual_Fal_A}} - 12$)/ 12 + Peso_desv*($\sigma_{\text{Atual_Fal_A}} - 12$)/12)</p>
Saídas	<p>Atualidade de Registros de Falhas (Atual_Fal_M)</p> $\mu_{\text{Atual_Fal_M}} = (1/\text{qtd_reg_fal_M}) * \sum \text{Max}(\text{dthr_ini_sistema} - \text{dthr_ini_fal}; \text{dthr_fim_sistema} - \text{dthr_fim_fal})$ $\sigma_{\text{Atual_Fal_M}} = ((1/\text{qtd_reg_fal_M}) * \sum (\text{Max}(\text{dthr_ini_sistema} - \text{dthr_ini_fal}; \text{dthr_fim_sistema} - \text{dthr_fim_fal}) - \mu_{\text{Atual_Fal_M}})^2)^{1/2}$ <p>Atual_Fal_M:= Se $\mu_{\text{Atual_Fal_M}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Fal_M}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Fal_M}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se Atual_Fal_M = 1 - (Peso_med*($\mu_{\text{Atual_Fal_M}} - 12$)/ 12 + Peso_desv*($\sigma_{\text{Atual_Fal_M}} - 12$)/12)</p>

Quadro 3.11 – Ficha do atributo Atualidade (continua)

	<p>Atualidade de Registros de Falhas (Atual_Fal_B)</p> $\mu_{\text{Atual_Fal_B}} = (1/\text{qtd_reg_fal_B}) * \sum \text{Max}(\text{dthr_ini_sistema} - \text{dthr_ini_fal}; \text{dthr_fim_sistema} - \text{dthr_fim_fal})$ $\sigma_{\text{Atual_Fal_B}} = ((1/\text{qtd_reg_fal_B}) * \sum (\text{Max}(\text{dthr_ini_sistema} - \text{dthr_ini_fal}; \text{dthr_fim_sistema} - \text{dthr_fim_fal}) - \mu_{\text{Atual_Fal_B}})^2)^{1/2}$ <p>Atual_Fal_B:= Se $\mu_{\text{Atual_Fal_B}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Fal_B}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Fal_B}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se Fim Se $\text{Atual_Fal_B} = 1 - (\text{Peso_med} * (\mu_{\text{Atual_Fal_B}} - 12) / 12 + \text{Peso_desv} * (\sigma_{\text{Atual_Fal_B}} - 12) / 12)$</p>
Saídas	<p>Atualidade de Registros de Manutenção (Atual_Man_C)</p> $\mu_{\text{Atual_Man_C}} = (1/\text{qtd_reg_man_C}) * \sum \text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv})$ $\sigma_{\text{Atual_Man_C}} = ((1/\text{qtd_reg_man_C}) * \sum (\text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv}) - \mu_{\text{Atual_Man_C}})^2)^{1/2}$ <p>Atual_Man_C:= Se $\mu_{\text{Atual_Man_C}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Man_C}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Man_C}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se Fim Se $\text{Atual_Man_C} = 1 - ((\mu_{\text{Atual_Man_C}} - 12) * \text{Peso_med} / 12 + (\sigma_{\text{Atual_Man_C}} - 12) * \text{Peso_desv} / 12)$</p>

Quadro 3.11 – Ficha do atributo Atualidade (continua)

	<p>Atualidade de Registros de Manutenção (Atual_Man_A)</p> $\mu_{\text{Atual_Man_A}} = (1/\text{qtd_reg_man_A}) * \sum \text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv})$ $\sigma_{\text{Atual_Man_A}} = ((1/\text{qtd_reg_man_A}) * \sum (\text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv}) - \mu_{\text{Atual_Man_A}})^2)^{1/2}$ <p>Atual_Man_A := Se $\mu_{\text{Atual_Man_A}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Man_A}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Man_A}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se Atual_Man_A = 1 - (($\mu_{\text{Atual_Man_A}} - 12$)*Peso_med/12 + ($\sigma_{\text{Atual_Man_A}} - 12$)*Peso_desv/12)</p>
Saídas	<p>Atualidade de Registros de Manutenção (Atual_Man_M)</p> $\mu_{\text{Atual_Man_M}} = (1/\text{qtd_reg_man_M}) * \sum \text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv})$ $\sigma_{\text{Atual_Man_M}} = ((1/\text{qtd_reg_man_M}) * \sum (\text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv}) - \mu_{\text{Atual_Man_M}})^2)^{1/2}$ <p>Atual_Man_M := Se $\mu_{\text{Atual_Man_M}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Man_M}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Man_M}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se Atual_Man_M = 1 - (($\mu_{\text{Atual_Man_M}} - 12$)*Peso_med/12 + ($\sigma_{\text{Atual_Man_M}} - 12$)*Peso_desv/12)</p>

Quadro 3.11 – Ficha do atributo Atualidade (continua)

Saídas	<p>Atualidade de Registros de Manutenção (Atual_Man_B)</p> $\mu_{\text{Atual_Man_B}} = (1/\text{qtd_reg_man_B}) * \sum \text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv})$ $\sigma_{\text{Atual_Man_B}} = ((1/\text{qtd_reg_man_B}) * \sum (\text{Max}(\text{dthr_fim_serv} - \text{dthr_reg_serv}) - \mu_{\text{Atual_Man_B}})^2)^{1/2}$ <p>Atual_Man_B := Se $\mu_{\text{Atual_Man_B}} \geq 72$ Então Peso_med = Peso_med_+72 Senão Se $\mu_{\text{Atual_Man_B}} \geq 12$ Então Peso_med = Peso_med_72 Senão Peso_med = 0 Fim Se Fim Se Se $\sigma_{\text{Atual_Man_B}} \geq 12$ Então Peso_desv = Peso_desvpad Senão Peso_desv = 0 Fim Se $\text{Atual_Man_B} = 1 - ((\mu_{\text{Atual_Man_B}} - 12) * \text{Peso_med} / 12 + (\sigma_{\text{Atual_Man_B}} - 12) * \text{Peso_desv} / 12)$</p>
	<p>Atualidade (Atual)</p> $\text{Atual} = (\text{peso_fal} * (\text{peso_Rv_C} * \text{Atual_Fal_C} + \text{peso_Rv_A} * \text{Atual_Fal_A} + \text{peso_Rv_M} * \text{Atual_Fal_M} + \text{peso_Rv_B} * \text{Atual_Fal_B}) + \text{peso_man} * (\text{peso_Rv_C} * \text{Atual_Man_C} + \text{peso_Rv_A} * \text{Atual_Man_A} + \text{peso_Rv_M} * \text{Atual_Man_M} + \text{peso_Rv_B} * \text{Atual_Man_B})) / (\text{peso_Rv_C} + \text{peso_Rv_A} + \text{peso_Rv_M} + \text{peso_Rv_B}) * (\text{peso_fal} + \text{peso_man})$
Fatores utilizados	<p>peso_Rv_C → Fator para registros com Relevância crítica (100%). peso_Rv_A → Fator para registros com Relevância alta (90%) peso_Rv_M → Fator para registros com Relevância média (75%) peso_Rv_B → Fator para registros com Relevância baixa (50%) peso_fal → Fator aplicado para ponderação dos registros de falhas. peso_man → Fator aplicado para ponderação dos registros de manutenção. peso_med_72 → fator de degradação da Atualidade para defasagem média de 12 até 72 horas peso_med_+72 → fator de degradação da Atualidade para defasagem média maior ou igual a 72 horas peso_desvpad → fator de degradação da Atualidade para desvio padrão maior ou igual a 12 horas</p>
Tipo de medida	Baseado nos registros do sistema de gestão da manutenção

Quadro 3.11 – Ficha do atributo Atualidade (conclusão)

Unidade de medida	Valor percentual (varia de 0% a 100%)
Frequência de medição	Mensal (gerado no dia 4 de cada mês)
Fonte	Sistema de gestão da manutenção

3.5.5 Acurácia

A verificação de Acurácia deve ser realizada por amostragem, analisando os registros de falhas e manutenções do sistema de gestão concluídos dentro do período escolhido para auditoria, para estimar a veracidade e exatidão de suas informações. É necessário utilizar a amostragem porque a avaliação de Acurácia de cada registro armazenado no sistema seria muito dispendiosa, uma vez que demanda comparar informações armazenadas no sistema com outras fontes, como as memórias dos funcionários ou matérias de jornal.

O primeiro passo para uma boa amostragem consiste em saber como escolher amostras em diversidade e quantidade suficientes para obter um resultado que seja representativo da população total, ou seja, que possibilite a inferência de generalizações para a população total (N), com base nas características da distribuição normal (Rea e Parker, 2002). É a hipótese de normalidade que permite que, com base em uma amostra, sejam feitos julgamentos probabilísticos a respeito da população.

Para a seleção de uma amostra é necessário identificar a população a ser amostrada e definir dentre ela, qual será a população útil. Para o escopo deste trabalho, a população útil abrange toda a população dos registros de falha e manutenção armazenados no sistema de gestão. Sobre ela deverá ser aplicada uma amostragem randômica sem substituição e estratificada de acordo com a avaliação de Relevância (crítica, alta, média e baixa) dos registros, produzindo oito sub-populações compostas por quatro grupos de registros de falha e quatro de registros de manutenção.

Ou seja, as sub-populações escolhidas com base na amostragem aleatória estratificada simples são separadas em dois grupos (os registros de falha e de manutenção) que por sua vez são divididos em quatro subgrupos, em função da classificação de Relevância dos registros. Essa divisão é para refletir as diferenças esperadas na avaliação de Acurácia dos registros de cada subgrupo devido às diferenças entre suas metas e pesos.

Isso é suficiente para a primeira versão de avaliação do atributo Acurácia. Para versões mais evoluídas, podem ser incluídos outros atributos para a estratificação das sub-populações, gerando assim uma estratificação cruzada ao invés da estratificação simples.

Para cada sub-população deve ser escolhido um tamanho de amostra proporcional à quantidade total de registros estudados. Essa estratégia, chamada de amostragem estratificada proporcional, consiste em dividir proporcionalmente a população em subgrupos homogêneos (estratos) dos quais são retiradas amostra aleatórias simples.

Definida a estratégia de amostragem a ser utilizada, é necessário determinar o tamanho mínimo necessário para as amostras para que o valor médio de Acurácia obtido seja representativo da população.

Com base no Teorema do Limite Central, sabe-se que a média de uma amostra (\bar{x}) se aproxima da média (μ) da população real e, por isso, é uma boa estimativa dela. Com o desvio padrão (s) de uma amostra contendo n elementos é possível estimar o erro padrão da média ($\sigma_{\bar{x}}$) com:

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} \quad (1)$$

Sendo s calculado por:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad (2)$$

O erro padrão da média mede o desvio padrão da distribuição das médias de amostras com n elementos (Cooper e Schindler, 2003), e a expressão $\sqrt{(N-n)/(N-1)}$ é conhecida como correção de população finita e se aproxima de 1 para populações grandes, situação na qual é normalmente desconsiderada (Rea e Parker, 2002). Na prática esse fator só é usado quando a amostra constitui pelo menos 5% da população (Freund, 2006).

Segundo o Teorema do limite central, para amostras suficientemente grandes ($n \geq 30$) as médias de amostras são distribuídas pela média da população aproximadamente em uma distribuição normal, mesmo se a população não for normalmente distribuída (Cooper e Schindler, 2003).

Sabe-se que a média da população (μ) raramente será a mesma que a média da amostra (\bar{x}) e o desvio padrão da população (σ) dificilmente será o mesmo da amostra (s) devido aos erros de amostragem que ocorrem de forma independente do rigor científico com que a amostra foi

selecionada e implementada (Rea e Parker, 2002). Esses erros podem ser estimados através do uso de intervalos de confiança.

É conhecido que em uma distribuição normal, existe uma probabilidade de 0,6826 de que a média de qualquer amostra esteja dentro de um erro padrão da média verdadeira, ou seja, pode-se dizer que o nível de confiança é de 68,26%. Normalmente é requerido um nível de confiança maior, sendo muito utilizado um nível de confiança de 95%, para o qual o intervalo de confiança é de $\pm 1,96$ vezes o valor do erro padrão. Essas pontuações em unidades fracionárias de desvio padrão são conhecidas como Estatística Z (Rea e Parker, 2002).

Segundo Rea e Parker (2002) existem dois fatores inter-relacionados que o pesquisador precisa especificar antes de prosseguir com a especificação do tamanho de uma amostra¹: o nível de confiança (risco de erro que o pesquisador está disposto a aceitar no estudo) e o intervalo de confiança (que determina o nível de precisão da amostragem que o pesquisador obtém). O intervalo de confiança (C) é determinado pela multiplicação do nível de confiança (Z) pelo desvio padrão da amostra (s).

A quantidade necessária de registros para a amostra aumenta quanto maior for sua dispersão ou variância, sua precisão estimativa desejada, o nível de confiança na estimativa e o número de subgrupos de interesse dentro da amostra – uma vez que cada subgrupo também deve atender as exigências mínimas de amostra. A precisão é medida pela amplitude do intervalo no qual é esperado encontrar a estimativa de parâmetro e pelo grau de confiança que querem ter na estimativa.

O tamanho (n) de uma amostra pode ser calculado por:

$$n = \left(\frac{Z \cdot \sigma}{\sigma_{\bar{x}}} \right)^2 \quad (3)$$

Onde:

- Z limite para o nível de confiança desejado
- n tamanho da amostra
- σ desvio padrão da população
- $\sigma_{\bar{x}}$ erro padrão tolerável ou admitido (também referenciado como E)

Quando o desvio padrão da população (σ) é desconhecido, pode ser substituído pelo desvio padrão de uma amostra de pré-teste (s), conforme exibido na equação 1. Sugere-se que na primeira medição da Acurácia, seja realizada uma amostragem piloto (com, por exemplo, 30

¹ – que está diretamente relacionado à precisão da média da amostra como estimativa da média da população real.

amostras) para que o seu desvio padrão possa ser utilizado para determinar o tamanho necessário para que uma amostra atinja a precisão requerida.

Uma vez determinada a amostra a ser utilizada, deve ser realizada a análise de Acurácia dos seus elementos. Para isso, devem ser considerados também alguns aspectos descritos em outros atributos como, Objetividade, Consistência, Completeza e Credibilidade.

Como a Objetividade visa avaliar o quanto a amostra selecionada é representativa da população, quando a amostra é selecionada segundo princípios estatísticos essa análise pode ser desconsiderada, uma vez que nesse caso as amostras possuem idealmente 100% de Objetividade.

A Consistência e a Completeza são avaliadas em conjunto para determinar se os registros selecionados possuem todos os relacionamentos previstos com os demais registros do sistema de gestão, através dos valores de seus campos. Assim a ausência de uma informação em um campo, que seria necessária para manter a Consistência dos registros, é também um problema de Completeza. Além disso, mesmo que os registros estejam completos e consistentes, os eventos por eles reportados devem ter credibilidade, ou seja, a informação deve ser considerada verdadeira e acreditável pelo auditor. Qualquer um desses problemas será considerado e avaliado como um problema de Acurácia.

A Acurácia deve ser avaliada em relação aos valores preenchidos nos campos de cada registro e, assim como na análise de Completeza, existem alguns campos nos quais a presença de erros ou ausência de informação compromete o registro inteiro. Como o indicador não objetiva medir a dificuldade ou facilidade de corrigir tais registros, quando encontrados problemas em campos com dados obrigatórios o percentual de Acurácia do registro inteiro será 0%.

Quando os problemas de Acurácia forem detectados em dados adicionais, o percentual de Acurácia do registro será igual ao percentual de campos do registro considerados acurados pelo auditor. De forma prática a verificação de Acurácia se concentra em detectar ausências de valores requeridos ou relacionamentos necessários, valores inválidos, combinações inválidas de valores válidos e resultados não razoáveis. No quadro 3.12 é apresentada a ficha do atributo Acurácia.

Quadro 3.12 – Ficha do atributo Acurácia (continua)

Sigla	Acur
Nome	Acurácia
Instrumentos de coleta	Ficha de registro de falhas detectadas Ficha de registro de manutenções efetuadas

Quadro 3.12 – Ficha do atributo Acurácia (continua)

Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção
Variáveis	<p>Relação de falhas de Relevância crítica, relação de falhas de Relevância alta, relação de falhas de Relevância média, relação de falhas de Relevância baixa, relação de manutenções de Relevância crítica, relação de manutenções de Relevância alta, relação de manutenções de Relevância média, relação de manutenções de Relevância baixa.</p> <p>Dados obrigatórios de manutenção: Número do registro de Manutenção, Identificação/Tag do Equipamento, Número do registro de falha, Datas da manutenção, Categoria de manutenção, <i>Down time</i>, Tempo de manutenção ativa.</p> <p>Dados adicionais de manutenção: Prioridade da manutenção, Periodicidade (manutenção oriunda de plano), Atividade de Manutenção, Impacto da manutenção na operação da planta, Subunidade atendida pela manutenção, Item(s) manutenível(is) atendido(s) pela manutenção, Sobressalente utilizado, Localização do sobressalente, Homem-hora de manutenção por disciplina, Homem-hora total de manutenção, Equipamentos de manutenção utilizados, <i>Lead Time</i>, <i>Turn Around Time</i> (TAT), Atrasos da manutenção, Informações Adicionais necessárias para a gestão da manutenção.</p> <p>Dados obrigatórios de falha: Número do registro de falha, Identificação/Tag do Equipamento, Datas da Falha, Modo de Falha, Método de Detecção, Impacto da falha na função do equipamento.</p> <p>Dados adicionais de falha: Condição de operação quando a falha foi detectada, Impacto da falha na segurança da planta, Impacto da falha nas operações da planta, Mecanismo de Falha, Causa da Falha, Subunidade em pane, Item(s) Manutenível(is) em pane, Informações Adicionais (descrição do evento em texto livre), Outras informações adicionais necessárias.</p>
Variáveis Calculadas	<p>\bar{X}_{Falha_Crit} → Acurácia média dos registros de falha com Relevância crítica</p> <p>\bar{X}_{Falha_Alt} → Acurácia média dos registros de falha com Relevância alta</p> <p>\bar{X}_{Falha_Med} → Acurácia média dos registros de falha com Relevância média</p> <p>\bar{X}_{Falha_Bai} → Acurácia média dos registros de falha com Relevância baixa</p> <p>\bar{X}_{Manut_Crit} → Acurácia média dos registros de manutenção com Relevância crítica</p> <p>\bar{X}_{Manut_Alt} → Acurácia média dos registros de manutenção com Relevância alta</p> <p>\bar{X}_{Manut_Med} → Acurácia média dos registros de manutenção com Relevância média</p> <p>\bar{X}_{Manut_Bai} → Acurácia média dos registros de manutenção com Relevância baixa</p>

Quadro 3.12 – Ficha do atributo Acurácia (conclusão)

Fatores utilizados	<p>peso_Rv_C → Fator para registros com Relevância crítica (100%).</p> <p>peso_Rv_A → Fator para registros com Relevância alta (90%)</p> <p>peso_Rv_M → Fator para registros com Relevância média (75%)</p> <p>peso_Rv_B → Fator para registros com Relevância baixa (50%)</p> <p>peso_fal → Fator aplicado para ponderação dos registros de falhas.</p> <p>peso_man → Fator aplicado para ponderação dos registros de manutenção.</p>
Saídas	<p>Acurácia dos Registros de Falhas (Acur_Fal_{95%})</p> $Acur_Fal_{95\%} = (\bar{x}_{Falha_Crit} * peso_Rv_C + \bar{x}_{Falha_Alt} * peso_Rv_A + \bar{x}_{Falha_Med} * peso_Rv_M + \bar{x}_{Falha_Bai} * peso_Rv_B) / (peso_Rv_C + peso_Rv_A + peso_Rv_M + peso_Rv_B)$
	<p>Acurácia dos Registros de Manutenção (Acur_Man_{95%})</p> $Acur_Man_{95\%} = (\bar{x}_{Manut_Crit} * peso_Rv_C + \bar{x}_{Manut_Alt} * peso_Rv_A + \bar{x}_{Manut_Med} * peso_Rv_M + \bar{x}_{Manut_Bai} * peso_Rv_B) / (peso_Rv_C + peso_Rv_A + peso_Rv_M + peso_Rv_B)$
	<p>Acurácia do banco (Acur_{95%})</p> $Acur = (Acur_Fal_{95\%} * peso_fal + Acur_Man_{95\%} * peso_man) / (peso_fal + peso_man)$
Tipo de medida	Baseado nos registros do sistema de gestão da manutenção
Unidade de análise	Registros de falha e manutenção.
Unidade de medida	Valor percentual (varia de 0% a 100%)
Frequência de medição	Anual
Fonte	Sistema de gestão da manutenção

3.5.6 Reputação

Cada fonte de informação que participa do processo de manutenção contribui para a qualidade do banco de dados resultante. Conforme explicado no item 3.3, é necessário definir a divisão de responsabilidades sobre os dados solicitados pelo sistema de gestão e, conforme o item 2.2, o sistema de gestão deve ser capaz de identificar o responsável por cada dado inserido.

Os principais papéis desempenhados pelos usuários que interagem com o sistema de gestão de manutenção são: registrador de falha detectada; planejador de tarefas de manutenção; responsável pela execução de tarefas de manutenção; e analista de falhas de equipamentos.

Este atributo não tem o objetivo de verificar se os papéis estão sendo desempenhados apenas pelos responsáveis definidos. Qualquer desvio nesse sentido deve ser apontado nas auditorias de Acurácia, mas apenas como uma informação extra, sem gerar nenhum ônus para o valor de qualidade da informação.

Para cada fonte, esse atributo irá avaliar a qualidade da informação obtida a partir de seus dados, com base apenas na avaliação dos demais atributos passíveis de avaliação automática (que nessa primeira versão são: Relevância; Controle de Duplicação; Completeza e Atualidade), uma vez que a realização de auditorias para qualificar cada fonte seria muito dispendiosa.

Para o registrador de falha detectada, o atributo Atualidade ($Atual_Rg_Fal_{Fonte}$) é gerado como resultado da comparação das datas de aberturas dos registros em relação às datas de detecção das falhas. Sua avaliação de Completeza ($Comp_Rg_Fal_{Fonte}$) é realizada considerando os campos: número do registro de falha; identificação/Tag do equipamento; data e hora de detecção da falha; modo de falha; método de detecção; condição de operação quando a falha foi detectada; impacto da falha na segurança da planta; impacto da falha nas operações da planta; impacto da falha na função do equipamento; informações adicionais (texto livre); e informações adicionais (outros campos). Por fim, sua avaliação de Controle de Duplicação ($Ctrl_Dup_Fal_{Fonte}$) verifica os registros de falha abertos por cada fonte.

Para o planejador de tarefas de manutenção, o atributo Atualidade ($Atual_Pl_Man_{Fonte}$) não é analisado nessa primeira versão do indicador. Em uma próxima versão, poderia ser gerado como resultado da comparação das datas de conclusão ou encerramento do registro e de encerramento da manutenção, e de eventuais diferenças entre a data de início da manutenção e a data de abertura do registro de manutenção. Sua avaliação de Completeza ($Comp_Pl_Man_{Fonte}$) é realizada considerando os campos do registro de manutenção usados para o planejamento da manutenção: número do registro de manutenção; identificação/Tag do equipamento; número do registro de falha; datas da manutenção; categoria de manutenção; prioridade da manutenção; periodicidade (manutenção oriunda de plano); impacto da manutenção na operação da planta; sobressalentes utilizados; localização dos sobressalentes; equipamentos de manutenção utilizados; e informações adicionais. Por fim, sua avaliação de Controle de Duplicação ($Ctrl_Dup_Man_{Fonte}$) verifica os registros de manutenção abertos por cada fonte.

Para o responsável pela execução de tarefas de manutenção, o atributo Atualidade ($Atual_Ex_Man_{Fonte}$) é gerado como resultado da comparação das datas de registro e conclusão da manutenção. Sua avaliação de Completeza ($Comp_Ex_Man_{Fonte}$) é realizada

considerando os campos do registro de manutenção relativos ao registro da tarefa de manutenção efetivamente realizada: datas de execução da manutenção; atividade de manutenção realizada; impacto da manutenção na operação da planta; subunidade atendida pela manutenção; item(s) manutenível(is) atendido(s) pela manutenção; sobressalentes; localização dos sobressalentes; homem-hora de manutenção por disciplina; homem-hora de manutenção total; equipamentos de manutenção utilizados; tempo de manutenção ativa; *down time*; *lead time*; *turn around time* (TAT); atrasos da manutenção; e informações adicionais. Como o papel de executante de manutenção não atua criando registros, o mesmo não é avaliado em relação ao atributo Controle de Duplicação.

Para o analista de falhas de equipamentos, o atributo Atualidade (Atual_Af_Fal_{Fonte}) não é analisado nessa primeira versão do indicador. Para uma próxima versão, poderia ser computado pela diferença entre a data de conclusão da análise de falha e a data de conclusão do serviço de manutenção, após terem sido definidos os intervalos aceitáveis e os percentuais de degradação de Atualidade aplicáveis para quando esses limites forem excedidos. Sua avaliação de Completeza (Comp_Af_Fal_{Fonte}) é realizada considerando os campos do registro de falha relativos à análise da falha: mecanismo de falha; causa da falha; subunidade em pane; item(s) manutenível(is) em pane; informações adicionais (texto livre); e informações adicionais (outros campos). Como o papel de analista de falhas de equipamentos não atua criando registros, o mesmo não é avaliado em relação ao atributo Controle de Duplicação.

Como este atributo avalia a Reputação de cada fonte, é possível inferir como está o comportamento das diversas fontes que interagem no sistema através do valor da média e desvio padrão dos percentuais de Reputação obtidos por cada fonte. O Quadro 3.13 apresenta a ficha do atributo Reputação.

Quadro 3.13 – Ficha do atributo Reputação (continua)

Sigla	Reput
Nome	Reputação
Instrumentos de coleta	Ficha de registro de falhas detectadas Ficha de registro de manutenções efetuadas
Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção
Fatores utilizados	<p>peso_Rv_C → Fator para registros com Relevância crítica (100%).</p> <p>peso_Rv_A → Fator para registros com Relevância alta (90%)</p> <p>peso_Rv_M → Fator para registros com Relevância média (75%)</p> <p>peso_Rv_B → Fator para registros com Relevância baixa (50%)</p> <p>peso_fal → Fator aplicado para ponderação dos registros de falhas.</p> <p>peso_man → Fator aplicado para ponderação dos registros de manutenção.</p>

Quadro 3.13 – Ficha do atributo Reputação (continua)

Variáveis	<p>Relação de falhas de Relevância crítica, relação de falhas de Relevância alta, relação de falhas de Relevância média, relação de falhas de Relevância baixa, relação de manutenções de Relevância crítica, relação de manutenções de Relevância alta, relação de manutenções de Relevância média, relação de manutenções de Relevância baixa.</p> <p>Registrador de falha detectada: Identificação do registrador da falha, data e hora de abertura do registro de falha, número do registro de falha; identificação/Tag do equipamento; data e hora de detecção da falha; modo de falha; método de detecção; condição de operação quando a falha foi detectada; impacto da falha na segurança da planta; impacto da falha nas operações da planta; impacto da falha na função do equipamento; informações adicionais (texto livre); e informações adicionais (outros campos).</p> <p>Planejador de tarefas de manutenção: Número do registro de manutenção; identificação/Tag do equipamento; número do registro de falha; datas da manutenção; categoria de manutenção; prioridade da manutenção; periodicidade (manutenção oriunda de plano); impacto da manutenção na operação da planta; sobressalentes utilizados; localização dos sobressalentes; equipamentos de manutenção utilizados; e informações adicionais.</p> <p>Responsável pela execução de tarefas de manutenção: Data e hora de registro da manutenção; data e hora de conclusão da manutenção; data e hora de registro do término da falha; data e hora do término da falha; datas de execução da manutenção; atividade de manutenção realizada; impacto da manutenção na operação da planta; subunidade atendida pela manutenção; item(s) manutenível(is) atendido(s) pela manutenção; sobressalentes; localização dos sobressalentes; homem-hora de manutenção por disciplina; homem-hora de manutenção total; equipamentos de manutenção utilizados; tempo de manutenção ativa; <i>down time</i>; <i>lead time</i>; <i>turn around time</i> (TAT); atrasos da manutenção; e informações adicionais.</p> <p>Analista de falhas de equipamentos: Mecanismo de falha; causa da falha; subunidade em pane; item(s) manutenível(is) em pane; informações adicionais (texto livre); e informações adicionais (outros campos).</p>
Variáveis calculadas	<p>Atual_Rg_Fal_{Fonte} → Atualidade dos campos de abertura dos registros de falha da fonte analisada.</p> <p>Comp_Rg_Fal_{Fonte} → Completeza dos campos de abertura dos registros de falha da fonte analisada.</p> <p>Ctrl_Dup_Fal_{Fonte} → Controle de Duplicação dos registro de falha da fonte analisada.</p> <p>Comp_Pl_Man_{Fonte} → Completeza dos campos de abertura dos registro de manutenção da fonte analisada.</p> <p>Ctrl_Dup_Man_{Fonte} → Controle de Duplicação dos registro de manutenção da fonte analisada.</p> <p>Atual_Ex_Man_{Fonte} → Atualidade dos campos de registro das manutenções realizadas sob responsabilidade da fonte analisada.</p> <p>Comp_Ex_Man_{Fonte} → Completeza dos campos de registro das manutenções realizadas sob responsabilidade da fonte analisada.</p> <p>Comp_Af_Fal_{Fonte} → Completeza dos campos de análise de falha sob responsabilidade da fonte analisada.</p>

Quadro 3.13 – Ficha do atributo Reputação (conclusão)

Saídas	<p>Reputação da fonte (Reput_{Fonte})</p> <p>Reput_{Fonte} := Se papel_fonte = “Registrador de falha detectada” Então R=1 Senão R=0 Fim Se Se papel_fonte = “Planejador de tarefas” Então P=1 Senão P=0 Fim Se Se papel_fonte = “Responsável pela execução” Então E=1 Senão E=0 Fim Se Se papel_fonte = “Responsável pela análise” Então A=1 Senão A=0 Fim Se</p> $\text{Reput}_{\text{Fonte}} = (\text{Peso_fal} * (\text{R} * (\text{Atual_Rg_Fal}_{\text{Fonte}} + \text{Comp_Rg_Fal}_{\text{Fonte}} + \text{Ctrl_Dup_Fal}_{\text{Fonte}}) + \text{A} * \text{Comp_Af_Fal}_{\text{Fonte}}) + \text{Peso_man} * (\text{P} * (\text{Comp_Pl_Man}_{\text{Fonte}} + \text{Ctrl_Dup_Man}_{\text{Fonte}}) + \text{E} * (\text{Atual_Ex_Man}_{\text{Fonte}} + \text{Comp_Ex_Man}_{\text{Fonte}}))) / ((\text{Peso_fal} * (\text{R} + \text{A} - \text{R} * \text{A}) + \text{Peso_man} * (\text{P} + \text{E} - \text{P} * \text{E})) * (\text{3R} + \text{A} + \text{2P} + \text{2E}))$
	<p>Reputação (Reput)</p> $\mu_{\text{Fonte}} = \frac{1}{n} \sum \text{Reput}_{\text{Fonte}}$ $\sigma = \sqrt{\frac{\sum (x - \mu_{\text{Fonte}})^2}{n}}$ <p>Reput = $\mu_{\text{Fonte}} \pm \sigma$</p>
Tipo de medida	Baseado nos registros do sistema de gestão da manutenção
Unidade de análise	Registros de falha e manutenção.
Unidade de medida	Valor percentual (varia de 0% a 100%)
Frequência de medição	Mensal (gerado no dia 4 de cada mês)
Fonte	Sistema de gestão da manutenção

Este atributo avalia a Reputação de cada fonte em relação aos registros criados nos últimos 12 meses (de acordo com a estratégia de cada empresa, podem ser escolhidos períodos diferentes) e em sua primeira versão, todos os seus atributos receberam o mesmo peso. Com o uso, podem ser definidos pesos diferentes para computar a contribuição dos atributos que sejam julgados como mais importantes para a qualidade da coleta.

3.6 Primeira versão do indicador de qualidade da informação da fase de coleta

Com base nas medições dos seis atributos selecionados é então gerado o valor de qualidade da informação em relação à fase de coleta. Relembrando, o atributo Relevância define 8 sub-populações, 4 para os registros de falha e 4 para os registros de manutenção, com pesos diferentes de contribuição para a qualidade da informação. Cada sub-população é então avaliada em relação ao Controle de Duplicação, Completeza, Atualidade e Acurácia e sofre os efeitos relativos à Reputação de cada uma de suas fontes.

Para os registros de falha, a Acurácia do nível é multiplicada pela soma do Controle de Duplicação com a soma de duas somatórias: a dos produtos da Reputação de cada registrador de falha com as somas dos valores de Atualidade e Completeza dos seus registros para aquele nível; e a dos produtos da Reputação de cada analista de falhas com a Completeza dos seus registros para aquele nível. O resultado é então dividido pelo triplo da quantidade de registros de falha daquele nível mais 1.

Para os registros de manutenção, a Acurácia do nível é multiplicada pela soma do Controle de Duplicação com a soma de duas somatórias: a dos produtos da Reputação de cada planejador de tarefas de manutenção com a Completeza dos seus registros para aquele nível; e a dos produtos da Reputação de cada responsável pela execução de tarefas de manutenção com as somas dos valores de Atualidade e Completeza dos seus registros para aquele nível. O resultado é então dividido pelo triplo da quantidade de registros de manutenção daquele nível mais 1.

Os resultados das avaliações de falha e manutenção de cada nível são então ponderados pelos pesos relativos aos registros de falhas e manutenções, respectivamente, para gerar o valor final de qualidade de informação para aquele nível. E, por fim, as QI resultantes de cada nível são então combinadas para gerar o valor de QI relativo ao o processo de coleta.

A ficha completa da primeira versão do Indicador de Qualidade da Informação para a fase de Coleta de dados é exibida no Quadro 3.14.

Quadro 3.14 – Ficha do Indicador de Qualidade da Informação – Coleta de dados (continua)

Sigla	QI _{Coleta}
Nome	Qualidade da Informação – Coleta de dados
Atributos Avaliados	Relevância, Controle de Duplicação, Completeza, Atualidade, Acurácia e Reputação.
Fatores utilizados	<p>peso_Rv_C → Fator para registros com Relevância crítica (100%).</p> <p>peso_Rv_A → Fator para registros com Relevância alta (90%)</p> <p>peso_Rv_M → Fator para registros com Relevância média (75%)</p> <p>peso_Rv_B → Fator para registros com Relevância baixa (50%)</p> <p>peso_fal → Fator aplicado para ponderação dos registros de falhas.</p> <p>peso_man → Fator aplicado para ponderação dos registros de manutenção.</p>
Saídas	<p>Qualidade da Informação de registros com Relevância Crítica</p> $QI_C = \left(\left(\text{Peso_fal} * \text{Acur_fal_C} * (\text{Ctrl_Dup_Fal_C} + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Rg_Fal_C} + \text{Comp_Rg_Fal_C}))) + \sum \text{Reput}_{\text{Fonte}} * \text{Comp_Af_Fal_C} \right) / (1 + 3 * \text{qtd_reg_fal_C}) \right) + \left(\left(\text{Peso_man} * \text{Acur_man_C} * (\text{Ctrl_Dup_Man} + \sum (\text{Reput}_{\text{Fonte}} * \text{Comp_Pl_Man_C}) + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Ex_Man_C} + \text{Comp_Ex_Man_C}))) \right) / (1 + 3 * \text{qtd_reg_man_C}) \right) \right) / (\text{Peso_fal} + \text{Peso_man})$
	<p>Qualidade da Informação de registros com Relevância Alta</p> $QI_A = \left(\left(\text{Peso_fal} * \text{Acur_fal_A} * (\text{Ctrl_Dup_Fal_A} + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Rg_Fal_A} + \text{Comp_Rg_Fal_A}))) + \sum \text{Reput}_{\text{Fonte}} * \text{Comp_Af_Fal_A} \right) / (1 + 3 * \text{qtd_reg_fal_A}) \right) + \left(\left(\text{Peso_man} * \text{Acur_man_A} * (\text{Ctrl_Dup_Man} + \sum (\text{Reput}_{\text{Fonte}} * \text{Comp_Pl_Man_A}) + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Ex_Man_A} + \text{Comp_Ex_Man_A}))) \right) / (1 + 3 * \text{qtd_reg_man_A}) \right) \right) / (\text{Peso_fal} + \text{Peso_man})$
	<p>Qualidade da Informação de registros com Relevância Média</p> $QI_M = \left(\left(\text{Peso_fal} * \text{Acur_fal_M} * (\text{Ctrl_Dup_Fal_M} + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Rg_Fal_M} + \text{Comp_Rg_Fal_M}))) + \sum \text{Reput}_{\text{Fonte}} * \text{Comp_Af_Fal_M} \right) / (1 + 3 * \text{qtd_reg_fal_M}) \right) + \left(\left(\text{Peso_man} * \text{Acur_man_M} * (\text{Ctrl_Dup_Man} + \sum (\text{Reput}_{\text{Fonte}} * \text{Comp_Pl_Man_M}) + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Ex_Man_M} + \text{Comp_Ex_Man_M}))) \right) / (1 + 3 * \text{qtd_reg_man_M}) \right) \right) / (\text{Peso_fal} + \text{Peso_man})$
	<p>Qualidade da Informação de registros com Relevância Baixa</p> $QI_B = \left(\left(\text{Peso_fal} * \text{Acur_fal_B} * (\text{Ctrl_Dup_Fal_B} + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Rg_Fal_B} + \text{Comp_Rg_Fal_B}))) + \sum \text{Reput}_{\text{Fonte}} * \text{Comp_Af_Fal_B} \right) / (1 + 3 * \text{qtd_reg_fal_B}) \right) + \left(\left(\text{Peso_man} * \text{Acur_man_B} * (\text{Ctrl_Dup_Man} + \sum (\text{Reput}_{\text{Fonte}} * \text{Comp_Pl_Man_B}) + \sum (\text{Reput}_{\text{Fonte}} * (\text{Atual_Ex_Man_B} + \text{Comp_Ex_Man_B}))) \right) / (1 + 3 * \text{qtd_reg_man_B}) \right) \right) / (\text{Peso_fal} + \text{Peso_man})$
	<p>Qualidade da Informação</p> $QI = (QI_C * \text{peso_Rv_C} + QI_A * \text{peso_Rv_A} + QI_M * \text{peso_Rv_M} + QI_B * \text{peso_Rv_B}) / (\text{peso_Rv_C} + \text{peso_Rv_A} + \text{peso_Rv_M} + \text{peso_Rv_B})$

Quadro 3.14 – Ficha do Indicador de Qualidade da Informação – Coleta de dados (conclusão)

Instrumentos de coleta	Ficha de registro de falhas detectadas; ficha de registro de manutenções efetuadas e relatórios de auditoria.
Frequência de coleta	Fichas atualizadas diariamente no sistema de gestão da manutenção Auditorias realizadas semestralmente e anualmente
Tipo de medida	Baseado nos registros do sistema de gestão da manutenção e avaliações e estimativas geradas em auditorias
Unidade de análise	Registros de falha e manutenção
Unidade de medida	Valor percentual (0% a 100%)
Frequência de medição	Mensal (gerado no dia 4 de cada mês) Fatores provindos de auditoria atualizados semestralmente e anualmente.
Fonte	Sistema de gestão da manutenção Relatórios de auditoria

3.7 Ajustando a métrica de avaliação à percepção de qualidade da empresa

O indicador de QI proposto foi construído com base no “conhecimento expresso” na norma de qualidade (NBR ISO 9000), na norma de coleta e troca de dados de confiabilidade e manutenção para equipamentos (ISO 14224) e na literatura de qualidade da informação e manutenção. Mas para que o seu resultado possa gerar uma foto compreensível da realidade, é preciso antes que os especialistas de confiabilidade da empresa gerem ou forneçam pelo menos três cenários diferentes em relação ao controle dos eventos de falha e manutenção: um que seja ruim; um que seja considerado médio ou aceitável; e um em que seja bom ou excelente.

Para cada cenário deve-se proceder da seguinte forma: estabelecer qual o valor esperado para o indicador de qualidade da informação; comparar com o valor efetivamente gerado pelo indicador para esses cenários; e com base nas diferenças observadas, deve-se verificar e ajustar os pesos de contribuição de cada atributo utilizado.

Após essa etapa, é necessária uma apresentação para o corpo gerencial que irá efetivamente acompanhar e cobrar o uso desse indicador, de forma a confirmar ou ajustar os fatores de contribuição indicados por cada especialista de confiabilidade.

Nesse momento, é importante que seja enfatizado que não se trata apenas da criação de mais uma meta a ser cumprida, e sim do estabelecimento de prioridades claras em relação ao que é esperado da atividade de manutenção, pois cada controle tem um custo e uma carga de trabalho associados. De outra forma, na carga de trabalho total dos empregados de uma

planta, devem ser previstos os tempos necessários para o correto acompanhamento e registro de cada controle definido.

3.8 Conduzindo futuras evoluções baseadas em conhecimento

A primeira versão do indicador de QI aqui proposto concentra-se no conhecimento já formalizado sobre o problema de coleta de dados de campo visando sua aplicação em estudos de confiabilidade e manutenibilidade.

Isso é suficiente para avaliar e controlar alguns aspectos primários, mas conforme já mencionado, vários problemas de qualidade ainda poderão ocorrer sem serem apontados.

Detectar e acusar esses problemas, entretanto, exige a captura e aplicação de conhecimentos que normalmente estão diluídos ao longo da organização. Conforme mostrado na Figura 2.13, parte desse conhecimento pode ser convertido em propriedades de coluna, análises de estrutura, regras de dados simples e complexas e análises de regras de valores para aprimorar a indicação de qualidade da informação. Para essa etapa, é indicada a aplicação do método CommonKads.

Uma vez que o processo de medição de QI já tenha iniciado, pela construção dos modelos de contexto do CommonKads é possível inferir qual é a aderência atual ao processo de controle de eventos especificado pela empresa.

Além disso, a avaliação de Relevância prevista por este trabalho ajuda a identificar em quais classes de equipamentos o controle de falhas e manutenções é mais crítico e merece receber uma atenção diferenciada. Para esses casos, recomenda-se a aplicação de uma tarefa de Monitoramento, que tem como objetivo analisar um processo em andamento para descobrir se está se comportando de acordo com as expectativas.

É importante ressaltar que a tarefa de Monitoramento trabalha com dados históricos sobre o sistema monitorado e indica as discrepâncias encontradas, mas não analisa as suas causas.

Para criar uma tarefa de Monitoramento é necessário identificar: os parâmetros pelos quais um comportamento anormal pode ser percebido; as normas que especificam os valores ou faixas de valores esperados para os parâmetros em caso de comportamento normal (ou anormal); e a forma que a discrepância encontrada deve ser indicada.

Conforme descrito no item 2.4.1, essa tarefa é parte do modelo de conhecimento previsto pelo CommonKads, composto por três categorias de conhecimento (ou sub-modelos): o Conhecimento da Tarefa (contém as metas, e estruturas de decomposição e controle); o

Conhecimento (ou modelo) de Inferência (composto por inferências básicas e papéis); e o Conhecimento do Domínio (que abrange tipos, regras e fatos).

O método básico da tarefa de monitoramento é apresentado na Figura 3.2.

```

TAREFA monitoramento;
  PAPÉIS:
    ENTRADA:
      Dados-históricos: "dados dos ciclos de monitoramento anteriores";
    SAÍDA:
      Discrepância: "indicação de desvio do comportamento do sistema";
  FIM TAREFA monitoramento;

MÉTODO-TAREFA monitoramento-guiado-por-dados;
  REALIZA: monitoramento;
  DECOMPOSIÇÃO:
    INFERÊNCIAS:
      selecionar, especificar, comparar, classificar;
    FUNÇÕES-DE-TRANSFERÊNCIA: receber;
  PAPÉIS:
    INTERMEDIATE:
      observação: "algum dado observado sobre o sistema";
      parâmetro: "variável para examinar o comportamento anormal";
      norma: "valor normal esperado para o parâmetro";
      diferença: "uma indicação do desvio observado em relação a norma";
  ESTRUTURA-DE-CONTROLE:
    receber(nova-observação);
    selecionar(nova-observação -> parâmetro);
    especificar(parâmetro -> norma);
    comparar(norma + observação -> diferença);
    classificar(diferença + dados-históricos -> discrepância);
    dados-históricos := observação ADICIONAR dados-históricos;
  FIM MÉTODO-TAREFA monitoramento-guiado-por-dados;

```

Figura 3.2 – Especificação de método para o método guiado por dados para monitoramento.

Fonte: SCHREIBER *et al.* (2000), p. 144, Fig. 6.11

Esse método é guiado por evento, tornando-se ativo a cada entrada de novos dados. Isso ocorre através da função de transferência “receber” comandada por um agente externo (um usuário humano ou outro sistema). Cada novo dado recebido é processado por quatro inferências:

- Selecionar: Um parâmetro do sistema é selecionado para avaliar os novos dados.
- Especificar: Um valor esperado é especificado para o parâmetro. Normalmente, um sistema de monitoramento tem como conhecimento de domínio um modelo de sistema, consistindo de um conjunto de parâmetros. Para cada parâmetro deve ser fornecido conhecimento sobre os valores normais do parâmetro em diferentes contextos do sistema.
- Comparar: É feita uma comparação da nova observação com o valor normal esperado

e é gerada uma descrição para as eventuais diferenças encontradas.

- Classificar: A diferença é enquadrada em uma classe de discrepância. Frequentemente, dados de ciclos prévios de monitoramento são usados nesta inferência.

A estrutura de inferência para o método descrito na Figura 3.2 é exibida na Figura 3.3.

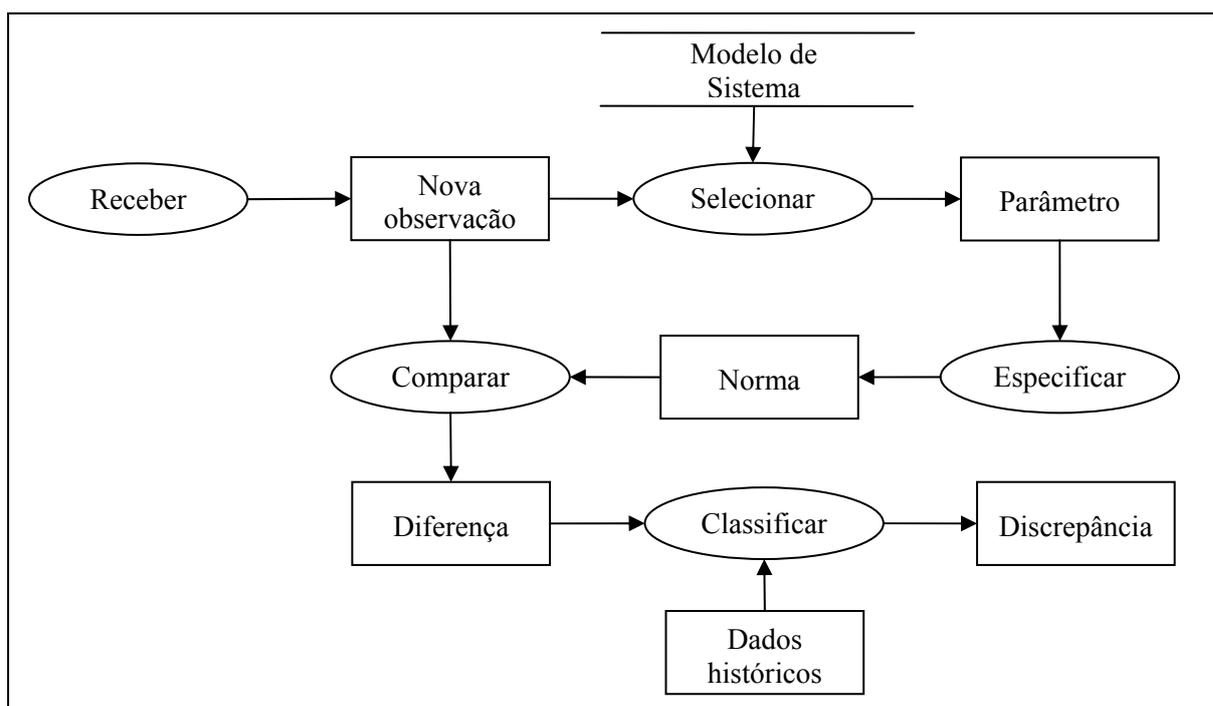


Figura 3.3 – Estrutura de inferência para o esquema de tarefa de monitoramento.

Fonte: SCHREIBER *et al.* (2000), p. 145, Fig. 6.12

Como o indicador de QI deve ser gerado periodicamente, a função de transferência “receber” pode ser substituída por uma função “obter”, programada para adquirir e avaliar os dados a intervalos regulares de tempo, por exemplo, todo quarto dia do mês. Nesse caso, o método passa a ser conhecido como monitoramento guiado por modelo (SCHREIBER *et al.*, 2000).

Dos atributos de QI previstos para a fase de coleta, a Consistência e a Credibilidade destacam-se como bons candidatos para uma tarefa de Monitoramento. Dentre as classes de equipamentos críticas, sugere-se escolher uma para servir de protótipo desse processo. Para essa classe, então, deve ser explicitado o comportamento esperado e os valores e relacionamentos de valores de atributos que não são pertinentes.

Por exemplo, o Quadro 2.12 oferece duas situações de combinações de valores de modo de falha e mecanismo de falha que poderiam ser monitoradas como problemas de Consistência, ou mesmo de Clareza:

- quando o modo de falha do equipamento for codificado como “vazamento”, deve-se usar um mecanismo de falha que seja mais relacionado com a causa da falha. Na ausência de maiores detalhes, deve ser escolhido o valor “desconhecido”;
- se o modo de falha no nível do equipamento também for vibração, deve-se usar um mecanismo de falha que seja mais relacionado com a causa da falha. Na ausência de maiores detalhes, deve ser escolhido o valor “desconhecido”.

Para facilitar o desenvolvimento da tarefa de monitoramento, SCHREIBER *et al.* (2000) sugere diversas atividades, técnicas e diretrizes, que também podem auxiliar nas sessões de elicitación de conhecimento, fundamentais a esse processo.

É importante ressaltar que como a elicitación permite gerar a partir do conhecimento das fontes disponíveis o conjunto de regras e informações empregado para avaliação de QI, esse material pode ser utilizado também para nivelar o conhecimento dos atores que participam do processo de manutenção e então incrementar a qualidade das informações geradas pelos controles de falha e manutenção.

4 CONCLUSÃO

4.1 Conclusões

O conhecimento utilizado para a análise da qualidade da informação relativa a um conjunto de dados coletados envolve a aplicação de conhecimentos oriundos de diversas áreas. Em uma tarefa dessa natureza, a primeira questão que precisa ser respondida é se os resultados esperados compensam o investimento demandado.

Para responder essa pergunta, foi analisada a necessidade das informações relativas aos eventos de falha e manutenção de equipamentos industriais aplicados na indústria petrolífera. Com base na revisão bibliográfica sobre a evolução da área de manutenção, foi constatado que a expectativa atual de uma ação pró-ativa exige o acompanhamento da “saúde” dos equipamentos e que o sucesso das ações empregadas para incremento da confiabilidade e disponibilidade das plantas industriais está diretamente relacionado à qualidade dos diagnósticos realizados com base no comportamento dos equipamentos.

Esse comportamento costuma ser registrado em um sistema informatizado, que é responsável por auxiliar a gestão do processo de manutenção, no qual ficam armazenados dados relativos às falhas detectadas e manutenções efetuadas.

Uma vez justificada a necessidade desse acompanhamento, é preciso estabelecer quais são as informações que devem ser obtidas. Para isso, foi utilizada a norma ISO 14224 que lista diversas informações que podem ser obtidas com base no controle de eventos de manutenção e o conjunto de dados requeridos para a produção dessas informações.

A partir desse cenário, havia a necessidade de explicitar como o conhecimento poderia ser aplicado para avaliar a qualidade das informações presentes nesse conjunto de dados. Para isso, primeiramente foram estabelecidas, com base em uma revisão bibliográfica, as diferenças entre os significados desses termos no universo desse trabalho:

- dados: informações fornecidas por alguma fonte, obedecendo um formato pré-determinado para utilização em um conjunto de aplicações previsto durante sua concepção;
- informação: idéias geradas como fruto de um processo de conhecimento e transmitidas em um processo de comunicação, mediado ou imediato. Termo usado também para designar um dado obtido a partir do processamento de outros dados;

- conhecimento: resultado emergente da interação entre um ente cognoscente e seu ambiente e que permite selecionar ações e gerar novas informações.

Então, foi apresentada a metodologia CommonKADS, que especifica todas as etapas necessárias para a construção de um Sistema Baseado em Conhecimento, e realizado um estudo para identificar quais aspectos devem ser envolvidos em uma avaliação de qualidade da informação.

Como ponto de partida foi examinado os requisitos especificados na norma NBR ISO 9000 sobre os aspectos envolvidos na medição de qualidade, seguido de uma revisão bibliográfica sobre a área de qualidade da informação.

Como principais frutos desse estudo foram identificados que:

- ✓ a definição do processo de registro das ocorrências de manutenção sofridas pelos equipamentos, dos fatores de influência desse processo e dos atores envolvidos é fundamental para o correto entendimento das informações armazenadas;
- ✓ a análise da qualidade da informação pode ser separada e avaliada de acordo com o ciclo de vida da informação, o que está de acordo com a proposta deste trabalho de restringir a avaliação da qualidade ao processo de coleta;
- ✓ a avaliação da qualidade da informação pode ser realizada em níveis crescentes de profundidade, permitindo que a implantação de um processo de avaliações periódicas foque, inicialmente, em um conjunto pequeno de aspectos, o qual deverá ser refinado de acordo com as melhoras obtidas na qualidade dos registros; e
- ✓ o conjunto de atributos relevantes para a avaliação da qualidade da informação relativa a um processo de coleta abrange: Clareza, Objetividade, Acurácia, Completeza, Consistência, Controle de Duplicação, Probidade do coletor, Reputação, Atualidade, Relevância, e Credibilidade.

Por fim, foi estudada a pertinência e aplicação desses atributos para o processo de avaliação dos registros de falha e manutenção de equipamentos da indústria petrolífera, com indicação clara de como os mesmos podem ser utilizados.

Devido às diferenças entre fatores de influência de cada atributo, estes podem ser separados em três níveis: aqueles que a avaliação necessita de um agente humano (Clareza, Acurácia e Probidade do coletor), aqueles que a avaliação pode ser auxiliada por um KBS (Credibilidade, Completeza e Consistência) e aqueles cuja avaliação pode ser realizada com base em um conjunto de regras mais simples (Objetividade, Controle de Duplicação, Reputação, Atualidade e Relevância).

Foi estabelecido então um conjunto inicial de atributos a ser utilizado e quais aspectos desses atributos poderiam ser considerados.

Uma vez estabelecido o esquema para essa análise de qualidade da informação, foi proposta a construção de um indicador cujo valor é gerado a partir da composição dos resultados das análises de cada atributo empregado. Para que seja possível associar o significado dos valores obtidos com esse indicador a realidade por ele retratada, devem ser gerados alguns cenários como exemplo e identificados que valores eles produziram no indicador. Por meio desse processo, deve ser realizado o ajuste de sensibilidade do indicador através do valor dos pesos utilizados.

A combinação entre os resultados gerados diretamente pelo sistema e aqueles oriundos de auditorias tanto fornece subsídios para a evolução do indicador quanto aumenta a confiança nos resultados por ele apresentados.

Dessa forma, a elicitação do conhecimento necessária para construção de um KBS pode ser realizada com base nos próprios relatórios de auditoria, simplificando o processo de modelagem do conhecimento. É importante ressaltar que esse processo, como depende da interpretação subjetiva do engenheiro do conhecimento, está sujeito a falhas e por isso demanda uma comparação entre os seus resultados e os que seriam produzidos pela análise de um especialista da área.

Finalizando, a estrutura criada neste trabalho permite tanto entregar um mecanismo de controle de qualidade das informações a gerência das empresas (com o qual é possível inferir a efetividade das ações para incremento da QI) quanto resguardar o conhecimento de seus especialistas relativo a essa análise de qualidade em uma estrutura automatizada.

4.2 Sugestões para Trabalhos Futuros

A avaliação da qualidade da informação relativa ao processo de coleta não parece diferir significativamente em relação a outros universos. Entretanto, pode ser interessante incluir alguns outros atributos, como por exemplo, a quantidade de pessoas envolvidas entre o relato e o registro de um evento, a precisão do dado registrado, o tipo e quantidade de mídias utilizadas entre a percepção do dado e seu registro no sistema de armazenamento e a influência da qualidade dos dados de cadastro de Tag e equipamento.

Ainda para os equipamentos da indústria petrolífera, existe a necessidade de identificar a qualidade final das informações geradas com base nos controles de evento, o que demanda um

estudo dos fatores de influência relativo às demais fases de tratamento dos dados e criação de indicadores capazes de qualificar cada informação produzida.

Outra linha a ser explorada abrange a influência dos aspectos subjetivos na qualidade como: a motivação dos funcionários, a qualidade da comunicação interpessoal, o nível de conhecimento e experiência da equipe, o tempo de trabalho na planta, o tempo sem progressão funcional, etc.

Por fim, é desejável que sejam realizados testes de aplicação do indicador, bem como uma pesquisa para avaliação de sua aplicabilidade.

5 REFERÊNCIAS

ABNT NBR ISO 9000. **Sistemas de gestão da qualidade – Fundamentos e vocabulário**. 2ª Ed. ABNT 2005. 35p.

ACKOFF, R. L. From data to wisdom. **Journal of Applied Systems Analysis**, New York, v. 16, p. 3-9, 1989.

ADLER, M. J. **A Guidebook to learning: for a lifelong pursuit of wisdom**. New York: Macmillan, 1986, p.163.

ANGELE, J.; FENSEL, D.; LANDES, D.; STUDER, R. Developing Knowledge-Based Systems with MIKE. **Automated Software Engineering**, v. 5, n. 4, p.389-418, 1998. Disponível em: <<http://www.springerlink.com/content/182212008q591p16/fulltext.pdf>>. Acesso em: 13 mar. 2007.

AI-HAKIM, L. **Information Quality Management: Theory and Applications**. Australia: University of Southern Queensland, Idea Group Publishing, 2007, 301p.

ALEXANDER, J.E.; TATE, M.A. **Web Wisdom: How to evaluate and create information quality on the Web**. 1. ed. Mahwah, NJ: Lawrence Erlbaum, 1999, 152p.

ALKAIM, J. L. **Metodologia para incorporar conhecimento intensivo às tarefas de manutenção centrada na confiabilidade aplicada em ativos de sistemas elétricos**. Florianópolis, SC: EdUFSC, 2003. Originalmente apresentada como Tese de Doutorado em Engenharia de produção, Universidade Federal de Santa Catarina programa de pós-graduação em engenharia de produção, 2003, 239p. Disponível em: <<http://www.tede.ufsc.br/teses/PEPS3100.pdf>>. Acesso em: 20 out. 2007.

BALLOU, D.P.; PAZER, H.L. Modeling data and process quality in multi-input, multi-output information systems. **Management Science**, v.31, n.2, p.150-62, 1985.

BANDURA, A. **Social learning theory**. Englewood Cliffs, New Jersey: Prentice-Hall, 1977. p. 247.

BARRETO JÚNIOR., J. **Qualidade de Software**. 1997. Disponível em: <<http://www.geocities.com/josedearaujo/>>. Acesso em 2 jan. 2008.

BATINI, C.; SCANNAPIECO, M. **Data Quality: Concepts, Methodologies and Techniques**. Berlin, Heidelberg: Springer-Verlag, 2006. p.262.

BELLINGER, G.; CASTRO, D.; MILLS, A. **Data, information, knowledge, and wisdom**. 2004. Disponível em: <www.systems-thinking.org/dikw/dikw.htm>. Acesso em: 5 fev. 2006.

BITTERER, A. **Gartner's Data Quality Maturity Model**. 2007. Disponível em: <<http://2007.dataqualitysummit.com/static/pdf/Andreas%20Bitterer%20GARTNER%20-%20DQS%20final.pdf>>. Acesso em: 04 abr. 2008.

BOBROWSKI, M.; MARRE, M.; YANKELEVICH, D. A Homogeneous framework to measure data quality. In: INTERNATIONAL CONFERENCE ON INFORMATION QUALITY, 1999, CAMBRIDGE: MIT. **Anais...** Cambridge: International Conference on Information Quality, 1999, p.115-125.

BOULDING, K. Notes on the information concept. **Explorations**, Pakistan, v. 6, p. 103-122, 1955.

BRIEN, J. **Introduction to Information Systems in Business Management**, Boston: Irwin, 1991. p. 213.

BUCKLAND, J. A.; FOWINKLE, R.; SHROYER, L.; RICE, F. V. **Total Quality Management in Information Services**. EUA: Wiley, p.324, 1991. ISBN: 9780471560463

CALAZANS, A. T. S. Qualidade da informação: conceitos e aplicações. **TransInformação**, v.20, n.1, p. 29-45, 2008.

CAPURRO, R.; HJORLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, 2007.

CARIBE, R. de C. do V. Sistema de indicadores: uma introdução. **Revista Digital de Biblioteconomia e Ciência da Informação**. Campinas. v. 6, n. 2. p. 1-23. 2009. ISSN: 1678-765X. Disponível em: <<http://polaris.bc.unicamp.br/seer/ojs/viewarticle.php?id=159&layout=abstract>>. Acesso em: 23 jun. 2009.

CARVALHO, E. N. **Uma Revisão Crítica do Emprego de Bancos de Dados de Falhas em Análises Probabilísticas de Segurança de Plantas Nucleares e Químicas**. Rio de Janeiro, RJ: UFRJ Editora, 1993. Originalmente apresentada como dissertação de Mestrado em Ciências e Engenharia Nuclear, Universidade Federal do Rio de Janeiro, Coordenação do Programa de pós-graduação em Engenharia, 1993. 106p.

CLANCEY, W. J. The Knowledge level reinterpreted: modeling how systems interact. **Machine Learning**, United States, v. 4, n. 3-4, p. 285-291, Dec. 1989. Disponível em:

<<http://www.springerlink.com/content/r3j5l76435q5w230/>>. Acesso em: 14 jun. 2007. doi: 10.1023/A:1022602705960.

CLEVELAND, H. Information as a resource. **Futurist**, v. 16, n. 6, p. 34-39, 1982.

CONNELL, N. A. D.; KLEIN, J. H.; POWELL P. L. It's tacit knowledge but not as we know it: redirecting the search for knowledge. **Journal of the Operational Research Society**, England, v. 54, n. 2, p. 140-152, 2003.

COOLEY, M. **Architect or bee**: the human price of technology. 2.ed. London: The Hogarth Press, 1987. p.357.

COOPER, D. R.; SCHINDLER, P. S. **Métodos de Pesquisa em Administração**, 7. ed. Porto Alegre: Bookman, 2003. 640 p. ISBN: 8536301171.

DANIEL, E.H. Quality Control of Documents. **Library Trends**. v.41, n.4, p. 644-664. 1993. Disponível em: <http://www.ideals.uiuc.edu/bitstream/handle/2142/7868/librarytrendsv41i4_opt.pdf?sequence=3#page=109>. Acesso em: 14 jun. 2008.

DRUCKER, P. **Sociedade pós-capitalista**. 5.ed. São Paulo: Pioneira, 1993. p.177.

ELIOT, T. S. **The rock**. London: Faber & Faber, 1934. Disponível em: <<http://www.wisdomportal.com/Technology/TSEliot-TheRock.html>>. Acesso em: 1 jan. 2009.

ENGLISH, L.P. Information Quality: Meeting Customer Needs, **DM review**, v.3, n.1, 1996.

ENGLISH, L.P. **Improving Data Warehouse and Business Information Quality**: Methods for Reducing Costs and Increasing Profits. New York, NY: John Wiley and Sons, 1999. p. 518.

EPPLER, M. **Qualitätsstandards – Ein Instrument zur Sicherung der Informationsqualität in Multimedia-Produktionen**, in: Merx, O. (Ed.) Qualitätssicherung in Multimedia-Projekten, Berlin: Springer Verlag, 1999, p. 129-148.

EPPLER, M. **Managing Information Quality**: Increasing the Value of knowledge-intensive Products and Processes. New York/Heidelberg: Springer, 2006. p. 395.

FERREIRA, A. B. H. **Dicionário Aurélio Eletrônico: século XXI**. Rio de Janeiro: Nova Fronteira, 1999.

FISHER, C., WANG, R., LAURIA, E., CHENGALUR-SMITH, S. **Introduction to Information Quality**. MIT Information Quality Program, 2005. ISBN-13: 9780977759903.

FLORIDI, L. Semantic conceptions of information. In: ZALTA, E. N. (Org.). **The Stanford Encyclopedia of Philosophy**. Stanford: The Metaphysics Research Laboratory, 2005. Disponível em: <<http://plato.stanford.edu/entries/information-semantic/>>. Acesso em: 11 nov. 2008.

FOX, C., LEVITIN, A., REDMAN, T. The notion of data and its quality dimensions. **Information Processing & Management**, v. 30, n. 1, p. 9-19, 1994.

FRAGOLA, J. R. Reliability and risk analysis data base development: An historical perspective. **Reliability Engineering and System Safety**, v. 51, n.2, p. 125-136, 1996.

FREUND, J. E. **Estatística Aplicada: Economia, Administração e Contabilidade**, 11.ed. Artmed Editora, p.536, 2006. ISBN: 9788536306674.

FRICKÉ, M. The knowledge pyramid: a critique of the DIKW hierarchy. **Journal of Information Science**, United Kingdom, v. 0, p. 1-12, Oct. 2008. Disponível em: <<http://jis.sagepub.com/cgi/rapidpdf/0165551508094050v1>> Acesso em: 11 nov. 2008.

FULLWOOD, R. R. **Probabilistic Safety Assessment in the Chemical and Nuclear Industries**. Oxford, UK: Butterworth-Heinemann, 1999. p. 514. ISBN 0750672080.

GACKOWSKI, Z. J.. Logical Interdependence of data/Information Quality Dimensions – A purpose-focused View on IQ. In: Proceedings of ninth International Conference on information Quality (ICIQ-04). 2004. Disponível em: <<http://www.iqconference.org/Documents/IQ%20Conference%202004/Papers/LogicalInterdependence.pdf>>. Acesso em: 22 jun 2008.

GACKOWSKI, Z. J. Informing systems in business environments: A purpose-focused view, **Informing Science Journal**, v.8, p.101-122, 2005a. Disponível em: <<http://inform.nu/Articles/Vol8/v8p101-122Gack.pdf>>. Acesso em: 18 ago. 2008.

GACKOWSKI, Z. J. **Operations Quality of Data and Information: Teleological Operations Research-Based Approach**, Call for Discussion. 2005b.

GERKES, M. **Information Quality Paradox of the Web**. 1997. Disponível em: <<http://izumw.izum.si/~max/paper.htm>>. Acesso em: 08 set. 2008.

GENNARI, J. H.; MUSEN, M. A.; FERGERSON, R. W.; GROSSO, W. E.; CRUBÉZY, M.; ERIKSSON, HY.; NOY, N. F.; TU, S. W. The evolution of Protégé: an environment for knowledge-based systems development. **International Journal of Human-Computer Studies**, v. 58, n. 1, 2003, p. 123. ISSN:1071-5819.

GINMAN, M. Quality information, and information for quality. In: WORMELL, I. **INFORMATION QUALITY DEFINITIONS AND DIMENSIONS**, 1989, Copenhagen. **Resumos...** Copenhagen: NORDINFO Seminar, Royal School of Librarianship, 1989. p.18.

HUANG, K.T.; LEE, Y.W.; WANG, R.Y. **Quality Information and Knowledge**. New Jersey: Prentice Hall, 1999. p.210.

ISO 14224:2006(E). **Petroleum, petrochemical and natural gas industries – Collection and exchange of reliability and maintenance data for equipment**.

ISO/TS 8000-110. **Data quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification**.

JARKE, M.; LENZERINI, M.; VASSILIOU, Y.; VASSILIADIS, P. **Fundamentals of Data Warehouses**. New York: Springer-Verlag. 2000. p.224. ISBN: 3-540-65365-1.

JARKE, M.; LENZERINI, M.; VASSILIOU, Y.; VASSILIADIS, P. **Fundamentals of Data Warehouses: 2nd revised and extended edition**. New York: Springer-Verlag. 2003. p.246. ISBN: 978-3-540-42089-7

JARKE, M.; VASSILIOU, Y. Data Warehouse Quality: A Review of the DWQ Project. In: **2ND CONFERENCE ON INFORMATION QUALITY**, 1997, Massachusetts Institute of Technology. **Resumos...** Massachusetts: 2nd Conference on Information Quality, 1997. p.185.

KAHN, B. K.; STRONG, D. M. Product and Service Performance Model for Information Quality: An Update. In: **THIRD CONFERENCE ON INFORMATION QUALITY MASSACHUSETTS INSTITUTE OF TECHNOLOGY**, 1998, Massachusetts. **Resumos ...** Third Conference on Information Quality, 1998, p. 102.

KAHN, B. K.; STRONG, D. M., WANG R. Y. Information quality benchmarks: product and service performance. **Communications of Association of computing Machinery**, v.4, p. 184-192, 2002.

KAY, R.; CECEZ-KECMANOVIC, D. When knowledge becomes information: a case of mistaken identity. In: WORKSHOP ON DATABASE AND EXPERT SYSTEMS APPLICATIONS (DEXA), 11., London, Uk, 2000. **Anais ...** Eleventh International Workshop on Database and Expert Systems Applications (DEXA). Greenwich, Computer Society Press, 2000. p. 1128-1133. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=875168>. Acesso em: 11 dec. 2008.

KÖNIGER, P.; REITHMAYER, W. **Management unstrukturierter Informationen**. Frankfurt: Campus Fachbuch. 1998, p. 284, ISBN: 3593360519.

LANZ, P.; MCFARLAND, D. Philosophical perspectives on representation, goals, and cognition. In: DEAN, J.; RITTER, H.; CRUSE, H. (Comp.). **Prerational intelligence: interdisciplinary perspectives on the behavior of natural and artificial systems**. Netherlands: Kluwer Academic Publishers, 2000. p. 247-264.

LEE, Y.W.; STRONG, D.M.; KAHN, B.; WANG, R.Y. AIMQ: a methodology for information quality assessment. **Information & management**, v.40, p. 133-146, 2002.

LEE, Y. W.; PIPINO, L. L.; FUNK, J. D.; WANG, R. Y. **Journey to Data Quality**. Massachusetts: The MIT Press, 2006, p.240.

LEHNER, F.; MAIER, R. Can information modelling be successful without a common perception of the term "information"? In: KANGASSALO, H. et al. (Comp.). **Information modelling and knowledge bases VIII**. Amsterdam, Netherlands: Ios Press, 1997. p. 181-198.

LESCA, H.; LESCA, E. **Gestion de l'information, qualité de l'information et performances de l'entreprise**. Paris: Litec, 1995. p. 209.

LIU, L.; CHI, L. N. Evolutional Data Quality: A Theory-Specific View. In: THE SEVENTH INTERNATIONAL CONFERENCE ON INFORMATION QUALITY (ICIQ-02), Cambridge, MA. **Resumos...** The Seventh International Conference on Information Quality (ICIQ-02), 2002. p. 302.

LOSHIN D. **Enterprises knowledge management: The data quality approach**. San Francisco: Morgan Kauffman, 2001. p. 493.

LOSHIN, D. **Master Data Management**. San Francisco: Morgan Kaufmann, 2008. p. 288. ISBN: 0123742250.

MACHLUP, F.; MANSFIELD, U. **The Study of information: interdisciplinary messages**. New York: Wiley, 1983. p.743.

MADNICK, S. E. Integration technology: The reinvention of the linkage between information systems and computer science. **Decision Support Systems**, v.13, p. 373-380, 1995.

MAIER, R. **Knowledge management systems**: information and communication technologies for knowledge management. 3. ed. Berlin: Springer-Verlag, 2007. p.720.

MARCHAND, D. Managing information quality. In: WORMELL, I. INFORMATION QUALITY DEFINITIONS AND DIMENSIONS, 1989, Copenhagen. **Resumos...** Copenhagen: NORDINFO Seminar, Royal School of Librarianship, 1989. p.7.

MARCONDES, C. H. Representação e economia da informação. **Ciências da Informação**, Brasília, v. 30, n. 1, abr. 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S010019652001000100008&lng=pt&nrm=iso>. Acesso em: 29 jan. 2008. doi: 10.1590/S0100-19652001000100008.

MASLOW, A. A Theory of human motivation. **Psychological Review**, United States, v. 50, n. 4, p. 370-396, Jul., 1943. Disponível em: <<http://psycnet.apa.org/journals/rev/50/4/>> Acesso em: 30 jan. 2008.

MAYER, R. J.; MENZEL, C. P.; PAINTER, M. K.; DEWITTE, P. S.; BLINN, T.; PERAKATH, B. **Information Integration for Concurrent Engineering (IICE) IDEF3 Process description Capture Method Report - Interim Technical Report for period April 1992-September 1995**. Disponível em: <<http://www.idef.com/idef3.html>>. Acesso em: 22 jan. 2009.

McGILVRAY, D. **Executing Data Quality Projects**: Ten Steps to Quality Data and Trusted Information. San Francisco: Morgan Kaufmann, 2008. p. 352. ISBN: 0123743699.

McINERNEY, C. Knowledge management and the dynamic nature of knowledge. **Journal of the American Society for Information Science and Technology**, United States, v. 53, n. 12, p. 1009-1018, 2002. Disponível em: <<http://www.q3.ca/articles/archives/articles/KNOWLEDGE%20MANAGEMENT.pdf>> Acesso em: 30 jan. 2008.

MEIRINHOS, M. F. A. **Desenvolvimento profissional docente em ambientes colaborativos de aprendizagem a distância**: estudo de caso no âmbito da formação contínua. Braga, Portugal: Editora da Universidade do Minho, 2006. Originalmente apresentada como Tese de Doutorado –Instituto de Estudos da Criança, Universidade do Minho, Braga, 2006. p. 362. Disponível em: <<http://hdl.handle.net/1822/6219>>. Acesso em: 12 ago. 2007.

MILLER, F. J. $I = 0$ (Information has no intrinsic meaning). **Information Research: an international electronic journal**, United Kingdom, v. 8, n. 1, paper n. 140, Oct. 2002. Disponível em: <<http://InformationR.net/ir/8-1/paper140.html>> Acesso em: 30 jan. 2008.

MONCHY, F. **A Função manutenção: formação para a gerência da manutenção industrial**. São Paulo: Ebras/Durban, 1989. p.424.

MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. 5 ed. New York, NY: John Wiley & Sons, 2004. p. 776. ISBN: 0471656313.

MORIN, E. **Introdução ao pensamento complexo**. 2. ed. Lisboa : Instituto Piaget, 1990. p.177.

MOTTA, E. The knowledge modelling paradigm in knowledge engineering. **Handbook of Software Engineering and Knowledge Engineering**, Singapore, v. 1, 2001. Disponível em: <<ftp://cs.pitt.edu/chang/handbook/27.pdf>>. Acesso em: 12 jan. 2008.

MOUBRAY, J. **Reliability-centered Maintenance**. 2. ed. Oxford: Butterworth-Heinemann, Industrial Press Inc., 2001. p. 448. ISBN 0831131462.

NAUMANN, F.; ROLKER, C. Do metadata models meet IQ requirements? In: PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION QUALITY (IQ), Cambridge, MA. **Resumos...**International Conference on Information Quality, 1999. p. 99.

NAUMANN, F., ROLKER, C. Assessment methods for information quality criteria. In: INTERNATIONAL CONFERENCE ON INFORMATION QUALITY, 5th, 2000, Cambridge. **Resumos...** Cambridge, MA: MIT Press, 2000, p. 148–162, 2000. Disponível em: <http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG_Naumann/publications/IQ2000.pdf>. Acesso em: 12 fev. 2008.

NBR ISO/IEC 9126-1. **Engenharia de software - Qualidade de produto – Parte 1: Modelo de qualidade**. 1ª Ed. ABNT 2003. 21p.

NEHMY, R. M. Q., PAIM, I. A desconstrução do conceito de “qualidade da informação”. **Ciência da Informação**, v. 27, n. 1, p. 36-45, 1998.

NEWELL, A. The Knowledge level: presidential address. **Artificial Intelligence Magazine**, United States, v. 2, n. 2, a. 33, p. 1-20, 1981. Disponível em: <<http://www.aaai.org/AITopics/assets/PDF/AIMag02-02-001.pdf>>. Acesso em: 20 jan. 2008.

NEWMAN, J. Some observations on the semantics of “information”. **Information Systems Frontiers**, Netherlands, v. 3, n. 2, p. 155-167, Jun. 2001. Disponível em: <<http://www.springerlink.com/content/q01661p087r53734/>>. Acesso em: 20 jan. 2008. doi: 10.1023/A:1011439305567.

NONAKA, I.; TAKEUCHI, H. **Criação de conhecimento na empresa**: como as empresas japonesas geram a dinâmica da inovação. Tradução de Ana Beatriz Rodrigues e Priscila Martins Celeste. Rio de Janeiro: Campus, 1997. p.372.

O'BRIEN, D. **Defining and Implementing an Effective Data Quality Strategy**. United Kingdom: Ark Group, 2008. p. 106. ISBN: 978-1-906355-14-2.

O'CONNOR, P. D. T., NEWTON, D.; BROMLEY, R. **Practical Reliability Engineering**. 4 ed. New York, NY: John Wiley & Sons, 2002. p. 513. ISBN 0470844620.

OFFSHORE RELIABILITY DATA (OREDA). **History**. Oslo, 2009. Disponível em: <<http://www.sintef.no/static/tl/projects/oreda/history.htm>>. Acesso em: 07 jul. 2009.

OLAISEN, J. Information quality factors and the cognitive authority of electronic information. In: WORMELL, I. (ED.). **INFORMATION QUALITY DEFINITIONS AND DIMENSIONS**, Copenhagen. **Resumos...**NORDINFO Seminar, Royal School of Librarianship, 1989. p. 91.

OLSON, J. **Data Quality: The Accuracy Dimension**. 5. Ed. San Francisco: Morgan Kaufmann, 2003. p. 293.

PAUTKE, R.W.; REDMAN, T.C. Techniques to control and improve quality of data large databases. **Proc. of Statistics Canada Symp.** 90, Canada, 1990.

PECHT, M. Reliability engineering in the 21st century—a focus on predicting the reliability of electronic products and systems. In: 5th INTERNATIONAL CONFERENCE ON RELIABILITY, MAINTAINABILITY & SAFETY, 2001, Da Lian, China. **Resumos...** Da Lian: 5th International Conference On Reliability, Maintainability & Safety 2001, v. 1, 2001. p. 1–19.

PEIRCE, C. S. **Semiótica**. São Paulo: Perspectiva, 1977. p.352.

PERRY, W. L.; SIGNORI, D.; BOON, J. E. Jr. **Exploring Information Superiority: A Methodology for Measuring the Quality of Information and Its Impact on Shared Awareness**. Santa Monica, CA: RAND Corporation, 2004, p.175. ISBN: 0833034898.

PIATTINI, M.; CALERO, C.; GENERO, M. **Information and Database Quality (Advances in Database Systems)**. Boston: Kluwer Academic Publishers, 2002. p. 223.

POLANYI, M. **The tacit dimension**. Gloucester: Peter Smith, 1983. p.119.

REA, L. M.; PARKER, R. **A Metodologia De Pesquisa: do planejamento a execução**. São Paulo: Pioneira. 2002. p. 272.

REDMAN, T. C. **Data quality: management and technology**. New York, NY: Bell Laboratories, Murray Hill, Publisher Bantam Books Inc., 1992. p. 308. ISBN: 0-553-09149-2.

REDMAN, T. C. **Data quality for the information age**. Boston, MA: Artech House, 1996. p.303.

REDMAN, T. C. **Data Quality: The Field Guide**. Boston: Digital Press, 2001. p. 325.

REDMAN, T. C. **Data Driven: Profiting from Your Most Important Business Asset**. Harvard: Harvard Business School, 2008. p. 272. ISBN: 1422119122.

ROWLEY, J. The wisdom hierarchy: representations of the DIKW hierarchy. **Journal of Information Science**, United States, v. 33, n. 2, p. 163-18, Apr. 2007. Disponível em: <<http://jis.sagepub.com/cgi/content/abstract/33/2/163>>. Acesso em: 20 jan. 2008. doi: 10.1177/0165551506070706.

RUß-MOHL, S. **Qualitätssicherung im amerikanischen Journalismus: Modell für Europa?** Zürich, Osnabrück: Edition Interfrom, 1994. p. 374.

SALEH, J.H.; MARAIS, K. Highlights from the early (and pre-) history of reliability engineering. **Reliability Engineering and System Safety**, v.91, p. 249–256, 2006.

SAUSSURE, F. **Curso de Lingüística Geral**. 29ª ed. São Paulo: Cultrix, 2008. p. 279.

SCANNAPIECO M.; MISSIER, P.; BATINI, C. Data Quality at a Glance. **Datenbank-Spektrum**, v.14, p. 6-14, 2005.

SCHEER, A. W. **ARIS - Business Process Modeling**. 3ed. New York: Springer., p.238, 2000. ISBN:3540658351

SCHREIBER, G.; AKKERMANS, H.; ANJEWIERDEN, A.; HOOG, R. De; SHADBOLT, N.; VELDE, W. Van de; WIELINGA, B. Knowledge Engineering and Management: The CommonKADS Methodology. **MIT Press**, Cambridge, 2000.

SHANNON, C. E. A Mathematical theory of communication. **The Bell System Technical Journal**, United States, v. 27, p. 379–423; 623–656, Jul./Oct., 1948. Disponível em: <<http://plan9.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>>. Acesso em: 23 abr. 2004.

SHARMA, N. **The origin of the “Data Information Knowledge Wisdom” hierarchy**, 2008. Disponível em: <http://www-personal.si.umich.edu/~nsharma/dikw_origin.htm>. Acesso em: 23 dez. 2008.

SIRIHAL, A. B.; LOURENÇO, C. de A. Informação e conhecimento: aspectos filosóficos e informacionais. **Informação & Sociedade: estudos**, João Pessoa, v. 12, n. 1, p. 67-92, 2002. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/viewFile/154/148>>. Acesso em: 9 dez. 2007.

SPIEGLER, I. Knowledge management: a new idea or a recycled concept? **Communications of the AIS**, United States, v. 3 n. 4, article n. 2, Jun. 2000 Disponível em: <<http://portal.acm.org/citation.cfm?id=374521>>. Acesso em: 9 dez. 2007.

STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge engineering: Principles and methods. **Data & Knowledge Engineering**, p. 161-197, 1998. Disponível em: <<http://www.sciencedirect.com/science/article/B6TYX-3SYXJ6S-G/2/67ea511f5600d90a74999a9fef47ac98>>. Acesso em: 15 mar. 2008.

TAYLOR, R.S. **Value-Added Processes in Information Systems**, Norwood: Ablex, Series: Communication and Information Science. Westport: Greenwood Publishing Group, 1986. p.264.

TOZER, G. V. **Information Quality Management**. Oxford, Cambridge: NCC Blackwell, 1995. p.172.

TUOMI, I. Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory. In: ANNUAL HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 32., Maui, HI, USA, 1999b. **Resumos...** Annual Hawaii International Conference On System Sciences, 1999a. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=772795>. Acesso em: 12 nov. 2007. doi: 10.1109/HICSS.1999.772795.

TUOMI, I. Data is more than knowledge: implications of the reversed knowledge hierarchy for knowledge management and organizational memory. **Journal of Management Information Systems**, United States, v. 16 n. 3 p. 103-117, winter, 1999b. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.88.3389>>. Acesso em: 13 nov. 2007.

TURNER, N.; EVANS, D. **Data Quality? Don't Waste your Time**. Apresentado em: 12th International Conference on Information Quality, 2007.

TYSON, P.; MCPARTLAND, D. **Quality management considerations for implementing SGML**. In: XML Europe 2000.

UMAR, A.; KARABATIS, G.; NESS, L.; HOROWITZ, B.; ELMAGARDMID, A. Enterprise Data Quality: A Pragmatic Approach. **Information Systems Frontiers**, v.1, n.3, p.279-301, 1999.

VENZIN, M.; VON KROGH, G.; ROOS, J. Future research into knowledge management. In: VON KROGH, G.; ROOS, J.; KLEINE, D. (Comp.). **Knowing in firms: understanding, managing and measuring knowledge**. Thousand Oaks, Ca, USA: Sage Publications, 1999.

WALLACE, D. P. **Knowledge management: historical and cross-disciplinary Themes**. Westport: Libraries Unlimited, 2007. p.235.

WAND, Y.; WANG, R. Anchoring Data Quality Dimensions Ontological Foundations. **Communications of the Association of Computing Machinery**, v.39, n.11, p.234-256, 1996.

WANG, Y. R.; MADNICK, S. E. A Polygen Model for Heterogeneous Database Systems: The Source Tagging Perspective. In: 16TH VERY LARGE DATA BASES CONFERENCE, Brisbane, Australia. **Resumos...** Brisbane. 16th Very Large Data Bases Conference, 1990, p. 519.

WANG, R.Y.; KON, H.B. Towards total data quality management, *Information Technology Action: Trends and Perspectives*, ed. Englewood Cliffs, N.J.: Prentice Hall, 1993.

WANG, R.Y.; KON, H. B.; MADNICK, S. E. Data Quality Requirements Analysis and Modeling. **ICDE**, p. 670-677, 1993.

WANG, R.Y.; REDDY, M.P.; KON, H.B. Toward quality data: an attribute-based approach. **Dec Supp Sys**, v. 13, p. 349-372, 1995a.

WANG, R. Y.; STOREY, V. C.; FIRTH, C. P. A Framework for Analysis of Data Quality Research. *IEEE Trans. Knowledge of Data Engineering*, v.7, n.4, p. 623-640, 1995b.

WANG, R. W.; STRONG, D. M. Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, v.12, n.4, p.25-34, 1996.

WANG, R. A Product Perspective on Total Data Quality Management. *Communications of the ACM*, v.41, n.2, p. 58-65, 1998.

WANG, R.Y., STRONG, D. M., KAHN, B. K. LEE, Y. "**AIMQ: A Methodology for Information Quality Assessment**", (TDQM-98-01), Total Data Quality Management (TDQM) Research, MIT Sloan School of Management. 1998.

WANG, R. Y.; ZIAD, M.; LEE, Y. W. **Data Quality**. Kluwer, 2001.

WANG, R.Y.; PIERCE, D.; MADNICK, S.E.; FISHER, C. **Advances in Management Information Systems (AMIS): Information Quality** . USA: M.E. Sharpe, p. 265, 2005. ISBN-13: 978-0765611338.

WENGER, E. **Communities of practice: learning, meaning, and identity**. Cambridge University Press, 1999.

WILSON, T. D. The nonsense of 'knowledge management'. **Information Research: an International Electronic Journal**, United Kingdom, v. 8, n. 1, paper n. 144, Oct. 2002. Disponível em: <<http://informationr.net/ir/8-1/paper144.html>>. Acesso em: 30 jan. 2008.

WITTGENSTEIN, L. **Investigações filosóficas**. Tradução de Marcos G. Montagnoli. 2. ed. Petrópolis: Vozes, 1996.

ZACK, M. Developing a Knowledge Strategy. *California Management Review*; v.41, n.3, spring, 1999, p. 125- 145.

ZACK, M. Epilogue: Developing a Knowledge Strategy. In: CHOO, C. W.; BONTIS, N. **The Strategic Management of Intellectual Capital and Organizational Knowledge**. New York, NY, USA: Oxford University Press, Mar. 2002. p. 268-276. ISBN: 019515486X

ZAPPA, F. Packard gosse. In: **Joe's Garage: Act II&III**, 1979. (Álbum).

ZHANG, C.; YANG, Q.; LIU, B. Guest Editors' Introduction: Special Section on Intelligent Data Preparation. *IEEE Computer Society*; 2005, p. 1163-1165. ISSN: 1041-4347

ZELNY, M. Management support systems: towards integrated knowledge management. **Human Systems Management**, Netherlands, v. 6, n. 7, p. 59-70, 1987.

ZHU, X.; WU, X. Cost-Constrained Data Acquisition for Intelligent Data Preparation. **IEEE Transactions on Knowledge and Data Engineering**, v. 17, n. 11, p. 1542-1556, 2005.