

**UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E  
ENGENHARIA DE MATERIAIS**

MINERAÇÃO DE TEXTOS DE PATENTES PARA MONITORAMENTO  
TECNOLÓGICO, IDENTIFICAÇÃO DE APLICAÇÕES E SELEÇÃO DE  
MATERIAIS: ESTUDO DE CASO EM DESENVOLVIMENTOS DE AÇOS

Bráulio Salumão de Oliveira

São Carlos-SP  
2020



**UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E  
ENGENHARIA DE MATERIAIS**

MINERAÇÃO DE TEXTOS DE PATENTES PARA MONITORAMENTO  
TECNOLÓGICO, IDENTIFICAÇÃO DE APLICAÇÕES E SELEÇÃO DE  
MATERIAIS: ESTUDO DE CASO EM DESENVOLVIMENTOS DE AÇOS

Bráulio Salumão de Oliveira

Tese apresentada ao Programa de  
Pós-Graduação em Ciência e Engenharia  
de Materiais como requisito parcial à  
obtenção do título de DOUTOR EM  
CIÊNCIA E ENGENHARIA DE MATERIAIS

Orientador: Dr. Daniel Rodrigo Leiva.

Agência Financiadora: CAPES - Processo: 88882.332704/2019-01 (PROEX).

São Carlos-SP  
2020



## DEDICATÓRIA

A todos que dedicam seus esforços na busca de compreender o ainda incompreensível no universo da ciência e engenharia de materiais.

## VITAE

Mestre em Ciência e Engenharia de Materiais pela UFSCar (2016).

Bacharel em Engenharia de Materiais pela UFSCar (2010).





# UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia  
Programa de Pós-Graduação em Ciência e Engenharia de Materiais

---

## Folha de Aprovação

---

Defesa de Tese de Doutorado do candidato Bráulio Salumão de Oliveira, realizada em 26/11/2020.

### Comissão Julgadora:

Prof. Dr. Daniel Rodrigo Leiva (UFSCar)

Prof. Dr. Tomaz Toshimi Ishikawa (UFSCar)

Prof. Dr. Leandro Innocentini Lopes de Faria (UFSCar)

Prof. Dr. Maurício Dwek (ANSYS)

Prof. Dr. Yuri Walter (UFES)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência e Engenharia de Materiais.





## AGRADECIMENTOS

Agradeço a Deus e Nossa Senhora Aparecida.

Aos meus pais Francisco e Maria do Carmo, minhas irmãs Flávia e Daniela e aos sobrinhos e afilhados Maria Fernanda e Daniel, por todo o apoio e amor.

Ao orientador, Prof. Dr. Daniel Rodrigo Leiva por manter viva minha motivação para elaboração do trabalho, pelas contribuições e, sobretudo, por compreender meus esforços e minhas limitações.

Aos colegas do NIT/Materiais pelo companheirismo ao longo da jornada e contribuição ao conhecimento que tornou possível a realização deste trabalho.

Aos funcionários da secretaria do PPGCEM, Geraldo, Marcelo e Valéria, pela dedicação e presteza em todas as solicitações.

À CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo apoio financeiro para realização desse trabalho com bolsa de estudos, processo nº 88882.332704/2019-01, e pela disponibilização e manutenção do Portal de Periódicos que permitiu a obtenção dos dados da matéria prima deste trabalho, principalmente a partir da assinatura da *Derwent Innovations Index*.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.



## RESUMO

O crescente número de opções em materiais torna cada vez mais complexo os processos de seleção daquele mais adequado a exercer uma função em um produto, ou de escolher uma aplicação competitiva para um novo material desenvolvido. Recursos computacionais e bases de dados vêm sendo utilizados para minimizar incertezas, mas ainda carecem de alternativas, principalmente para avaliar oportunidades que possam ainda não ter se consolidado como opções comerciais. Os registros bibliográficos de documentos de patentes se apresentam como uma fonte de informação atualizada, abrangente e com detalhes técnicos que podem ser disponibilizados a partir de técnicas de bibliometria e mineração de textos. Esta tese objetivou elaborar uma planilha de dados a partir da mineração de textos em registros de patentes de aços indexados na base *Derwent Innovations Index*. Os métodos empregados permitiram recuperar dados, a partir de segmentos de textos, de 85% entre 22.911 registros no período entre os anos de 2009 e 2018. Foi possível avaliar os desenvolvimentos recentes de aços a partir da composição química, processos, microestrutura, propriedades e aplicações. Os resultados mostram um avanço no detalhamento de indicadores para atividades de monitoramento tecnológico. No âmbito da seleção de materiais, foi possível elaborar diagramas de propriedades de materiais ao longo do tempo e com outras informações extraídas dos textos dos registros. Os procedimentos utilizados podem ser aperfeiçoados no futuro para um escopo temporal maior, avaliação de outras propriedades ou sua extensão para diferentes classes de materiais.

**Palavras-chave:** Seleção de Materiais; Indicadores tecnológicos; Patentes; Aços



## ABSTRACT

### TEXT MINING IN PATENTS FOR TECHNOLOGICAL MONITORING, APPLICATIONS IDENTIFICATION, AND MATERIALS SELECTION: A CASE STUDY IN STEELS DEVELOPMENTS

The growing number of options in materials leads to increasing complexity of the process to select the most suitable to perform a function in a product or to choose a competitive application for a newly developed material. Computational resources and databases are used to minimize uncertainties, but they still lack alternatives to evaluate opportunities that may not present themselves as commercial options. The bibliographic records of patent documents are presented as comprehensive and updated information sources, with technical details that can be retrieved from bibliometrics and text mining techniques. This thesis aimed to develop a data sheet based on text mining in steels patent records indexed in the Derwent Innovations Index. The methods employed made it possible to retrieve 85% of text segments among 22,911 records between 2009 and 2018. It was possible to evaluate recent steel developments based on chemical composition, processes, microstructure, properties, and applications. The results show the progress in the development of more detailed indicators for technological monitoring activities. In the scope of materials selection, it was possible to create diagrams of materials properties over time and other information retrieved. The procedures used in this study can be improved for a longer time scope, other properties analysis, or application to different classes of materials.

**Keywords:** Materials selection; Technological Indicators; Patents; Steels



## PUBLICAÇÕES

### Em periódicos

- SANTOS, D.P.; PELISSARI, P.I.B.G.B.; OLIVEIRA, B.S.; LEIVA, D.R.; MELLO, R.F.; PANDOLFELLI, V.C. Materials selection of furnace linings with multi-component refractory ceramics based on an evolutionary screening procedure. **Ceramics International**, v. 46, p. 4113-4125, 2020.
  
- OLIVEIRA, B.S.; MILANEZ, D.H.; LEIVA, D.R.; FARIA, L.I.L.; BOTTA, W.J.; KIMINAMI, C.S. Thermal spraying processes and amorphous alloys: Macro-indicators of patent activity. **Materials Research-Ibero-american Journal of Materials**, v. 20, p. 89-95, 2017.
  
- MILANEZ, D.H.; OLIVEIRA, B.S.; NOYONS, E.C.M; FARIA, L.I.L; BOTTA, W.J. Assessing collaboration and knowledge flow on coatings of metallic glasses obtained from thermal spraying processes using bibliometrics and science mapping. **Materials Research-Ibero-american Journal of Materials**, v. 20, p. 71-80, 2017.

### Em anais de eventos

- SILVA, R.F.; OLIVEIRA, B.S.; LEIVA, D.R. Indicadores De Patentes Em Aços De Baixa Densidade E Aços De Alto Módulo Elástico. *In: Encontro Nacional dos Estudantes de Engenharias Metalúrgica, de Materiais e Mineração*. 19. 2019. São Paulo. **Anais [...]**. Associação Brasileira de Metalurgia, Materiais e Mineração, 2019. p. 98–108.
  
- OLIVEIRA, B.S.; MILANEZ, D.H.; LEIVA, D.R. Indicadores sobre desenvolvimentos recentes em aços a partir da análise de documentos de patente. *In: Congresso Anual da ABM, 73.*, 2018, São Paulo. **Anais [...]**. Associação Brasileira de Metalurgia Materiais e Mineração, 2018. p. 1607-1616.

- ANDRADE, M. A.; OLIVEIRA, B.S.; MILANEZ, D. H.; LEIVA, D. R.; ZEPON, G. *In: Materials Science & Technology*. 2018. **Anais [...]**. Columbus, OH, Estados Unidos da América, 2018. p. 341-347.

- PEREIRA, B. C. F.; OLIVEIRA, B. S.; LEIVA, D. R. Análise bibliométrica da produção científica em métodos de tomada de decisão para seleção de materiais. *In: Congresso Brasileiro de Engenharia e Ciências dos Materiais*, 23., 2018, Natal-RN. **Anais [...]**. Metallum, 2018. p. 9144-9155.

- SANTOS, C. V.; OLIVEIRA, B. S.; MILANEZ, D. H.; LEIVA, D. R.; FARIA, L. I. L. Contribuições da bibliometria para a seleção de materiais no desenvolvimento de produtos. *In: Encontro Brasileiro de Bibliometria e Cientometria: A Ciência em Rede*, 6. 2018, Rio de Janeiro. **Anais [...]**. Rio de Janeiro, 2018. p. 453-463.

- OLIVEIRA, B.S.; MILANEZ, D.H.; LEIVA, D.R.; FILHO, W.J.; KIMINAMI, C.S. Indicadores tecnológicos sobre recobrimentos para tubos metálicos utilizados na indústria petroquímica. *In: Congresso Anual da ABM*, 72., 2017, São Paulo. **Anais [...]**. São Paulo, 2017. p.2044-2053.

- SANTOS, C.V.; OLIVEIRA, B.S.; LEIVAD.R.; FARIA, L.I.L; GREGOLIN, J.A.R. Seleção de materiais para fabricação de componentes aeronáuticos em compósitos integrando aspectos ambientais e econômicos. *In: Congresso Anual da ABM*, 72. 2017, São Paulo. **Anais [...]**. São Paulo, 2017. p. 2438-2449.



## ÍNDICE DE ASSUNTOS

	Pág.
FOLHA DE APROVAÇÃO.....	i
AGRADECIMENTOS .....	iii
RESUMO.....	v
ABSTRACT .....	vii
PUBLICAÇÕES.....	ix
ÍNDICE DE ASSUNTOS .....	xi
ÍNDICE DE TABELAS .....	xiii
ÍNDICE DE QUADROS .....	xv
ÍNDICE DE FIGURAS .....	xvii
SÍMBOLOS E ABREVIATURAS.....	xxi
1 INTRODUÇÃO.....	1
1.1 Objetivos.....	4
2 REVISÃO DA LITERATURA.....	5
2.1 Seleção de Materiais e Seleção de Aplicações .....	5
2.1.1 Fontes de informação para atividades de seleção de materiais.....	13
2.2 Informação Tecnológica e Engenharia de Materiais .....	16
2.2.1 Mineração de textos em patentes para elaboração de indicadores .....	20
2.3 Desenvolvimentos recentes em aços.....	27
3 MATERIAIS E MÉTODOS .....	35
3.1 Recuperação e preparação da amostra de dados .....	35
3.2 Desenvolvimento de filtros em software bibliométrico.....	38
3.2.1 Composição química.....	44
3.2.2 Propriedades físicas e mecânicas.....	51
3.2.3 Microestrutura, processos e aplicações .....	54
3.3 Organização da planilha de dados e elaboração dos indicadores .....	56
4 RESULTADOS E DISCUSSÃO .....	61
4.1 Avaliação dos dados recuperados e indicadores gerais do patenteamento recente em aços.....	61
4.2 Avanços no detalhamento de indicadores para atividades de monitoramento tecnológico .....	67

4.3 Contribuições para a Seleção de Materiais.....	73
5 CONCLUSÕES.....	81
6 SUGESTÕES PARA FUTUROS TRABALHOS .....	85
7 REFERÊNCIAS BIBLIOGRÁFICAS.....	87
APÊNDICE A Lista de marcadores e tesouros de composição química .....	97
APÊNDICE B Lista de marcadores e tesouros de processos, microestruturas, propriedades e aplicações.....	103
APÊNDICE C Instruções detalhadas de atualização dos filtros e da planilha de dados .....	111
APÊNDICE D Instruções detalhadas de elaboração de indicadores.....	123
ANEXO A Códigos da Classificação Internacional de Patentes referentes a aços .....	129

## ÍNDICE DE TABELAS

	Pág.
Tabela 2.1 – Exemplos de bases de dados de materiais gratuitas e comerciais disponíveis para consulta.....	14
Tabela 2.2 – Exemplo de expressão de busca desenvolvida para recuperação de registros de patentes em aços de baixa densidade. ....	26
Tabela 3.1 – Procedimento de limpeza de registros para a definição da amostra de registros de patentes em aços entre os anos de 2009 e 2018.....	37
Tabela 3.2 – Exemplo de substituição de termos por marcadores em um texto do campo “Novidade” de um registro de patente. ....	41
Tabela 3.3 – Sequência geral de tratamentos de textos do campo “Novidade” para a recuperação do teor de manganês (à esquerda) em um registro de documento de patente.....	46
Tabela 3.4 – Sequência geral de tratamentos de textos do campo “Novidade” para a recuperação do teor de manganês (à direita) em um registro de documento de patente.....	48
Tabela 3.5 – Expressões utilizadas como marcadores e unidades de propriedades físicas e mecânicas. ....	52
Tabela 3.6 – Sequência geral de tratamentos de textos do campo “Vantagem” para a recuperação do Limite de escoamento (à direita) em um registro de documento de patente.....	53
Tabela 3.7 – Sequência geral de tratamentos de textos do campo “Uso” para a recuperação de possíveis aplicações em um registro de documento de patente. ....	55
Tabela 3.8 – Código de cores atribuído aos registros após verificação manual dos dados recuperados.....	58
Tabela 4.1 – Dados dos elementos de liga recuperados na planilha. ....	62
Tabela 4.2 – Dados das propriedades recuperadas na planilha. ....	63
Tabela 4.3 – Principais processos recuperados no desenvolvimento dos filtros e apresentados na planilha de dados.....	64
Tabela 4.4 – Principais microestruturas apresentadas na planilha de dados...	65

Tabela 4.5 – Principais aplicações para aços apresentadas na planilha de dados. .....	65
Tabela 4.6 – Comparação do número de registros em aços de baixa densidade. .....	68

## ÍNDICE DE QUADROS

Quadro A.1 – Marcadores utilizados para a seleção de texto nos filtros de recuperação das informações de teores de elementos químicos em aços nos registros de patentes.....	97
Quadro A.2 – Marcadores utilizados para os elementos de ligas em aços utilizados na elaboração da planilha. ....	97
Quadro A.3 – Expressões utilizadas para eliminação prévia à atribuição de marcadores. ....	97
Quadro A.4 – Marcadores utilizados para limites superiores e inferiores dos elementos de liga e propriedades. ....	102
Quadro B.1 – Expressões utilizadas para eliminação prévia à atribuição de marcadores das propriedades.....	103
Quadro B.2 – Substituições de denominações da norma API 5L para valores de Limite de Escoamento e Limite de Resistência à Tração.....	104
Quadro B.3 – Expressões utilizadas no tesouro de termos de processos. ....	105
Quadro B.4 – Códigos da CIP utilizados no tesouro de processos.....	106
Quadro B.5 – Expressões utilizadas no tesouro de termos de aplicações.....	107
Quadro B.6 – Códigos da CIP utilizados no tesouro de aplicações. ....	108
Quadro B.7 – Expressões utilizadas no tesouro de termos de microestrutura. ....	109



## ÍNDICE DE FIGURAS

	Pág.
Figura 2.1 – Estágios e atividades gerais de um processo de seleção de materiais.....	6
Figura 2.2 – Exemplo de diagrama de Ashby para seleção de materiais a partir de valores de resistência mecânica e densidade, e o conceito de índice de mérito para comparação. ....	7
Figura 2.3 – Abordagens com objetivo de encontrar aplicações para um novo material.....	9
Figura 2.4 – Rede de co-ocorrência de palavras-chave dos autores de artigos de seleção de materiais e aços em função do índice médio de citações. ....	11
Figura 2.5 – Representação percentual dos principais elementos de liga abordados em patentes de aços entre 2007 e 2016 a partir de códigos da CIP. ....	19
Figura 2.6 – Total de documentos de patente de diferentes classes de nanocelulose a partir de diferentes fontes de informação. ....	22
Figura 2.7 – Exemplo de um trecho de registro de documento de patente na base DII com destaques para as informações utilizadas para elaboração da base de dados. ....	24
Figura 2.8 – Trecho de documento original da patente com o registro bibliográfico apresentado na Figura 2.7. ....	25
Figura 2.9 – Evolução temporal do depósito de patentes e da produção mundial de aço bruto entre 2007 e 2016. ....	28
Figura 2.10 – Diagrama resistência em função do alongamento de aços utilizados na indústria automobilística. ....	30
Figura 2.11 – Diagrama resistência em função do alongamento de ligas Fe-C-Mn-Al.....	31
Figura 2.12 – Número de registros de patentes em aços para dutos de transporte de hidrocarbonetos classificados em códigos da CIP com percentual em peso acima de 1,5% de manganês. ....	33
Figura 3.1 – Fluxograma geral de elaboração de filtros utilizado na pesquisa. ....	40

Figura 3.2 – Fluxograma simplificado de etapas para a recuperação e padronização dos teores de alumínio em um registro de patente de aço. ....	43
Figura 4.1 – Porcentagem de registros em que foi possível recuperar informações sobre a composição química dos elementos avaliados nos registros de patentes em aços com prioridade entre 2009 e 2018 e indexados na <i>Derwent Innovations Index</i> . ....	61
Figura 4.2 – Diagrama de caixas dos valores máximos de carbono citados nos registros de patentes em aço entre 2009 e 2018. ....	69
Figura 4.3 – Número de registros com informação de depósitos de patentes com manganês acima de 1,5% em peso, obtidos por códigos da CIP e pela planilha de dados. ....	70
Figura 4.4 – Número de registros de patentes de aços para dutos de transporte de hidrocarbonetos entre 2009 e 2014 e média dos teores máximos de manganês. ....	71
Figura 4.5 – Relação entre a densidade de aços e o percentual médio em peso de alumínio e manganês a partir de dados apresentados na planilha. ....	72
Figura 4.6 – Relação entre limite de resistência à tração e dureza a partir de dados apresentados na planilha. ....	74
Figura 4.7 – Relação entre limite de resistência à tração e alongamento a partir de dados apresentados na planilha. ....	75
Figura 4.8 – Relação entre limite de resistência à tração e alongamento, com detalhes de microestruturas, a partir de dados apresentados na planilha. ....	76
Figura 4.9 – Relação entre limite de escoamento e alongamento em registros com aplicação requisitada nos mercados de automóveis, dutos e construção civil. ....	77
Figura 4.10 – Relação entre limite de escoamento e alongamento em registros de documentos de patentes e aços comerciais. ....	78
Figura C.1 – Tela de escolha de filtros no software VantagePoint e de seus campos para a elaboração e atualização da planilha de dados. ....	112
Figura C.2 – Exemplo de lista gerada no VantagePoint para os teores de carbono no campo novidade, com sua seleção total e comando de copiar. ....	113



Figura C.3 – Exemplo de lista após copiada e colada para o MSEXcel com destaque para as colunas a serem excluídas. ....	115
Figura C.4 – Ilustração de procedimentos no menu “Texto para colunas”. ....	115
Figura C.5 – Disposição das colunas na planilha de apoio após comando “Texto para colunas” para dados numéricos. ....	116
Figura C.6 – Exemplo de verificação manual de valores apresentados além da coluna “D”. ....	117
Figura C.7 – Exemplo de funcionamento da planilha de apoio Ajustador de Limites. ....	118
Figura C.8 – Disposição da planilha após o comando “Texto para colunas” para dados de caracteres alfabéticos. ....	119
Figura C.9 – Planilha de apoio para compilar os dados recuperados de teor de carbono dos três diferentes campos. ....	120
Figura C.10 – Planilha de apoio para compilar os dados recuperados de teor de carbono de três diferentes campos. ....	121
Figura C.11 – Detalhes da organização da planilha principal “NIT SPD” e de suas fórmulas “PROCV”. ....	121
Figura D.1 – Exemplo de apresentação de filtros nas colunas do MSEXcel para operação da planilha de dados. ....	124
Figura D.2 – Etapas para criar uma cópia da planilha principal para elaboração de gráficos e tabelas. ....	125
Figura D.3 – Etapas para exclusão de linhas que não apresentam valores... ..	126
Figura D.4 – Exemplo de tabela dinâmica a partir de dados de carbono (média dos valores máximos), processos e aplicações selecionadas. ....	127
Figura D.5 – Exemplo de busca avançada por códigos GA na <i>Derwent Innovations Index</i> .....	128



**SÍMBOLOS E ABREVIATURAS**

<b>AHP</b>	<i>Analytic Hierarchy Process</i>
<b>API</b>	<i>American Petroleum Institute</i>
<b>ASTM</b>	<i>American Society for Testing and Materials</i>
<b>CAPES</b>	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
<b>CIP</b>	Classificação Internacional de Patentes
<b>EPO</b>	Escritório Europeu de Patentes
<b>EPS</b>	<i>Earliest Priority Selector</i>
<b>Ksi</b>	<i>Pounds per square inch</i>
<b>LE</b>	Limite de escoamento
<b>LRT</b>	Limite de resistência à tração
<b>MADM</b>	<i>Multiple Attribute Decision Making</i>
<b>MCDM</b>	<i>Multi Criteria Decision Making</i>
<b>MS</b>	Microsoft®
<b>INPI</b>	Instituto Nacional da Propriedade Industrial
<b>MPa</b>	Mega Pascal
<b>NIT</b>	Núcleo de Informação Tecnológica em Materiais
<b>TOPSIS</b>	<i>Technique for Order of Preference by Similarity to Ideal Solution</i>
<b>UNS</b>	<i>Unified numbering system</i>
<b>VIKOR</b>	<i>VlseKriterijumska Optimizacija I Kompromisno Resenje</i>
<b>%p.</b>	Porcentagem em peso



## 1 INTRODUÇÃO

Com o inestimável número de possibilidades em termos de composição química e melhorias nas rotas de processamento, os materiais de engenharia atingem uma vasta gama de propriedades, o que torna cada vez mais complexo o processo de escolha daquele mais adequado a uma aplicação. Embora bases de dados disponibilizem informações de até 450.000 opções comerciais [1], certamente o número real de possibilidades em materiais é maior e tem ascendência diária, resultado de atividades em pesquisa e desenvolvimento por pesquisadores em universidades, centros de pesquisas e empresas. Diferentes procedimentos e algoritmos visam diminuir o risco na tomada de decisão na seleção de materiais, de acordo com o projeto mecânico e desempenho funcional esperado [2–4]. Entretanto, o volume de opções pode gerar ainda mais incertezas, principalmente em etapas iniciais do desenvolvimento de produtos [5,6].

Por outro lado, as pesquisas em materiais não são somente motivadas para a solução de um problema de engenharia específico, ou seus resultados destoam do desempenho teórico esperado. Dessa maneira, encontrar aplicações para novos materiais também é um desafio para cientistas e engenheiros. As propriedades estudadas, após confirmadas por caracterizações e testes, podem se apresentar como alternativas a serem consideradas para a solução de novos ou recorrentes desafios técnicos. Uma outra discussão, ainda presente por pesquisadores em materiais, trata de alternativas para avaliar soluções similares já desenvolvidas, principalmente para evitar esforços repetitivos e para buscar aplicações com proteção da propriedade industrial ainda não requerida, principalmente por meio de patentes.

Os documentos de patentes são considerados uma importante fonte de informação técnica pois contêm detalhes sobre as invenções que se pretende proteger como um requisito necessário à sua concessão. No caso de novos materiais, por exemplo, apresentam detalhes sobre a composição química, parâmetros de processamento e, em alguns casos, sobre suas propriedades resultantes que justificam o uso a ser protegido. Dessa maneira, possuem um grande potencial para atividades de seleção de materiais, caso essas

informações possam ser disponibilizadas e, depois, aproveitadas para aplicação em métodos de tomada de decisão já utilizados para tal fim. Além disso, trata-se de documentos em domínio público e com informações disponíveis em bases de textos completos ou de registros bibliográficos.

Um desafio para a aplicação dos métodos de seleção de materiais é encontrar informações detalhadas sobre possíveis candidatos, com detalhamento até o nível de composição química, condições de processamento e propriedades dos materiais. Geralmente, essas informações não estão concentradas em uma única base de dados [7] ou as opções apresentadas são de custo elevado para acesso de acordo com a situação econômica do país ou instituição. Outros desafios tratam da quantidade, qualidade e da atualização dessas informações nessas bases ou softwares, e o histórico de aplicação dos materiais candidatos para uma avaliação qualitativa das opções candidatas. Uma oportunidade de uso das informações contidas em documentos de patentes para a seleção de materiais está em encontrar e recuperar os dados de interesse e relacioná-los e, assim, buscar oportunidades de substituição em produtos atualmente manufaturados com outros materiais. Além disso, para pesquisadores acadêmicos que buscam encontrar possíveis aplicações de novos materiais, avaliar opções semelhantes já patenteadas pode indicar oportunidades de novas pesquisas, estímulo à proteção intelectual do conhecimento adquirido, criação de novas empresas de base tecnológica ou parcerias de licenciamento comercial com a iniciativa privada.

Estudos que utilizam informações de documentos de patentes para atividades de ciência e engenharia de materiais se baseiam em códigos da Classificação Internacional de Patentes (CIP) e no processamento de linguagem natural em informações contidas em registros de bases referenciais. Avanços recentes na mineração de texto em documentos completos [8,9] e em registros bibliográficos apresentam potencial para alimentar algoritmos para tomada de decisão em seleção de materiais com informações de materiais ainda não necessariamente produzidos em escala comercial. Um segundo benefício potencial para o uso de informações mais detalhadas recuperadas de documentos de patentes se encontra em atividades de monitoramento

tecnológico. O uso de dicionários para agrupamento e classificação de palavras e códigos da CIP têm permitido identificar tendências de desenvolvimento em classes específicas de materiais. Entretanto, novas técnicas de mineração de textos podem refinar esses indicadores com a apresentação de um maior nível de detalhes para análise de cenários tecnológicos [8]. Como exemplo, além de evidenciar os principais atores no desenvolvimento de materiais, sua ação geográfica e focos de suas invenções, é também interessante demonstrar como evoluem as propriedades dessas invenções ao longo do tempo, de que maneira estão sendo alcançadas, e para quais aplicações são requeridas sua proteção por meio desses pedidos de patentes.

Para tanto, a organização de um grande volume de informação em uma mesma base de dados que permita comparações e análise dessas invenções é desafiadora. Dessa maneira, o desenvolvimento de uma ferramenta única que englobasse dados do patenteamento de todos os inventos relacionados a materiais parece inviável. Por outro lado, após a elaboração de um método que atenda uma classe de materiais específica, a adaptação para outras poderia ser facilitada.

Embora consista como um dos materiais mais utilizados em engenharia durante décadas, ainda existem oportunidades para o desenvolvimento de tecnologias que envolvam os aços. Uma estimativa inicial aponta para uma média de 3 depósitos de patentes por dia, nos últimos 10 anos, de invenções relacionadas a novas ligas, condições de processamento, melhorias de propriedades e oportunidades de aplicações desses metais ferrosos [10]. Adicionalmente, trata-se de uma classe com uma definição simples e bem estabelecida de ligas Ferro com adição de até 2,14% de carbono, embora seus desenvolvimentos possam envolver tanto o alto teor de outros elementos quanto a microadição de outros, de acordo com as propriedades desejadas e condições de manufatura. Adicionalmente, os processos de produção de aços são bem conhecidos e explorados na literatura, bem como suas tradicionais aplicações. Assim, os aços constituem uma promissora classe de materiais em um estudo de caso para o desenvolvimento de um método para mineração de textos em registros de documentos de patentes, que permitirá a avaliação de sua eficácia.

A partir desta primeira experiência, será possível avaliar oportunidades para a expansão para outras ligas metálicas e classes de materiais.

## 1.1 Objetivos

Baseado nos desafios atuais referentes ao uso de informações contidas em documentos de patentes para estudos de seleção de materiais e monitoramento tecnológico, esta tese visa contribuir com avanços na mineração de textos dos registros bibliográficos destes documentos, focados em recuperar informações sobre composição química, microestrutura, processos de produção, propriedades atingidas e aplicações requeridas. Assim, os pesquisadores acadêmicos e industriais da área de materiais terão à disposição uma nova ferramenta para avaliar o potencial de aplicação dos materiais resultantes de suas pesquisas e compará-los com os que já requisitaram a proteção da propriedade intelectual por patentes.

Logo, o objetivo geral desta tese é extrair informações de materiais em registros bibliográficos de patentes e compilar os segmentos de textos em uma única planilha, que permita tanto a apresentação dos dados recuperados de uma única patente quanto a comparação entre dados de diferentes documentos.

Para atingir este objetivo, foram definidos as seguintes etapas e objetivos específicos:

a) Comparar como as informações de composição química, microestrutura, processamento, propriedades e aplicações de materiais são apresentadas em documentos de patentes e em seus registros bibliográficos.

b) Elaborar procedimentos em software bibliométrico para recuperar os segmentos de textos de interesse.

c) Criar mecanismos para organizar as informações recuperadas em uma única planilha de dados.

d) Avaliar as contribuições para avanços nos estudos de monitoramento tecnológico em materiais.

e) Avaliar o potencial de uso das informações recuperadas em estratégias de identificação de aplicações e seleção de materiais.



## 2 REVISÃO DA LITERATURA

### 2.1 Seleção de Materiais e Seleção de Aplicações

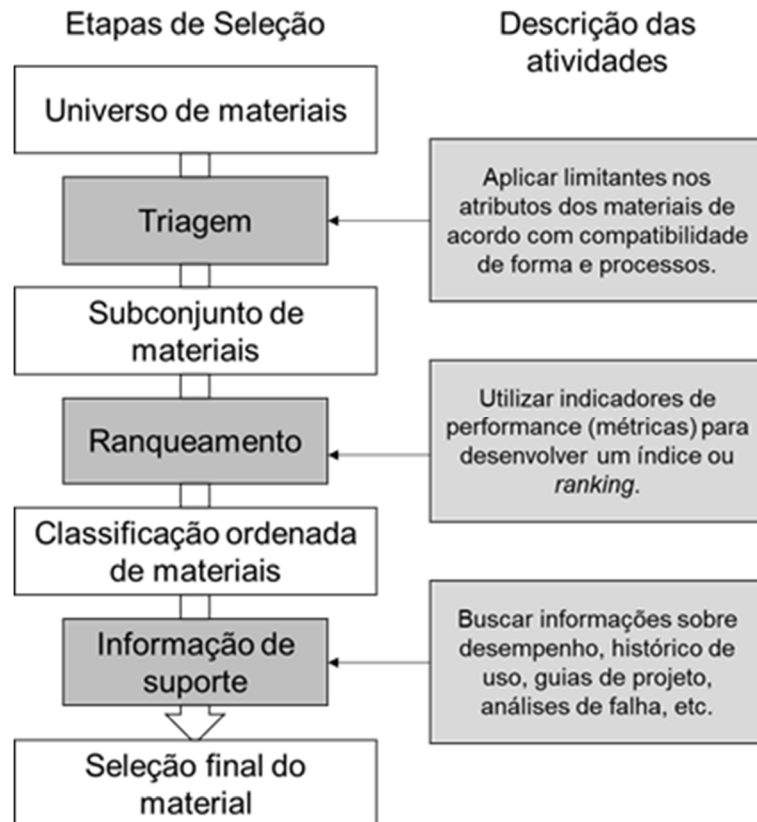
A Seleção de Materiais consiste no processo de escolha de um conjunto de opções de materiais mais adequados de acordo com o projeto mecânico e outras características funcionais necessárias ao desempenho esperado de um produto ou de alguma de suas partes [11,12]. A abordagem aos problemas de seleção pode ser dividida a partir de três categorias principais:

1 - Substituição de materiais a partir de um produto com restrições estabelecidas em termos de forma, sem a possibilidade de alterações de desenho de projeto. Assim, a escolha de um material substituto é a única variável livre e limitada àqueles com um desempenho bastante similar, em que a vantagem da nova opção seja independente da forma, como um menor custo de aquisição ou processamento.

2 - Melhoria de desempenho em um projeto com algumas variáveis fixas e outras restritas com pouca liberdade de alteração, como por exemplo a espessura de uma placa. Dessa maneira, o processo de escolha de um novo material pode levar a alterações no desenho, em troca da otimização de desempenho, o que torna a escolha dos materiais menos restrita.

3 - Novo produto, em que as variáveis de projeto, desenho e opções de materiais são livres. Por consequência, os projetistas podem optar por qualquer opção a partir do universo de materiais e processos disponível.

O caminho para a decisão final da escolha do material pode ser realizado a partir de diversas estratégias e algoritmos, desenvolvido por diversos autores de acordo com o tipo de problema e seus desafios [13–17]. Os mais utilizados possuem uma lógica em comum e suas etapas principais são sintetizadas na Figura 2.1. De maneira geral, parte-se de todo um universo de opções de materiais disponível, aplica-se restrições para a eliminação de alternativas não viáveis, estabelece-se critérios para avaliações dos candidatos restantes e pesos para o seu ranqueamento. A seguir, os candidatos remanescentes são submetidos a um processo de documentação, de maneira a acrescentar uma avaliação qualitativa ao que os métodos de ranqueamento apresentaram e, assim, minimizar incertezas na escolha.



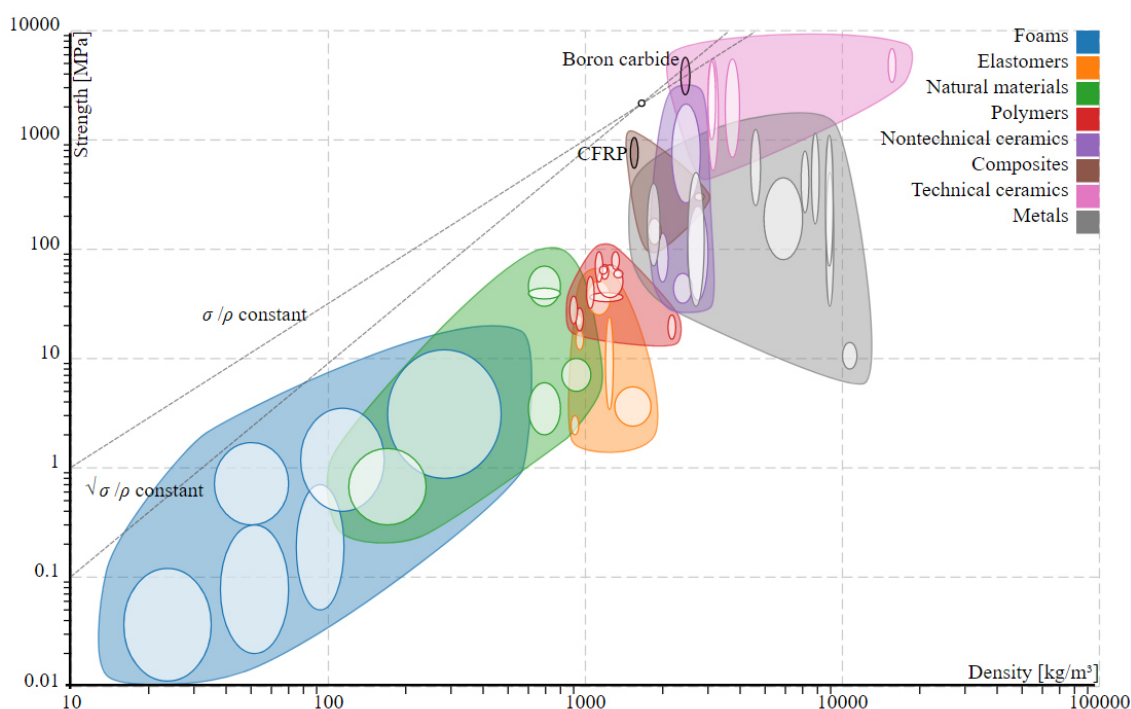
Fonte: Adaptado de [2].

Figura 2.1 – Estágios e atividades gerais de um processo de seleção de materiais.

A etapa de triagem parte dos requisitos do projeto mecânico que podem ser desde as mais restritas por documentos específicos como normas e regulamentos, até as mais livres no caso de alternativas para otimização de desempenho ou de novos produtos. O método mais consagrado para essa triagem inicial é proposto a partir dos diagramas desenvolvidos por Ashby [12] que relaciona duas propriedades de interesse e mapeia o universo de materiais entre elas. Um exemplo desses diagramas é apresentado na Figura 2.2. Diversos diagramas podem ser encontrados em livros, artigos e softwares específicos para tal análise.

De acordo com a solicitação mecânica e parâmetros geométricos do produto é possível estabelecer uma função chamada índice de mérito, que corresponde a uma relação entre as duas propriedades alvo com o objetivo de ser maximizada. No exemplo da Figura 2.2, duas linhas são apresentadas, e sua

escolha depende da geometria e da solicitação mecânica esperada para o produto ou componente. Todas as alternativas que se encontram na mesma linha de índice de mérito possuem o mesmo desempenho, o que auxilia os processos de comparação e triagem.



*Strength*: Resistência; *Density*: Densidade; *Boron carbide*: Carbetto de boro; *CFRP*: Polímero de fibra de carbono reforçado; *Foams*: Espumas; *Elastomers*: Elastômeros; *Polymers*: Polímeros; *Nontechnical ceramics*: cerâmicas não-técnicas; *Composites*: Compósitos. *Technical Ceramics*: Cerâmicas técnicas; *Metals*: Metais.

Fonte: [18].

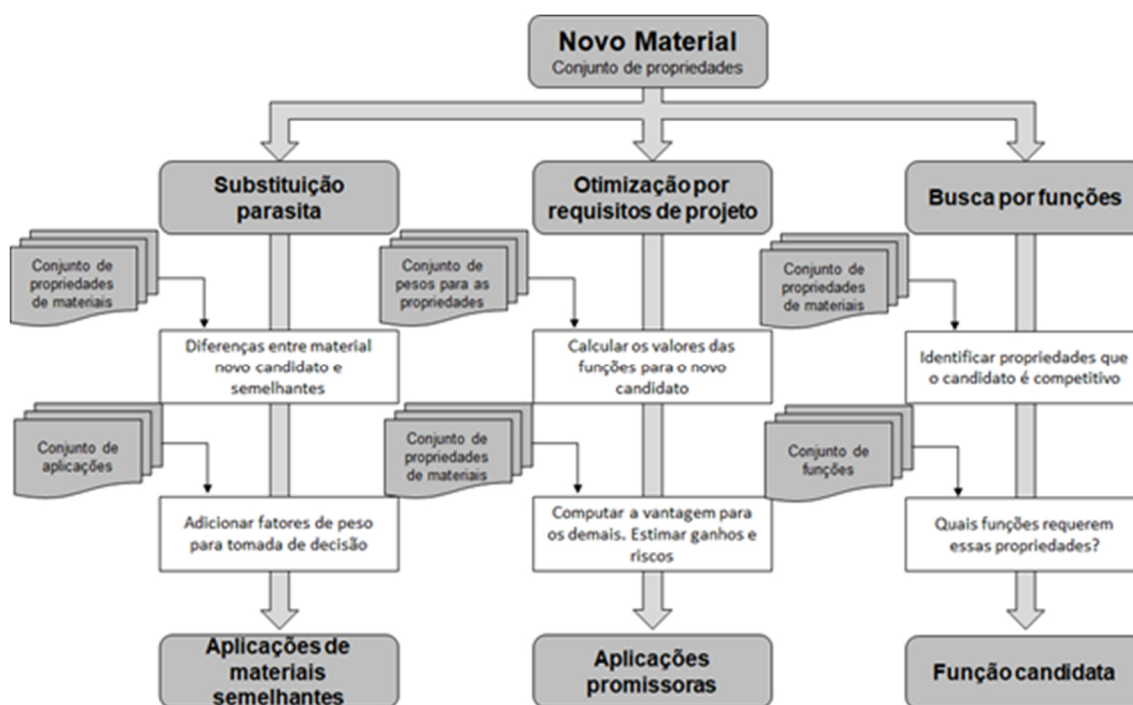
Figura 2.2 – Exemplo de diagrama de Ashby para seleção de materiais a partir de valores de resistência mecânica e densidade, e o conceito de índice de mérito para comparação.

A etapa de ranqueamento é facilitada pela aplicação de métodos de tomada de decisão multicritério ou multiatributo (MCDM e MADM nas siglas em inglês) [4,19]. Basicamente, trata-se de algoritmos que relacionam os critérios para tomada de decisão (propriedades), os materiais (alternativas) e seu desempenho (variáveis) a partir de matrizes. A seguir, são estabelecidos pesos para cada critério e as alternativas são avaliadas pelo seu desempenho em cada uma das propriedades. De acordo com a ponderação estabelecida pelo

algoritmo, cada material terá um desempenho final avaliado, que será finalmente classificado do mais adequado para o de menor chance de sucesso.

A informação de suporte pode ser obtida a partir de diferentes fontes, de maneira a confirmar se os materiais que apresentaram as melhores classificações são realmente os mais indicados para desempenhar as funções requeridas. Itens como disponibilidade, histórico de uso e a própria ausência de dados concretos para análise podem levar os projetistas a desconsiderar uma ou mais alternativas, estabelecer novos critérios para a tomada de decisão ou alterar os pesos atribuídos nos algoritmos de auxílio ao ranqueamento. Assim, a Figura 2.1 não representa um processo necessariamente linear, mas também o caráter iterativo dos processos de projeto de produtos, em que a experiência dos projetistas é também considerada na tomada de decisão.

Embora muitos métodos para seleção de materiais estejam bem desenvolvidos e sejam aplicados em atividades acadêmicas e industriais, a seleção de aplicação para um dado material, geralmente novo, trata-se de um problema mais desafiador pois consiste em partir de uma única alternativa e compará-la com um universo de outros materiais, funções desempenhadas e aplicações já estabelecidas [20–23]. Três abordagens principais são propostas para encontrar uma aplicação para um novo material, apresentadas de maneira sintetizada na Figura 2.3.



Fonte: Adaptado de [20,21].

Figura 2.3 – Abordagens com objetivo de encontrar aplicações para um novo material.

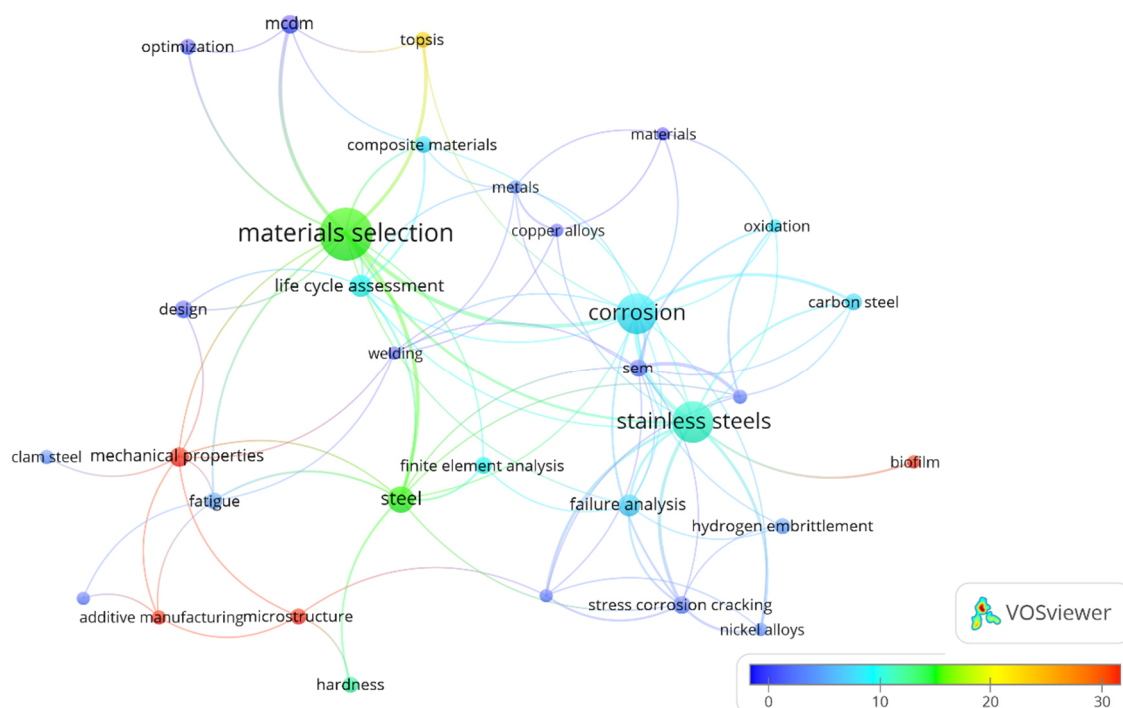
A substituição parasita é a opção mais simples, porém somente se apresenta viável quando um novo material difere de seus semelhantes por uma única característica marcante [20,21], nem sempre relacionada ao desempenho mecânico ou funcional, como por exemplo custo ou disponibilidade. Dessa maneira, não necessariamente representa um ganho de desempenho no conceito de índice de mérito, mas pode representar uma otimização em um item pontual considerando toda a cadeia de produção e ciclo de vida de um produto.

A busca por otimizações de requisitos de projetos é mais complexa pois consiste na busca em um universo de aplicações conhecidas e a tradução de seus objetivos em funções e pesos aplicados para seleção dos materiais atualmente utilizados. Assim, seria necessário o acesso a bases de dados de aplicações para que então sejam verificadas àquelas cujo as propriedades do novo material seriam potencialmente promissoras [20,21]. A concepção e atualização dessas bases seria complexa devido a características específicas de cada segmento de mercado [20,21]. Por exemplo, um material resistente e leve

é desejado tanto pela indústria automobilística quanto aeronáutica, entretanto, os requisitos adicionais e suas ponderações levam a opções finais distintas.

As alternativas anteriores representam extremos no que tange a acesso de informação e complexidade da tomada de decisão. Assim, a busca por funções se apresenta como uma abordagem intermediária que pode ser traduzida em uma verificação de oportunidades a partir de índices de mérito conhecidos [20,21]. Listas de índices de mérito para requisitos mecânicos e geométricos mais comuns podem ser obtidas em livros, artigos, memoriais de projetos etc. ou calculados a partir dos conceitos de, por exemplo, mecânica dos sólidos. Dessa maneira, uma triagem inicial de oportunidades em que o novo material candidato é competitivo se torna mais simples, pois a segunda etapa consiste em avaliar quais aplicações valorizam esses índices de mérito e seriam mais promissoras.

Uma análise dos trabalhos mais recentes em seleção de materiais que envolvem aços apresenta um destaque para os temas de corrosão e aplicação de aços inoxidáveis, métodos de otimização para tomada de decisão, design, propriedades mecânicas e ciclo de vida dos produtos. A Figura 2.4 apresenta uma rede de co-ocorrência entre as 32 palavras-chave de autores mais presentes nos artigos relacionados a seleção de materiais e aços nos últimos 10 anos, após uma busca na *Web of Science*.



Elaborado pelo autor com auxílio do software VOSviewer [24] a partir da co-ocorrência de palavras-chave dos autores. Atribuição de cores pela citação média dos artigos que contém os termos. Expressão de busca:  $TS=(steel^*)$  AND  $TS=(\text{"material* selection"}$  OR  $\text{"material* choice"}$  OR  $\text{"selection of material*"})$ . Anos de publicação 2011 a 2020. Total de 656 registros. Apresentadas as 32 palavras-chave com 5 ou mais ocorrências.

Figura 2.4 – Rede de co-ocorrência de palavras-chave dos autores de artigos de seleção de materiais e aços em função do índice médio de citações.

Ao avaliar as palavras-chave com maior número de citações em média dos artigos (coloração mais próxima a vermelha na Figura) nota-se um maior índice de referências de trabalhos que contém os termos de propriedades mecânicas, microestrutura, processos (destaque para manufatura aditiva) e aplicações (destaque para biofilmes). Uma hipótese para tal comportamento pode estar associada à busca de informações sobre esses temas a partir dessas publicações. Pelo fato de possivelmente apresentarem dados sobre experimentos, estudos de caso e avaliação de situações reais de análises de falhas, é notável que os artigos científicos apresentam uma quantidade de informação confiável que pode ser utilizada para atividades de seleção de materiais. Entretanto, assim como para documentos de patentes, ainda se trata

de um desafio lidar com tal volume de publicações e na extração mais automatizada desses dados de interesse.

Entre os trabalhos que envolvem as propriedades dos materiais, um problema comum trata da manutenção (ou melhoria) de propriedades mecânicas com foco na redução de peso dos materiais ou de componentes, muito estimulados pela indústria de transporte [25,26], em que não se espera mudanças de formas muito bruscas, devido ao nível de avanço de simulações mecânicas e aerodinâmicas desses setores. Nesse quesito, os materiais compósitos vêm apresentando destaque por conciliarem altos limites de resistência à tração e/ou impacto com uma baixa densidade, embora ainda apresentem desafios em termos de processamento, resíduos em sua manufatura e disposição no final do ciclo de vida [27]. Pelo fato de o aço ser um material com facilidade de reprocessamento, embora com um alto custo energético para tal [28], sua citação em estudos de seleção de materiais pode tratar tanto da sua substituição, como das possibilidades de aplicação a partir de melhorias de suas propriedades.

O avanço de recursos computacionais trouxe a possibilidade da utilização de algoritmos mais robustos para as atividades de seleção de materiais [16]. Entretanto, ainda não é possível estabelecer um método multicritério geral e totalmente eficaz para atividades de seleção de materiais e de aplicações, o que torna cada vez mais comum o desenvolvimento e aplicação de métodos integrados ou híbridos [29,30] para minimizar incertezas no processo de escolha. Os procedimentos com maior número de publicações recentes em atividades de seleção de materiais são os métodos AHP (*Analytic Hierarchy Process*), TOPSIS (*Technique for Order of Preference by Similarity to Ideal Solution*) e VIKOR (*Vlsekriterijumska Optimizacija I Kompromisno Resenje*) [31]. Os três métodos partem do princípio da não-existência de uma solução ideal, e avaliam a diferença de desempenho dos candidatos ao que seria o desempenho ótimo através de diferentes tratamentos matemáticos e estatísticos para a elaboração de uma ordem classificatória. Assim, a resposta obtida é o material mais adequado, não necessariamente o ideal, entre as soluções apresentadas, de acordo com o desempenho informado e os pesos estabelecidos.



Por fim, atividades de seleção de materiais envolvendo aços, seja para sua substituição ou que avaliem os benefícios de suas propriedades são ainda alvo de muitos estudos recentes, visto o volume de resultados da pesquisa que originou a Figura 2.4. Logo, no contexto de um mercado que movimenta cerca de 800 bilhões de dólares anuais [32], tanto o risco de erros quanto oportunidades de ganhos são evidenciados, seja em monitoramento de tecnologias ou na correta escolha de suas aplicações.

### **2.1.1 Fontes de informação para atividades de seleção de materiais**

Ao abordar os atuais desafios da seleção de materiais e aplicação a partir de suas etapas (Figuras 2.1 e 2.3), um desafio comum é encontrar informações sobre materiais, processos e aplicações em nível de detalhe suficiente para análise e tomada de decisão. Algumas soluções comerciais e não comerciais são disponíveis, com diferentes recursos de apoio de análises gráficas, abrangência de escopo e custos. Ramalhe et. al. (2010) [7] avaliou diversas ferramentas digitais utilizadas para seleção de materiais. A Tabela 2.1 apresenta exemplos de bases de dados estudadas no decorrer da elaboração desta tese, algumas de suas características e limitações avaliadas pelo autor.

Tabela 2.1 – Exemplos de bases de dados de materiais gratuitas e comerciais disponíveis para consulta.

Base de dados de materiais <sup>1</sup> [referência]:	Características:	Limitantes:
ANSYS GRANTA [33]	<ul style="list-style-type: none"> <li>- Apresenta ferramentas gráficas para comparações entre opções selecionadas.</li> <li>- Apresenta informações para etapas de documentação como custos e fornecedores.</li> <li>- Realiza triagem de opções a partir de informações inseridas.</li> <li>- Disponibiliza diferentes ferramentas para análise gráfica.</li> <li>- Base de funções para busca de aplicações.</li> <li>- Atualizações constantes (versões anuais do software).</li> </ul>	- Não gratuita.
Total Materia [1]	<ul style="list-style-type: none"> <li>- Se apresenta como a base de dados mais completa com 20 milhões de propriedades de 450.000 opções disponíveis.</li> <li>- Apresenta pacotes de dados específicos para classes específicas de materiais como metais e polímeros.</li> <li>- Permite comparações de múltiplas opções.</li> </ul>	- Não gratuita
Matweb [34]	- Gratuita com recursos adicionais para usuários registrados.	- Cobertura focada para classes específicas de metais e polímeros.

<sup>1</sup> A avaliação das bases de dados comerciais (não-gratuita) foi realizada a partir das informações contidas nos sites de seus desenvolvedores, sem acesso às versões completas dos produtos.

	<ul style="list-style-type: none"> <li>- Apresenta informações sobre composição química e propriedades físicas.</li> <li>- Alimentada por empresas e associações de classe.</li> <li>- Realiza triagem básica por composições químicas e propriedades físicas e mecânicas.</li> </ul>	<ul style="list-style-type: none"> <li>- Limitação em opções de comparação entre várias opções.</li> <li>- Não apresenta detalhes de informações para etapas de documentação.</li> </ul>
Make it From [35]	<ul style="list-style-type: none"> <li>- Gratuita.</li> <li>- Realiza comparações entre opções selecionadas.</li> <li>- Interface gráfica para comparação entre opções.</li> <li>- Apresenta informações sobre composição química e propriedades estimadas a partir de normas.</li> <li>- Apresenta informações simples para documentação.</li> </ul>	<ul style="list-style-type: none"> <li>- Informações genéricas por classes de materiais, por exemplo em valores de densidade e limite de escoamento, sem detalhar variações específicas.</li> <li>- Não realiza triagem das opções a partir de valores inseridos pelo usuário.</li> </ul>
CustomPartNet [36]	<ul style="list-style-type: none"> <li>- Gratuita.</li> <li>- Realiza triagem de opções a partir de informações inseridas.</li> <li>- Alimentada por usuários, empresas e associações de classe.</li> <li>- Apresenta informações sobre custos e estudos de caso para etapas de documentação.</li> <li>- Apresenta opções a partir de aplicações específicas.</li> </ul>	<ul style="list-style-type: none"> <li>- Focada em materiais metálicos e poliméricos e em aplicações para ferramentais.</li> <li>- Dificuldades na comparação entre várias opções.</li> </ul>
Knovel [37]	<ul style="list-style-type: none"> <li>- Gratuita</li> <li>- Realiza triagem de opções a partir de informações inseridas.</li> <li>- Apresenta informações sobre substâncias.</li> </ul>	<ul style="list-style-type: none"> <li>- Abrangência do escopo não é clara.</li> <li>- Não apresenta detalhes de informações para etapas de documentação.</li> </ul>

Fonte: Elaboração do autor.

Para trabalhos de seleção de aplicação são desconhecidas bases de dados que contenham informações específicas sobre o conjunto de atributos em que um novo material poderia ser competitivo. As soluções apresentadas se baseiam em *guidelines* de setores específicos, como o automotivo [38]. Embora cada setor da indústria se responsabilizasse pela criação de bases de dados de aplicações específicas de suas atividades, a complexidade das mesmas ainda seria desafiadora em termos de manutenção, atualização [20,21] e disponibilidade de desenvolvedores de materiais em fornecer dados realistas sobre desempenho antes de protegidos seus ativos intelectuais pela atividade de patenteamento.

O acesso a informações confiáveis e o tratamento com grandes volumes de dados são apresentados como desafios complexos de engenharia [39], o que por definição engloba os processos de seleção de materiais. Assim, técnicas de bibliometria e mineração de dados e textos se apresentam como ferramentas úteis neste contexto em vista de facilitar a interpretação de grande volume de dados, detectar tendências não facilmente perceptíveis e orientar o uso de alternativas como novos materiais e/ou aplicações [5].

A etapa de documentação de materiais e aplicações também apresenta limitações quanto a disponibilidade de acesso a informações atualizadas, tanto em termos de custos quanto de abrangência de cobertura. A falha em acessar e compreender desenvolvimentos recentes, bem como seus sucessos e desafios em uma área específica, pode levar a investimentos desnecessários e invenção de soluções redundantes, inclusive de novos materiais e processos. Dessa maneira, a informação contida em documentos de patentes pode ser também uma alternativa para complementar estudos de seleção de materiais e/ou a busca por aplicação de novas composições químicas e processos.

## **2.2 Informação Tecnológica e Engenharia de Materiais**

Para Paganini (2004 apud SILVA et al., 2013, p.141) a Informação Tecnológica pode ser definida como “todo tipo de conhecimento sobre tecnologias de fabricação, de projeto e de gestão, que favoreça a melhoria contínua da qualidade e a inovação no setor produtivo” [40]. Por essa razão, o

conteúdo de documentos de patentes se apresenta como uma importante fonte de informação tecnológica por tratar, desde o título até a reivindicação proposta pelos inventores, de aspectos técnicos, legais e comerciais, em um formato padronizado e conhecido em todo o mundo [41]. A patente é um título de propriedade temporário que permite ao detentor o direito de exploração comercial exclusiva da tecnologia desenvolvida [42]. Para sua concessão, é necessário que o titular forneça detalhes de sua invenção com suficiência descritiva, e no que ela se difere das soluções que já existem no estado da técnica, além de indícios de aplicabilidade industrial [43]. Logo, também se trata de uma fonte de informação atualizada, e que geralmente não está disponível em outro tipo de registro de conhecimento, mesmo em livros técnicos e artigos científicos [43]. No âmbito da ciência e engenharia de materiais, essas informações podem tratar de uma nova composição química, novo processo ou condição de trabalho, ou a adaptação para uma aplicação ainda não explorada, entre outros.

O acesso ao conteúdo dos documentos pode ser realizado através de bases de textos completos dos documentos, geralmente disponibilizadas gratuitamente por escritórios de propriedade industrial de cada país, quanto em bases referenciais que compilam informações básicas em forma de registros bibliográficos, com cobertura internacional [44]. Devido ao crescente volume de informação disponível, torna-se cada vez mais desafiador coletar e analisar grandes volumes de dados e de textos, em que se faz necessário o uso de ferramentas computacionais para tornar essas etapas semiautomatizadas e auxiliar analistas e tomadores de decisão [45]. Logo, os dados e textos desses registros podem ser a base da elaboração de indicadores tecnológicos, que auxiliam a compreender o panorama tecnológico de um setor, observar tendências e suportar a tomada de decisão em atividades de pesquisa e desenvolvimento [46,47]. Essa atividade, conhecida como monitoramento tecnológico, trata da busca por responder às questões “quem está fazendo?”, “o quê?” e “onde?”, além de estimar o esforço científico, de pesquisa e de desenvolvimento em uma tecnologia de interesse [46,48].

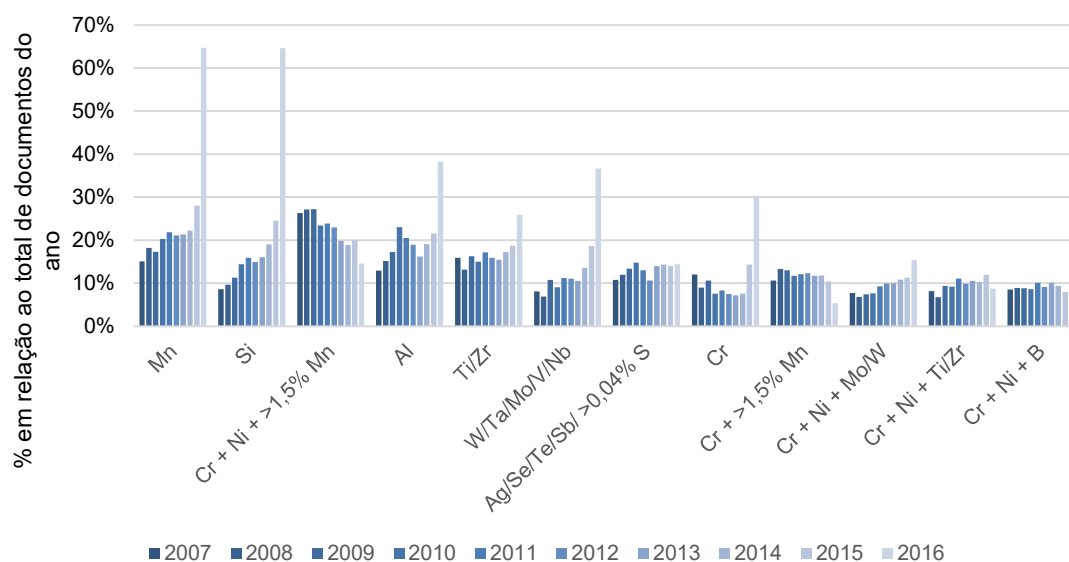
Para a elaboração desses indicadores são utilizados conhecimentos de bibliometria, que consiste em métodos estatísticos e ferramentas para elaboração de métricas de comunicação técnica e científica [49]. Trata-se de uma abordagem quantitativa para análise da informação registrada em diferentes meios, tendo sido aplicada com sucesso em artigos científicos e patentes como fonte de informação [50]. Para a manipulação desse conteúdo podem ser utilizados softwares para análise bibliométrica que facilitam a seleção de segmentos de textos de interesse e operações de agrupamento e contagem, e o processamento de linguagem natural. Esse último, é de fundamental importância para estudos bibliométricos e elaboração de indicadores a partir de patentes por transformar segmentos de textos de informação tecnológica em estruturas de linguagem mais simples e buscar relações entre eles [51].

Uma fonte de informação contida nos documentos e registros de patentes, bastante aplicada para elaboração de indicadores envolvendo materiais, processos e aplicações são os códigos da Classificação Internacional de Patentes (CIP) [52]. Trata-se de uma classificação elaborada com o objetivo de agrupar documentos referentes a um mesmo tema tecnológico. Sua estrutura é composta por 8 seções principais (Códigos que iniciam com as letras de A a H), seguida de classe (dois números, totalizando 3 dígitos), subclasse (mais uma letra, totalizando 4 dígitos), grupo principal (de 1 a 3 algarismos totalizando até 7 dígitos) e subgrupo, após uma barra oblíqua "/" e mais um ou dois algarismos totalizando até 9 dígitos [53]. Essas subdivisões, que atualmente contemplam cerca de 70 mil códigos distintos, permitem categorizar os inventos de interesse em termos de funções e aplicações, por exemplo [54], além de refinar informações sobre composição química como, por exemplo, elementos de liga em aços (grupo C22C/038<sup>1</sup>) e seus processos de beneficiamento [55]. Esses códigos são atribuídos por analistas de patentes quando da análise dos pedidos nos escritórios de propriedade industrial onde são depositados. Como esses pedidos são direcionados a especialistas da área de conhecimento a que se destina a invenção, a classificação tende a ser apurada e confiável. Como

---

<sup>1</sup> Como se trata do principal código utilizado na busca de registros de patentes para a elaboração desta tese, sua descrição é apresentada no Anexo A.

exemplo de uso dos códigos da CIP para elaboração de indicadores em estudos de monitoramento tecnológico em materiais, a Figura 2.5 apresenta dados sobre elementos de liga em patentes relacionadas a aços para inventos classificados até o nível de subgrupo.



Fonte: [10].

Figura 2.5 – Representação percentual dos principais elementos de liga abordados em patentes de aços entre 2007 e 2016 a partir de códigos da CIP.

Os códigos da CIP apresentam limitações quanto a abrangência de detalhes. Por exemplo, o código C22C-038/14 descreve ligas de aço contendo titânio ou zircônio, mas não especifica qual dos elementos é referente à invenção, se são ambos, nem mesmo qual seus teores na liga. Na Figura 2.5, também se nota que o percentual de patentes classificados nos códigos CIP referentes a alumínio e silício sofreu um expressivo crescimento no último período avaliado sem, entretanto, apresentar um detalhamento como o teor mínimo ou máximo desses elementos nas ligas. Dessa maneira, são necessários esforços na mineração de textos em documentos completos de patentes ou de seus registros bibliográficos para maior riqueza de detalhes que possam ser ainda mais úteis ao monitoramento tecnológico e a processos de seleção de materiais e de aplicações.

Os indicadores de patentes têm auxiliado na identificação de ciclo de vida de tecnologias, alternativas técnicas para um mesmo desafio, áreas e subáreas tecnológicas de interesse etc. [5]. Como exemplo, pesquisadores do Núcleo de Informação Tecnológica em Materiais (NIT/Materiais) da Universidade Federal de São Carlos têm trabalhado na interface entre o uso de indicadores a partir de registros de patentes e a área de seleção de materiais. Além de aços, destacam-se estudos em classes de materiais e tecnologias para armazenagem de hidrogênio [56], desafios técnicos em componentes para o mercado de óleo e gás [57], patenteamento por classes de nanocelulose [9,58], aspectos de seleção para compósitos sustentáveis para interiores de aeronaves [27], processos e tecnologias de metais amorfos [59] e de solidificação rápida [60], entre outros.

### **2.2.1 Mineração de textos em patentes para elaboração de indicadores**

O volume de informação gerado, registrado e armazenado em todo o mundo cresce anualmente, embora o conhecimento que é criado por essa informação não se dá na mesma velocidade [61]. Assim, ferramentas computacionais são necessárias para trabalhar com diferentes fontes de informação e com a forma com que se apresentam. Nesse contexto, os termos mineração de dados e mineração de textos se popularizaram e tratam de ferramentas que possibilitam extrair conteúdo a partir de textos, estruturados ou não. Geralmente, a mineração de dados lida com a informação contida em bancos ou bases de dados, com uma estrutura bem definida, que permite seu acesso a partir de sistemas gerenciadores [62]. Já a mineração de textos, embora se encontrem diversas definições para o termo, tem como objetivo geral [63]:

[...] explorar informações contidas em documentos textuais de várias maneiras, incluindo o tipo de análises que são tipicamente realizadas em mineração de dados: descoberta de padrões e tendências em dados, associações entre entidades, regras preditivas etc.<sup>1</sup>

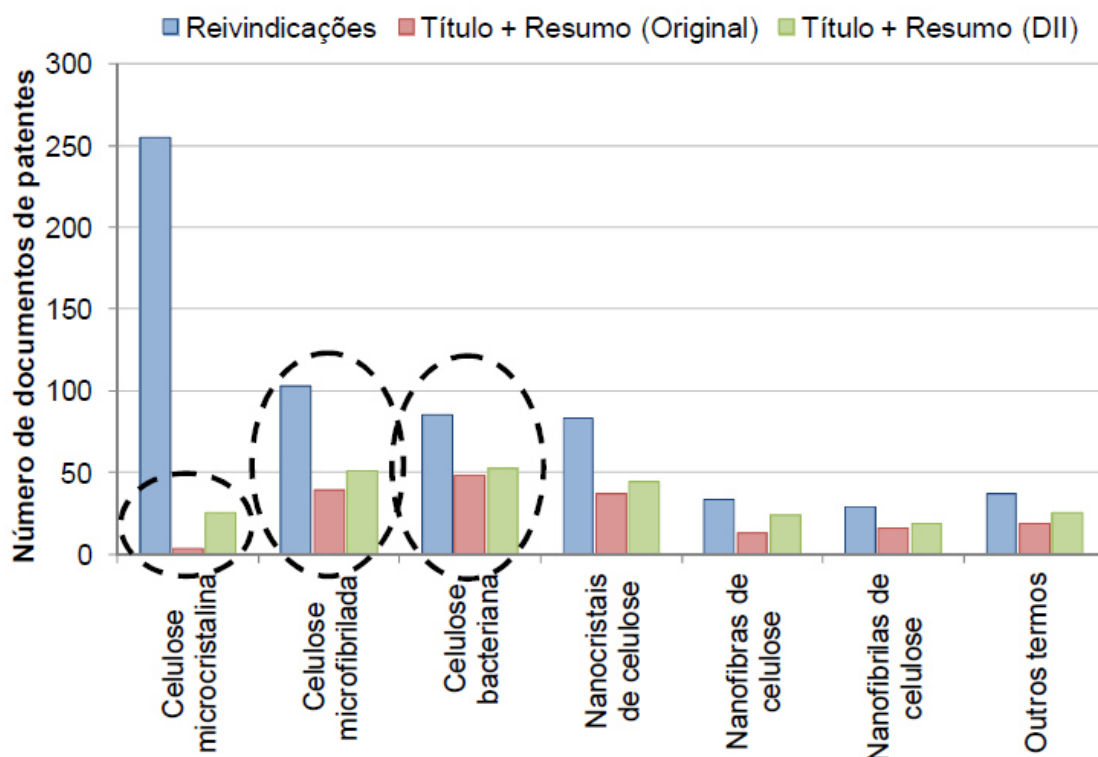
---

<sup>1</sup> Tradução realizada pelo autor do trecho original: “*The objective of Text Mining is to exploit information contained in textual documents in various ways, including the type of analyses that are typically performed in Data Mining: discovery of patterns and trends in data, associations among entities, predictive rules, etc.*” p.1.



Técnicas de mineração de textos podem ser utilizadas para a elaboração de indicadores baseados em documentos patentes, sendo a partir de registros bibliográficos ou da seção de reivindicações no texto completo [8,64]. Os avanços recentes nessa área têm se baseado em análise de termos de títulos e resumos, citação entre documentos, análise semântica e aproximação de termos [50,65]. Um outro desafio dos estudos envolvendo grande volume de informação trata da representação destes dados em uma forma visual que facilite a análise. Neste cenário, destacam-se análise por *clusters*, redes de citação e co-ocorrência, curvas de extrapolação e de ciclos de vida das tecnologias [66].

A principal vantagem no trabalho com texto completo é a riqueza de informações e detalhes técnicos apresentados, principalmente no setor de reivindicações desses documentos. Entretanto, esse procedimento apresenta limitações para recuperação das informações pois cada escritório de patentes nacional disponibiliza os documentos de uma maneira diferente, geralmente em arquivos de imagem ou no formato “.pdf”. Adicionalmente, é necessário um domínio no idioma em que os documentos são apresentados. Milanez et al. [8] avançaram na recuperação de informações em textos completos do escritório estadunidense de patentes (USPTO), e elaboraram indicadores a partir de informações contidas nos textos completos. A Figura 2.6 apresenta uma comparação entre os tipos de materiais identificados (nanocelulose) a partir do acesso ao texto das reivindicações, nos títulos e resumos do documento original, e do mesmo procedimento a partir de registros indexados da base de registros bibliográficos *Derwent Innovations Index*.



Fonte: [67].

Figura 2.6 – Total de documentos de patente de diferentes classes de nanocelulose a partir de diferentes fontes de informação.

Por outro lado, a mineração de textos a partir de registros bibliográficos de patentes se torna interessante por esses registros apresentarem uma abrangência maior de documentos depositados em outros países, a partir de bases de cobertura internacional. Entretanto, como observado na Figura 2.6, os resumos recuperados geralmente possuem uma riqueza de detalhes menor comparado aos textos completos.

Para a elaboração da planilha de dados desenvolvida nesta tese, optou-se por trabalhar com os registros bibliográficos de documentos de patentes para avaliar uma cobertura global para os desenvolvimentos recentes de açós. Assim, a base de dados escolhida foi a *Derwent Innovations Index* (DII), com cobertura de registros de patentes de 59 países [68] e com acesso gratuito para atividades de pesquisa pelo portal de periódicos da CAPES [69]. Outras bases de dados como o Google Patents [70] e a Espacenet [71], disponibilizada pelo Escritório Europeu de Patentes (EPO), possuem uma dificuldade maior, quando

comparadas à DII, para a exportação dos títulos, resumos, e demais conteúdos, alguns deles obtidos a partir de tradutores. Uma outra vantagem no uso da DII é seu processo de reescrita de títulos e resumos dos registros de documentos que visa facilitar a interpretação do invento proposto [68]. Isso também permitiu incluir na base de dados e nos indicadores elaborados as informações de documentos não escritos originalmente no idioma inglês. Nesse contexto, destacam-se as patentes de origem chinesa, com um número em crescimento recente e com pedidos geralmente não estendidos para outros países por questões de políticas internas [72,73], como também já observado em estudos anteriores em patenteamento de aços e outras classes de materiais [56,60,74]. A Figura 2.7 apresenta como exemplo o trecho de um registro como apresentado pela DII, enquanto a Figura 2.8 apresenta um trecho do documento original referente a este registro no idioma japonês, a título de exemplo, com destaque para as informações de interesse buscadas para a elaboração do procedimento desenvolvido nesta tese.

**TI** Hot-rolled sheet steel used for manufacturing electro-resistance-welded tubes for coiled tubing, has structure comprising martensite, retained austenite, and bainite, and preset yield strength, tensile strength, and elongation.

**GA** 201860390W

**AB** **NOVELTY** - A hot-rolled sheet steel comprises greater than 0.1 %mass to 0.16 %mass or less carbon, 0.1-0.5 %mass silicon, 1.6-2.5 %mass manganese, 0.02 %mass or less phosphorus, 0.005 %mass or less sulfur, 0.01-0.07 %mass aluminum, greater than 0.5 %mass to 1.5 %mass or less chromium, 0.1-0.5 %mass copper, 0.1-0.3 %mass nickel, 0.1-0.3 %mass molybdenum, 0.01-0.05 %mass niobium, 0.01-0.1 %mass vanadium, 0.005-0.05 %mass titanium, 0.005 %mass or less nitrogen, and remaining iron and unavoidable impurities. The sheet has a structure comprising 3-20 vol.% martensite, 10 vol.% or less retained austenite, and remaining bainite. The sheet steel has a yield strength of 600 MPa or more, a tensile strength of 950 MPa or more, and a uniform elongation of 7% or more.

**USE** - Hot-rolled sheet steel is used for manufacturing electro-resistance-welded tubes for coiled tubing used for various operations in well and in shale gas mining for hydraulic fracturing of shale layer.

**ADVANTAGE** - The hot-rolled sheet steel has excellent roll forming workability and yield strength, and enables economical manufacture of high-quality electro-resistance-welded tubes with high productivity.

**TF** **TECHNOLOGY FOCUS - METALLURGY** - Preferred Composition: The sheet steel further comprises 0.001-0.005 %mass tin and/or 0.001-0.003 %mass calcium.

**IP** C21D-009/08; C21D-009/46; C21D-009/50; C22C-038/00; C22C-038/58; C21D-008/02; C22C-038/02; C22C-038/06; C22C-038/42; C22C-038/44; C22C-038/46; C22C-038/48; C22C-038/50

#### Legenda:

**TI:** Título do registro; **GA:** Código de identificação único do registro; **AB:** Resumo; **NOVELTY:** Campo de "Novidade" do registro; **USE:** Campo de "uso" do registro; **ADVANTAGE:** Campo de "vantagem" do registro; **TF:** Campo de "Foco Tecnológico" do registro. **IP:** Campo dos códigos CIP do registro.

**Cores:** Composição química; Fases; Processos de fabricação; Propriedades; Aplicação.

Fonte: Elaboração do autor. Trecho do registro 201860390W na *Derwent Innovations Index*.

Figura 2.7 – Exemplo de um trecho de registro de documento de patente na base DII com destaques para as informações utilizadas para elaboração da base de dados.

## (57) 【特許請求の範囲】

## 【請求項 1】

質量 % で、C : 0.10 % 超 0.16 % 以下、  
 Si : 0.1 % 以上 0.5 % 以下、  
 Mn : 1.6 % 以上 2.5 % 以下、  
 P : 0.02 % 以下、  
 S : 0.005 % 以下、  
 Al : 0.01 % 以上 0.07 % 以下、  
 Cr : 0.5 % 超 1.5 % 以下、  
 Cu : 0.1 % 以上 0.5 % 以下、  
 Ni : 0.1 % 以上 0.3 % 以下、  
 Mo : 0.1 % 以上 0.3 % 以下、  
 Nb : 0.01 % 以上 0.05 % 以下、  
 V : 0.01 % 以上 0.10 % 以下、  
 Ti : 0.005 % 以上 0.05 % 以下、

N : 0.005 % 以下を含み、残部 Fe および不可避免的不純物からなる成分組成を有し、  
 体積分率で、3% 以上 20% 以下のマルテンサイトと、10% 以下の残留オーステナイトと、  
 残部としてベイナイトと、からなる組織を有し、降伏強度が 600MPa 以上であり、引張強度  
 が 950MPa 以上であり、一様伸びが 7.0% 以上である、コイルドチュービング用熱延鋼板。

## 【請求項 2】

Fonte: Recorte de JP6384635-B1 [75] em sua versão original obtida pelo *Japan Patent Office*.

Figura 2.8 – Trecho de documento original da patente com o registro bibliográfico apresentado na Figura 2.7.

Assim nota-se que, para os temas de interesse da elaboração de uma planilha de dados de materiais, devem ser desenvolvidos mecanismos para a mineração desses segmentos de textos nos registros bibliográficos. Logo, o próximo desafio consiste em buscar registros que contenham essas informações no escopo desejado, no caso desta tese, os aços. Geralmente, o desenvolvimento de expressões de buscas para encontrar registros de patentes em bases de dados envolve o uso de palavras-chave que descrevam o assunto de interesse, assim como possíveis sinônimos, combinados com códigos da Classificação Internacional de Patentes. A procura pelos termos mais adequados e das divisões da CIP para uma boa abrangência e precisão dos resultados é um processo subjetivo, que necessita conhecimento no tema estudado e requer tempo para testes dos termos e suas combinações até que os resultados permitam uma análise representativa e que seja reproduzível em estudos futuros. A Tabela 2.2 apresenta um exemplo de expressão de busca para a recuperação

de registros de patentes em uma classe específica de materiais, como aços de baixa densidade.

Tabela 2.2 – Exemplo de expressão de busca desenvolvida para recuperação de registros de patentes em aços de baixa densidade.

ETAPAS	TERMOS DA EXPRESSÃO DE BUSCA
#1	(IP=C22C-038*) AND (TI=steel*)
#2	TS=(((("lightweight" OR "light-weight" OR "low-density" OR "low density" OR "lower density" OR "light weight") NEAR/4 steel*) OR "k* carbide" OR "(Fe,Mn)3AlC" OR "Fe-Mn-Al-C" OR "Fe-Al-Mn-C" OR "Fe-C-Al-Mn")
#3	#1 AND #2

Fonte: [74].

Logo após a recuperação desses registros é necessário estudar a forma como são apresentadas as informações de interesse. No caso da DII, como observado na Figura 2.7, são apresentados campos estruturados como o título (sempre precedido pela sigla TI), códigos da CIP (IP) e o resumo (AB) e suas subdivisões (*Novelty*, *Use*, *Advantage*) entre outros. Assim, o processo de mineração de textos deve começar a partir da seleção dos textos desses campos, para em seguida partir para o refinamento da seleção de segmentos de interesse. Em alguns casos, é necessário utilizar o auxílio de marcadores, que consistem em caracteres conhecidos que serão utilizados nos comandos para iniciar ou parar a seleção de um segmento de texto. Esses marcadores podem ser expressões regulares, que consistem em uma sequência de caracteres que se repetem de forma constante em algum texto como, por exemplo, os próprios códigos de denominação dos campos dos registros da DII.

No trabalho com textos livres, o uso de tesouros pode ser útil para atividades de agrupamentos de termos que podem ser sinônimos ou descrevem o mesmo conceito ou denominação. No campo da ciência da informação, o termo tesouro é definido como “linguagem documentária alfabética que relaciona os termos, geralmente de uma área específica do conhecimento, cuja organização demonstra a estrutura conceitual dessa área, por meio de abreviaturas convencionadas.” [76]. No trabalho de mineração de textos, esses dicionários

podem ser utilizados para substituir termos e expressões conhecidas para um mesmo padrão, o que também facilita a posterior atribuição de marcadores para início e fim de seleção de segmentos de texto. Por outro lado, podem também serem utilizados em etapas de pré-processamento do texto, como a prévia exclusão de termos não desejados para as futuras etapas de seleção de trechos de interesse.

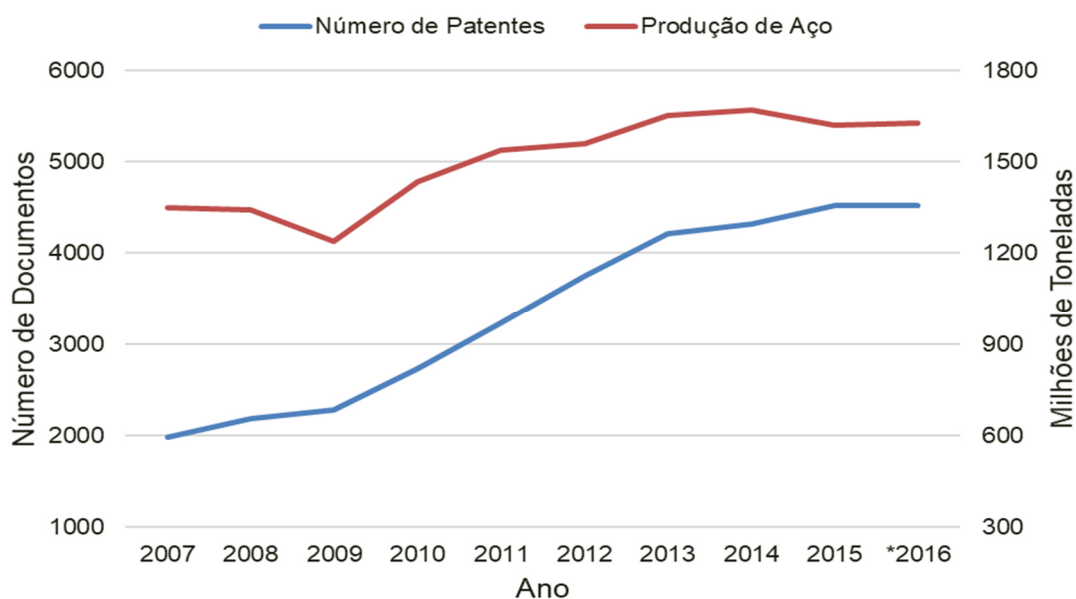
Entretanto, apesar de aparentemente não-estruturado, os textos contidos ou provenientes de documentos de patentes geralmente apresentam uma estrutura similar, pela própria característica de linguagem utilizada nesses documentos, o que torna o processo um pouco mais simples quando comparado a textos totalmente livres [51,77]. Por exemplo, ao observar o registro apresentado na Figura 2.7, nota-se que os elementos de liga são apresentados no campo “*Novelty*” após o termo “*comprising*”, e suas quantidades são expressas seguidas de uma sequência de caracteres “*%mass*” seguidas do nome dos elementos, separados por vírgulas. A identificação desses padrões é fundamental para a atribuição de marcadores para seleção de segmentos de texto.

Como colocado nos objetivos deste trabalho, a elaboração de uma planilha de dados de materiais depende de extrair as informações contidas em seus registros, através de técnicas de mineração de textos ou relacionadas. Assim, segmentos específicos a partir de textos semiestruturados poderão se tornar mais organizados e estruturados em uma planilha, para posterior uso em indicadores para o monitoramento tecnológico e atividades de seleção de materiais. O recorte para uma primeira versão deste procedimento pelo estudo de aços também se justifica pela maior facilidade esperada de se encontrar os padrões de linguagem utilizados nos registros de materiais semelhantes, que poderão futuramente ser expandidos para outras ligas metálicas e classes de materiais.

### **2.3 Desenvolvimentos recentes em aços**

Apesar de sua utilização secular e de um expressivo volume de produção e comércio a nível global, ainda existem desafios na compreensão da ciência

básica dos aços e de desenvolvimento de novas ligas e processos. Por se tratar de um setor que vive uma competição tecnológica para seus produtos de maior valor agregado, espera-se que a atividade de patenteamento seja bastante expressiva no mercado siderúrgico, sua cadeia de beneficiamento e de aplicações. A Figura 2.9 apresenta o número de depósitos de patentes relacionadas a aços e a produção bruta do material nos últimos 10 anos.



\* Dados que podem sofrer alterações do número de patentes visto período de confidencialidade e tempo para indexação na base. Expressão de busca: TS=(steel\*) AND IP=(C22C-038\*).

Fonte: [10].

Figura 2.9 – Evolução temporal do depósito de patentes e da produção mundial de aço bruto entre 2007 e 2016.

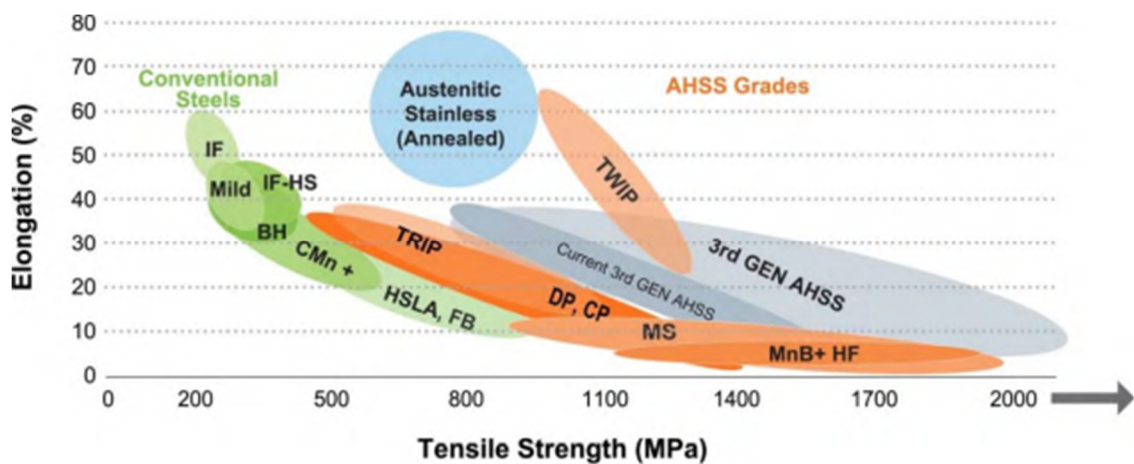
Enquanto a taxa média de evolução da produção de aço para o período avaliado foi de 2,30%, o crescimento do número de pedidos de patentes foi 9,83%, pouco mais de quatro vezes maior. Com os dados de 2016, obtém-se uma média de 12 depósitos por dia em todo o mundo. Como já estimado também a partir da Figura 2.4, estudos recentes em aços visam alternativas para os seus principais limitantes em projetos de seleção de materiais como sua relação entre limite de resistência e peso, além da resistência ao desgaste e à corrosão [38]. Adicionalmente, estudos relacionados ao impacto ambiental, eficiência



energética e custos também se apresentam como critérios industriais com grande atenção por parte de pesquisadores [78].

Um dos principais mercados que impulsiona novos desenvolvimentos em aços é o automobilístico pois o material ocupa, em diferentes ligas, a maior parte estrutural de um automóvel. Assim, as montadoras buscam, principalmente, a redução do peso dos automóveis para uma melhor eficiência energética de seus combustíveis [79]. Nesse cenário, materiais compósitos ganham destaque como concorrentes, embora a indústria de aço também busque soluções para manter seu domínio [80]. Diversas alternativas têm se mostrado promissoras para melhorar a relação entre resistência mecânica e peso das peças produzidas em aços, em que o foco em aumentar o limite de resistência à tração é mais comum. Entretanto, novos caminhos também são estudados como diminuir a densidade das ligas ferrosas ou aumentar seu módulo elástico, que levaria a estrutura do automóvel ser produzida a partir de componentes com menor área de seção e com um menor consumo de material [81,82]. Em ambos os casos, as soluções devem também atender a requisitos de processamento das peças automotivas, como ductilidade para estampagem profunda e resistência à corrosão, o que não se trata de um desafio trivial de ser solucionado [83].

Recentemente deu-se origem à terceira geração dos aços avançados de alta resistência (*Advanced High Strength Steels - AHSS*) que se apresentam como a nova fronteira de desenvolvimento para aplicações estruturais, em especial, no setor automotivo [38,79]. A Figura 2.10 apresenta um mapa de relação entre propriedades de alongamento e resistência à tração de aços tipicamente utilizados nesse setor, incluindo a terceira geração dos de alta resistência, em recente desenvolvimento.

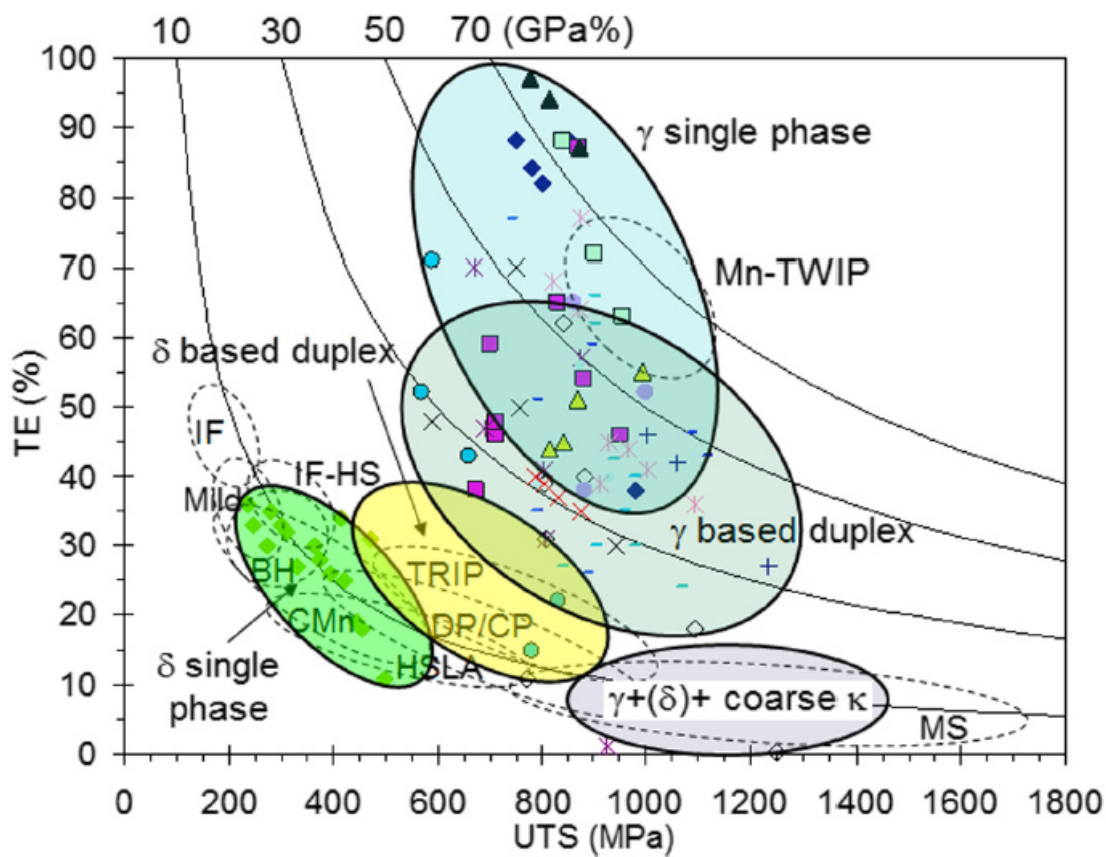


*Elongation*: Alongamento; *Tensile Strength*: Resistência a tração.

Fonte: [38].

Figura 2.10 – Diagrama resistência em função do alongamento de aços utilizados na indústria automobilística.

Embora um dos focos seja a melhoria da relação entre resistência e ductilidade na também denominada “curva da banana” do diagrama apresentado, um dos objetivos desses desenvolvimentos recentes é que essas novas soluções tenham custo competitivo [79,84]. Assim, não existe ainda uma única rota determinada, em que as propostas partem do desenvolvimento de microestruturas complexas pelas diferentes combinações de elementos de liga e condições de processamento, embora ainda desafiadores de serem aplicados em escala industrial [84]. A Figura 2.11 apresenta um compilado de resultados de estudos de ligas Fe-C-Mn-Al com detalhes que os diferenciam pela microestrutura apresentada.



TE – *Total Elongation*: Alongamento total. UTS – *Ultimate Tensile Strength*: Limite de resistência à tração.

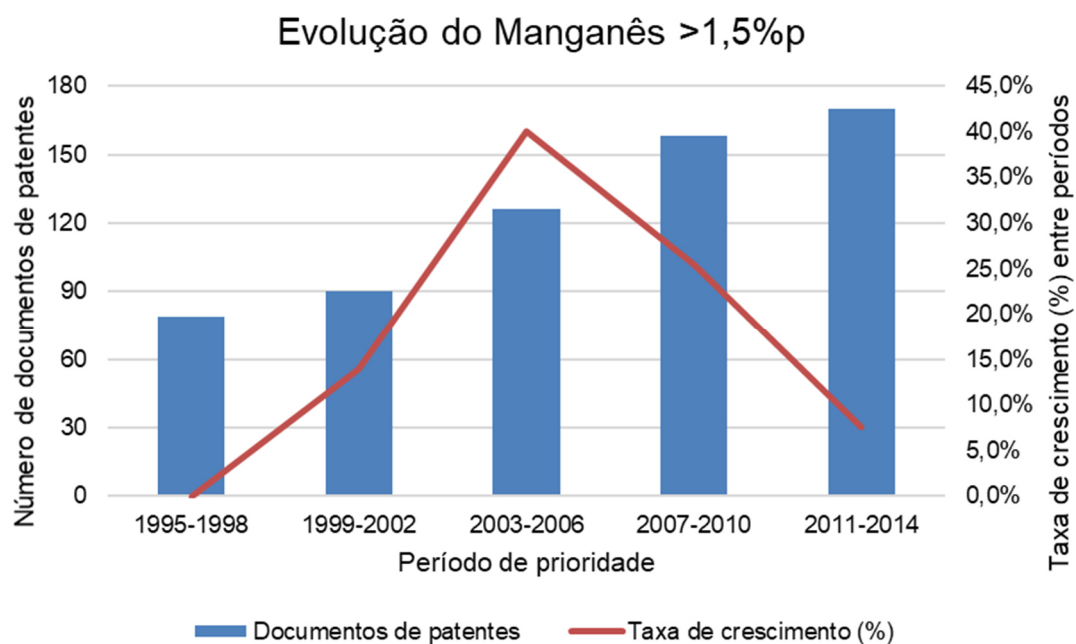
Fonte: [85].

Figura 2.11 – Diagrama resistência em função do alongamento de ligas Fe-C-Mn-Al.

A apresentação de informações integradas de composição química, processamento, microestrutura e propriedades é de grande valia tanto para atividades de seleção de materiais quanto de monitoramento tecnológico. Entretanto, a elaboração de figuras como a colocada acima requer o empenho na busca em diferentes tipos de publicações, e a análise de que os ensaios foram realizados nas mesmas condições para efeito de comparação, por exemplo. No caso de novas tecnologias, como os aços avançados de alta resistência, a oportunidade de consultar um conjunto de dados extraídos de uma fonte de informação de forma semiautomatizada pode contribuir para uma melhor compreensão de sua dinâmica de desenvolvimento.

Um outro mercado que demanda aços de alto valor agregado é o de petróleo, gás natural e seus derivados. Os dutos para transporte de hidrocarbonetos têm fundamental importância na matriz energética de vários países por permitirem o acesso viável de energia em locais distantes de suas principais fontes produtoras [86]. Por outro lado, quanto mais extenso um duto, maior o seu custo de construção, operação e manutenção. Adicionalmente, o transporte por grandes distâncias necessita de maior resistência dos tubos que os compõem, que podem ser obtidas por uma maior espessura, que leva a um maior consumo de material, ou pelo aumento da resistência mecânica do aço que os compõem. Não suficiente o desafio já colocado, os tubos são soldados uns aos outros, o que necessita também de controle da composição química para uma boa qualidade de solda. Os aços de alta resistência e baixa liga (ARBL) e microligados processados termomecanicamente são geralmente empregados para fabricar os tubos que compõem os dutos [87], com sua composição química limite, condições de processamento e propriedades desejadas bem estabelecidas por normas, como a API 5L [88].

Trabalhos anteriores avançaram na elaboração de indicadores para o monitoramento tecnológico desses aços e de outros componentes da indústria petrolífera [55,57]. Entretanto, o nível de detalhamento depende de avanços na recuperação e disponibilização para que os indicadores possam ser ainda mais precisos. A Figura 2.12 apresenta um exemplo de indicador de registros de patentes de aços para aplicação em dutos classificados em subgrupos da CIP com percentual em peso de manganês acima de 1,5%. Durante o período avaliado, houve revisões da norma API com o aumento do limite máximo de manganês nas composições químicas, principalmente para os produtos com maior limite de escoamento [89].



Fonte: [89].

Figura 2.12 – Número de registros de patentes em aços para dutos de transporte de hidrocarbonetos classificados em códigos da CIP com percentual em peso acima de 1,5% de manganês.

Logo, os avanços na mineração de textos são fundamentais para que se possa apresentar, por exemplo, o teor de manganês e de outros elementos de liga para essa aplicação específica ou de outras. Estudos de casos em bases de dados ou softwares de seleção de materiais que aplicassem os limites estabelecidos em normas certamente trariam muitas opções viáveis. Por outro lado, o conhecimento de novas soluções, ainda que não comerciais neste momento, também se faz importante. Assim, espera-se que os dados extraídos a partir de registros bibliográficos de documentos de patentes em desenvolvimentos recentes em aços contribuam para esses avanços.



### 3 MATERIAIS E MÉTODOS

#### 3.1 Recuperação e preparação da amostra de dados

O processo que culminou nos métodos e procedimentos desenvolvidos e descritos nessa tese se iniciou com a leitura e comparação do texto de reivindicações de documentos originais de patentes referentes a aços e de seus registros bibliográficos indexados pela base *Derwent Innovations Index*. Através destas comparações foi possível observar que as informações de interesse para a elaboração da planilha de dados, quando encontradas no documento de patente original, geralmente também são descritas em alguns dos campos dos registros bibliográficos, em especial no resumo e em suas subdivisões de “Novidade”, “Descrição Detalhada”, “Vantagem”, “Uso” e no campo “Foco Tecnológico”.

Durante essa etapa de leitura e comparações dos registros e documentos originais também foi observado que, em alguns dos registros bibliográficos de invenções anteriores ao ano de 2000, a riqueza de detalhes sobre as invenções era menor. Adicionalmente, testes com expressões de buscas para um recorte de tempo inicialmente estimado em 20 anos ultrapassava o número de 50 mil resultados, o que evidenciou dificuldades nas questões de tempo de processamento do software bibliométrico utilizado (VantagePoint 5.0). Esse grande volume de informação também poderia se apresentar como um limitante nos tratamentos posteriores para a elaboração da planilha de dados, como limpeza de registros fora do escopo do estudo e o tratamento de registros que não apresentam tal riqueza de detalhes. Dessa maneira, o recorte de tempo colocado para a elaboração dessa planilha de dados inicial foi de 10 anos, com os registros dos documentos de patentes com ano de primeiro depósito entre 2009 e 2018.

Assim, para um estudo de caso que envolvesse desenvolvimentos recentes de aços, que em si é uma classe de materiais que pode ser subdividida em diferentes maneiras, optou-se por uma expressão de busca mais simples possível. Dessa maneira, atualizações posteriores como o aumento do escopo temporal, partirão de um menor esforço na etapa de busca. Adicionalmente, estudos futuros em aços poderão realizar análises diretamente em uma interface

já pronta, que contempla um expressivo volume de informação sobre aços, sem a necessidade de realizar buscas e coletas diferentes de registros para cada escopo que desejam avaliar.

A expressão de busca utilizada neste trabalho recuperou registros de patentes indexados na DII que contivessem o termo “*steel*” em seus títulos reescritos pela base e documentos que foram classificados no código C22C-038 da CIP, que é descrito por “Ligas ferrosas, p. ex. ligas de aço”, porém que não contempla em si os ferros fundidos (C22C-037) e ligas à base de outros elementos metálicos. Os procedimentos e indicadores elaborados nesta tese se referem à última busca e seus 33.293 resultados de registros de patentes apresentados e recuperados no dia 10/02/2020. Além dos termos da expressão de busca, a única restrição imposta foram os anos de indexação na base entre 2009 e a data de busca. A seguir, os registros recuperados foram tratados com o *Earliest Priority Selector* (EPS) [90], desenvolvido pelo NIT/Materiais. A razão para tal etapa se justifica que os registros importados originalmente pela DII contemplam as famílias de patentes, ou seja, os depósitos realizados em diferentes países e suas datas. Para estudos de monitoramento tecnológico que envolvem informação contida em documentos de patentes é recomendado que sejam apresentados a data do primeiro depósito da patente, o que mais se aproxima do real momento de desenvolvimento da tecnologia [91,92]. Nesse caso, os registros recuperados na DII necessitam de um pré-tratamento antes da leitura pelo VantagePoint, como a busca e identificação da data e país de primeiro depósito, realizado pelo EPS.

Todos os registros foram importados no software VantagePoint (v5.0) em um filtro<sup>1</sup> previamente já desenvolvido pelo NIT/Materiais para leitura dos registros da DII e criação de listas dos principais campos como títulos, resumos, códigos da CIP, entre outros. A seguir, os códigos da CIP com maior nível de detalhe (9 dígitos) foram utilizados para a exclusão de registros não relacionados diretamente a proposta de inovação em aços como, por exemplo, equipamentos

---

<sup>1</sup> Nesta tese são apresentados dois conceitos distintos para a palavra “filtro”. No software VantagePoint, os filtros consistem em um conjunto de comandos que permitem selecionar ou modificar um trecho de texto, com sua posterior apresentação e quantificação. No MS Excel, a operação de filtrar dados corresponde a exibir apenas os itens das linhas da planilha que correspondem aos critérios escolhidos.



e dispositivos de controle utilizados na cadeia siderúrgica. Essa etapa foi realizada a fim de se obter um conjunto de dados com foco predominante em novas ligas, processos que alteram a microestrutura dos materiais, assim como as propriedades resultantes e aplicações requeridas dessas propostas de inovação. A Tabela 3.1 apresenta detalhes das sequências de etapas de limpeza iniciais de registros para a definição da amostra de dados avaliada nesta tese.

Tabela 3.1 – Procedimento de limpeza de registros para a definição da amostra de registros de patentes em aços entre os anos de 2009 e 2018.

Etapa	Registros após etapa	Descrição da ação realizada
1	33.293	Resultados da expressão de busca original. "TI=steel* AND IP=C22C-038*"
2	27.591	Limitação pelo ano de primeiro depósito entre 2009 e 2018.
3	25.999	Eliminação dos registros classificados no código CIP C23C-002: "Processo de imersão a quente ou a frio para aplicar um material de revestimento em estado fundido sem modificar a forma do objeto; Aparelhos para esse fim."
4	25.199	Eliminação dos registros classificados no código B21B-037: "Dispositivos de controle ou métodos especialmente adaptados para laminadores ou o trabalho produzido pelos mesmos (métodos ou dispositivos para medir, especialmente adaptados para laminadores.)"
5	24.509	Eliminação dos registros classificados nos códigos C22C-001 a C22C-029 e dos códigos C22C-032 e C22C-037, referentes a ligas à base de metais não ferrosos e ferros fundidos.
6	23.892	Eliminação dos registros classificados nos códigos B23K-001 a B23K-037 (com exceção de B23K-035), B23K-101 e B23K-103/06, referentes a processos e dispositivos de soldagem, e a ligas à base de metais não ferrosos e ferros fundidos a serem soldados.
7	23.321	Eliminação dos registros classificados nos códigos B22D-011/08 A B22D-011/22, referentes a acessórios, dispositivos e métodos de controle de lingotamento.
8	23.053	Eliminação dos registros classificados nos códigos C21D-009 /54 a C21D-009/70, referentes a fornos para tratamentos térmicos.

Fonte: Elaboração do autor.

Apesar de uma eliminação de mais de 4 mil registros no período de interesse (2009 e a 2018) apenas por códigos da CIP, essas etapas também excluíram alguns registros que poderiam conter informações relevantes, embora classificados em códigos em que a maioria dos registros não era adequado ao escopo previsto. Por outro lado, alguns registros que também não estão diretamente relacionados ao desenvolvimento recentes de ações não foram eliminados por esses processos, o que requer o cuidado e atenção dos analistas que venham a replicar este método e utilizar a planilha de dados. Essas duas limitações podem ser trabalhadas em atualizações futuras, de acordo também com o nível de precisão ou com a abrangência esperada de seus desenvolvedores, com a opção de expandir ou restringir sua abrangência. Ao longo das etapas posteriores para a elaboração da planilha de dados, também foram encontrados registros não condizentes com o escopo esperado, que foram eliminados através da atualização dos filtros e de tesouros, detalhados nos itens subsequentes.

### **3.2 Desenvolvimento de filtros em software bibliométrico**

O software VantagePoint disponibiliza uma ferramenta para desenvolvimento e edição de filtros que permitem recuperar trechos de textos de acordo com uma programação específica, a partir de cada registro bibliográfico importado. Pelo fato de a DII apresentar um registro de campos estruturados que, por sua vez, contêm frases elaboradas a partir dos textos dos documentos de patentes originais, o processo de recuperação das informações de interesse partiu da análise de quais partes dos registros disponibilizam essas informações. Para cada dado de interesse para elaboração da planilha, foram utilizados dois ou três campos dos registros em que melhor foram encontrados esses dados. Por exemplo, a composição química é geralmente expressada no campo “Novidade”, enquanto as propriedades mecânicas são mais facilmente encontradas no campo “Vantagem”.

Na primeira versão de planilha de dados elaborada nesta tese, foram buscados dados sobre a composição química do material e de algumas propriedades, além da microestrutura, processos de produção e aplicações

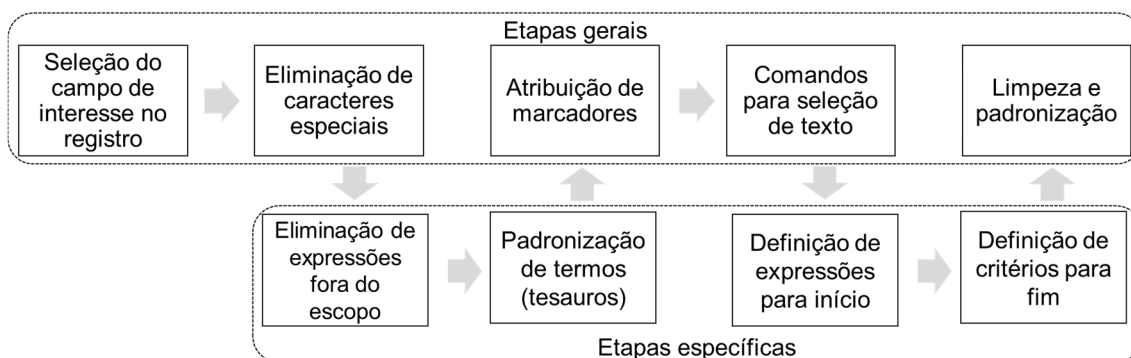
requeridas ou em potencial. Os dados sobre os elementos de liga foram buscados com os teores percentuais em peso de carbono, alumínio, manganês, silício, cromo, nióbio, vanádio e titânio. Esses elementos foram escolhidos por demonstrarem o potencial de uso da ferramenta para atividades de seleção de materiais e de indicadores para o monitoramento tecnológico nos exemplos avaliados nos resultados e discussões, como os aços de baixa densidade e aços para dutos de transporte de hidrocarbonetos.

O desenvolvimento de cada novo filtro partiu do estudo dos já previamente desenvolvidos pelo NIT/Materiais para a apresentação de listas das frases contidas em cada campo do resumo fornecido pela DII. Como exemplo, a partir do campo “Novidade” e do seu filtro que já apresentava os textos dessa divisão do registro, foram trabalhadas diferentes estratégias para recuperar os dados de composição química da liga, ou de sua microestrutura, processos, propriedades e aplicações. Para tal, foi preciso avaliar como as informações de interesse são comumente apresentadas nesses campos, em busca de padrões. No caso de composição química e das propriedades físicas e mecânicas, o dado de interesse se apresenta na forma de caracteres numéricos como os teores percentuais em peso de cada elemento na composição da liga ou o limite de escoamento do material em mega pascal, por exemplo. Por outro lado, as informações de microestrutura, processos e aplicações são apresentadas por caracteres alfabéticos ou códigos da CIP que os descrevem. Além do resumo e suas subdivisões, consideradas como campos específicos (novidade, descrição detalhada, uso e vantagem), foram também desenvolvidos filtros a partir dos campos título, foco tecnológico e de códigos da CIP com 9 dígitos.

A análise de como essas informações são apresentadas nos campos dos registros permitiu observar que nem todas as frases contidas nesses campos são úteis para a recuperação dos dados e textos de interesse. Logo, se tornou necessário a utilização de marcadores para segmentar os trechos das frases para limitar a leitura do campo e facilitar a extração de seus segmentos úteis, a exemplo de estudos anteriores em textos dos documentos completos [8,67]. No caso deste trabalho, os marcadores foram utilizados para destacar os segmentos de textos precedentes à apresentação da composição química da liga, os

próprios elementos químicos para busca de seus teores, e para as palavras que descrevem as propriedades, processos, microestruturas e aplicações buscadas.

O procedimento geral de elaboração dos filtros para recuperação das informações de interesse na elaboração da base de dados neste estudo foi baseado em 5 etapas principais: 1) Leitura do campo de interesse; 2) Limpeza de caracteres especiais e expressões indesejadas; 3) Atribuição de marcadores; 4) Seleção do fragmento de texto de interesse; 5) Limpeza de caracteres adicionais e padronização para apresentação final. A Figura 3.1 apresenta um fluxograma dessa sequência de etapas, intercaladas por etapas específicas que dependem do tipo de caractere a ser buscado: numéricos, para os casos de composição química e propriedades, e alfabéticos, para processos, microestruturas e aplicações.



Fonte: Elaboração do autor.

Figura 3.1 – Fluxograma geral de elaboração de filtros utilizado na pesquisa.

Após a limitação do campo de interesse, uma limpeza prévia de caracteres foi necessária para eliminar dos registros qualquer possível marcador utilizado para seleção de textos em etapas posteriores. No caso dos procedimentos desenvolvidos nesta tese, foram escolhidos os caracteres “@”, “#” e “&”, o que tornou necessário sua eliminação dos textos antes de qualquer tratamento a partir de um comando de “encontrar e substituir” do editor de filtros do VantagePoint. O marcador “@@” foi utilizado para substituir palavras e expressões que comumente precedem a lista de elementos químicos e seus teores, como os termos “*comprising*”, “*including*”, “*containing*” etc. A Lista de

todos os marcadores utilizados para tal fim na versão desenvolvida nesta tese se encontram no Apêndice A (Quadro A.1). Adicionalmente, o padrão “# Símbolo do elemento químico ##” foi utilizado para encontrar os elementos de liga, e sua posterior busca de informações numéricas de teor em peso no material. Como exemplo, o termo “*manganese*” ou o símbolo atômico “Mn” foram padronizados em # Mn ##. A Lista dos marcadores de elementos químicos utilizados na versão desenvolvida nesta tese também se encontra no Apêndice A (Quadro A.2). A Tabela 3.2 apresenta um trecho do campo “Novidade” do mesmo registro apresentado na Figura 2.7 e suas modificações após a substituição dos termos de interesse por marcadores.

Tabela 3.2 – Exemplo de substituição de termos por marcadores em um texto do campo “Novidade” de um registro de patente.

Etapa	Texto
Texto original	<p><i>A hot-rolled sheet steel comprises greater than 0.1 %mass to 0.16 %mass or less carbon, 0.1-0.5 %mass silicon, 1.6-2.5 %mass manganese, 0.02 %mass or less phosphorus, 0.005 %mass or less sulfur, 0.01-0.07 %mass aluminum, greater than 0.5 %mass to 1.5 %mass or less chromium, 0.1-0.5 %mass copper, 0.1-0.3 %mass nickel, 0.1-0.3 %mass molybdenum, 0.01-0.05 %mass niobium, 0.01-0.1 %mass vanadium, 0.005-0.05 %mass titanium, 0.005 %mass or less nitrogen, and remaining iron and unavoidable impurities.</i></p>
Após a substituição por marcadores	<p><i>A hot-rolled sheet steel @@ greater than 0.1 %mass to 0.16 %mass or less # CC ##, 0.1-0.5 %mass # Si ##, 1.6-2.5 %mass # Mn ##, 0.02 %mass or less # PP ##, 0.005 %mass or less # SS ##, 0.01-0.07 %mass # Al ##, greater than 0.5 %mass to 1.5 %mass or less # Cr ##, 0.1-0.5 %mass # Cu ##, 0.1-0.3 %mass # Ni ##, 0.1-0.3 %mass # Mo ##, 0.01-0.05 %mass # Nb ##, 0.01-0.1 %mass # VV ##, 0.005-0.05 %mass # Ti ##, 0.005 %mass or less # Ni ##, and remaining iron and unavoidable impurities.</i></p>

Fonte: Elaboração do autor.

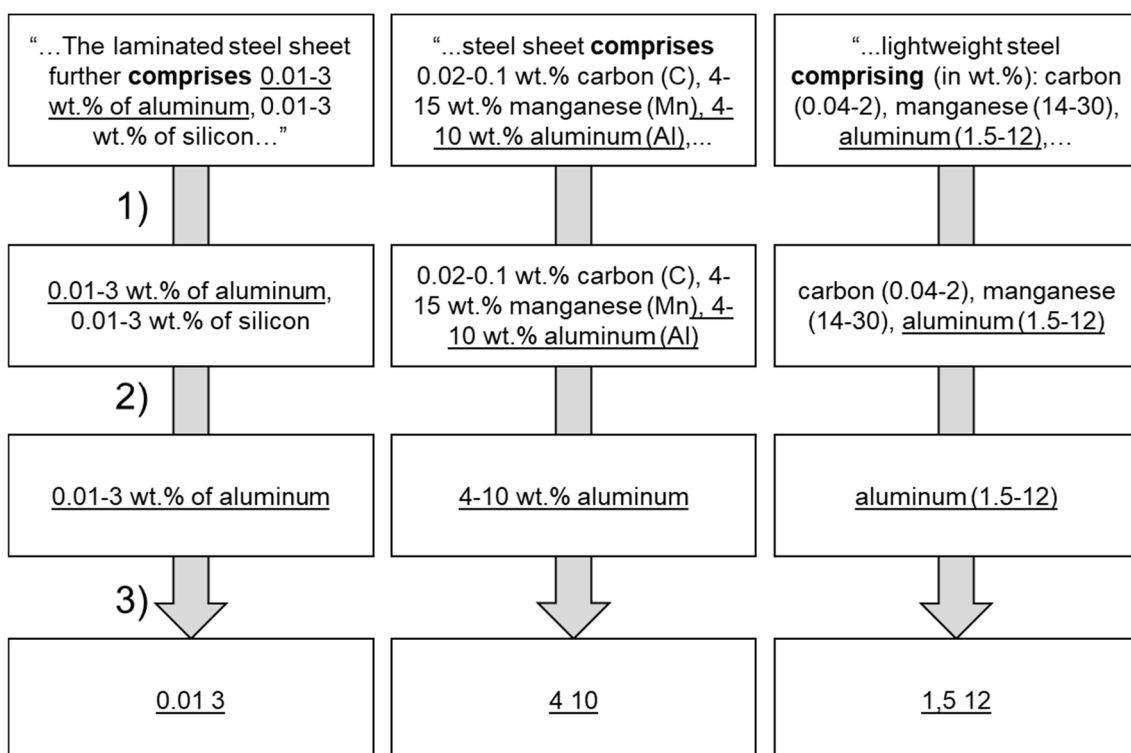
Entretanto, principalmente no trabalho de recuperação de teores de elementos de liga, diferentes fragmentos de texto tiveram que ser excluídos para

que a informação de interesse não fosse recuperada de forma incorreta. Por exemplo, os algarismos próximos à citação de óxidos, sulfetos, carbetos etc. de elementos metálicos não retornariam um valor numérico condizente com o teor de elementos de liga típicos de aços. Adicionalmente, informações de temperatura acompanhadas pela expressão “*degrees C*”, poderiam retornar valores incompatíveis com o teor de carbono esperado em aços, assim como valores de carbono equivalente, por exemplo. Por fim, situações semelhantes ocorrem com os marcadores que separam um trecho de texto anterior daquele que inicia a lista dos elementos de liga e seus teores. Como exemplo, a expressão “*impurities containing*” poderia iniciar a leitura de componentes de impurezas dos aços, e não os elementos planejados como parte da liga. Assim, a partir dessa exclusão prévia (etapa 2), a seleção dos textos de interesse (etapa 3) é facilitada e os procedimentos adotados para tal fim podem retornar resultados condizentes com o esperado para um aço.

Ao longo do desenvolvimento da planilha de dados, os diversos testes na tentativa de encontrar a melhor maneira de recuperar os dados resultaram na elaboração de tesouros para a eliminação dessas expressões antes da atribuição de marcadores. Os principais temas encontrados foram descrições de informações sobre revestimentos de proteção anticorrosiva, partes de material compósito não relacionadas ao aço, impurezas, composição química de escórias, de aditivos e de ferroligas utilizados no processo siderúrgico. Adicionalmente, também foram elaborados tesouros para expressões a serem eliminadas para cada elemento de liga apresentado, como de óxidos, carbetos, sulfetos, entre outros compostos, além de expressões que representam a relação entre elementos de liga, como carbono equivalente ou resultado da divisão entre teores de titânio e nitrogênio, por exemplo. Os tesouros referentes a lista de expressões excluídas na versão inicial dos filtros desenvolvidos nesta tese se encontram no Quadro A.3 do Apêndice A.

Um desafio adicional encontrado nas etapas de leitura, comparação e recuperação de informações se refere a como os dados e textos de interesse são apresentados de diferentes maneiras nos registros de patentes recuperados. Um dos principais desafios no tratamento de segmentos de textos com

caracteres numéricos, como os teores de elementos de liga ou de valores de propriedades mecânicas, é que o valor de interesse pode estar no trecho de texto anterior ou posterior ao marcador. De maneira mais simples, o segmento de texto de interesse pode estar à esquerda ou à direita dele quando se lê a frase. A Figura 3.2 apresenta exemplos dos teores de alumínio em diferentes registros de documentos sobre aços recuperados na DII e uma sequência de tratamentos necessários envolvidos para padronizar essas informações.



Etapas: 1) Delimitação do texto de interesse da informação; 2) Seleção do trecho específico; 3) Padronização da informação.

Fonte: Elaboração do autor.

Figura 3.2 – Fluxograma simplificado de etapas para a recuperação e padronização dos teores de alumínio em um registro de patente de aço.

Para superar esse tipo de inconveniente, inicialmente dois filtros distintos foram elaborados para a recuperação de dados numéricos, um para a leitura de valores à esquerda do marcador, e um para leituras à direita do marcador, que depois foram unificados em um único filtro. Como é esperado que sejam casos excludentes, ou seja, caso o teor do elemento químico esteja antes de sua

citação não se espera que a mesma informação esteja repetida também à direita, apenas um dos lados deve retornar o dado desejado. Além disso, para uma melhor organização da planilha de dados, o código único de cada registro gerado pela DII (código GA) foi utilizado em todos os filtros como identificador de cada registro e separado da informação recuperada por uma vírgula. Assim, cada resultado retornado pelo filtro desenvolvido para a recuperação de elementos de liga, e para cada campo de interesse, mesmo que não os encontre, apresenta como retorno o seguinte padrão: “Código GA,informação recuperada à esquerda,informação recuperada à direita”. Os detalhes específicos e processos adicionais para cada campo e tipos de segmentos de textos trabalhados para a elaboração da base são apresentados nos subitens seguintes.

### **3.2.1 Composição química**

Após a leitura de diferentes registros e estudo por onde são apresentados os teores de elementos de liga dos aços, os campos dos registros bibliográficos escolhidos para a recuperação das informações sobre a composição química foram “Novidade”, “Descrição Detalhada” e “Foco Tecnológico”. Para cada um desses campos, e para cada elemento de interesse, foi desenvolvido um filtro único, o que totalizou 24 filtros apenas para os 8 elementos químicos apresentados e estudados nesta tese. Em alguns casos, a mesma informação pode ser apresentada em dois ou nos três campos escolhidos, o que não se torna um problema. Por outro lado, para a elaboração da planilha em si, foi necessário estabelecer uma ordem de prioridade entre os campos de acordo com o volume de dados que recuperavam e definida, após testes, em 1- Novidade, 2 – Descrição detalhada e 3- Foco Tecnológico. Assim, caso o primeiro campo apresente a informação de interesse, essa é considerada. Em caso negativo, considera-se o segmento de texto no campo seguinte e, caso não tenha se encontrado valor numérico algum, as células correspondentes da planilha se encontrarão vazias.

De maneira geral, quando disponibilizadas, a composição da liga nos registros é apresentada em porcentagem em peso de cada elemento, além de um valor mínimo e máximo esperado, como os exemplos da Figura 3.1 e Tabela



3.2. Entretanto, em alguns casos, pode ser apresentado apenas um valor nominal único ou os limites máximos ou mínimos de cada elemento, sem a presença de uma faixa de valores possíveis. Para situações que se encontram neste último caso foram criados marcadores que substituem trechos específicos de texto em caracteres definidos como ">" e "<", de maneira a representar expressões como "*not greater than*" e "*less than*" para valores apresentados como conteúdo máximo de um elemento, ou "*greater than*" e "*or more*" para valores mínimos. O tesouro elaborado para essas substituições se encontra no Quadro A.4 do Apêndice A.

Como já mencionado, a informação de interesse pode se encontrar à esquerda ou à direita do marcador do elemento químico específico de cada filtro. Assim, as Tabelas 3.3 e 3.4 apresentam o tratamento desenvolvido e exemplos para ambos os casos, desde a seleção do texto de interesse até a apresentação final dos teores de manganês para o campo Novidade.

Tabela 3.3 – Sequência geral de tratamentos de textos do campo “Novidade” para a recuperação do teor de manganês (à esquerda) em um registro de documento de patente.

Etapa	Descrição da etapa	Texto após a etapa
1	Leitura do campo “Novidade”	<i>X70 pipeline steel comprises 0.025-0.08 wt.% carbon, 0.1-0.4 wt.% silicon, 1.1-1.6 wt.% manganese, less than or equal to 0.02 wt.% phosphorus, less than or equal to 0.003 wt.% sulfur, less than or equal to 0.2 wt.% copper, 0.31-0.5 wt.% nickel, less than or equal to 0.2 wt.% chromium, less than or equal to 0.2 wt.% molybdenum, 0.04-0.08 wt.% niobium, 0.005-0.02 wt.% titanium, 0.001-0.005 wt.% calcium, 0.02-0.05 wt.% aluminum, less than or equal to 0.0005 wt.% boron, 0.001-0.006 wt.% nitrogen, less than or equal to 0.005 wt.% oxygen, and remaining amount of iron and inevitable impurities. The steel has (carbon+manganese/6) value of 0.25-0.33%, and (molybdenum+chromium) value of 0.1-0.4%.</i>
2	Eliminação de caracteres especiais (#, & e @)	Inalterado (neste caso)
3	Eliminação de termos e expressões prévias aos marcadores	<i>X70 pipeline steel comprises 0.025-0.08 wt.% carbon, 0.1-0.4 wt.% silicon, 1.1-1.6 wt.% manganese, less than or equal to 0.02 wt.% phosphorus, less than or equal to 0.003 wt.% sulfur, less than or equal to 0.2 wt.% copper, 0.31-0.5 wt.% nickel, less than or equal to 0.2 wt.% chromium, less than or equal to 0.2 wt.% molybdenum, 0.04-0.08 wt.% niobium, 0.005-0.02 wt.% titanium, 0.001-0.005 wt.% calcium, 0.02-0.05 wt.% aluminum, less than or equal to 0.0005 wt.% boron, 0.001-0.006 wt.% nitrogen, less than or equal to 0.005 wt.% oxygen, and remaining amount of iron and inevitable impurities. The steel has <del>(carbon+manganese/6)</del> value of 0.25-0.33%, and <del>(molybdenum+chromium)</del> value of 0.1-0.4%.</i>

4	Atribuição de marcadores	X70 pipeline steel @@ 0.025-0.08 wt.% # CC ##, 0.1-0.4 wt.% # Si ##, 1.1-1.6 wt.% # Mn ##, < 0.02 wt.% # PP ##, < 0.003 wt.% # SS ##, < to 0.2 wt.% # Cu ##, 0.31-0.5 wt.% # Ni ##, < to 0.2 wt.% # Cr ##, < 0.2 wt.% # Mo ##, 0.04-0.08 wt.% # Nb ##, 0.005-0.02 wt.% # Ti ##, 0.001-0.005 wt.% # Ca ##, 0.02-0.05 wt.% # Al ##, < 0.0005 wt.% # BB ##, 0.001-0.006 wt.% # NN ##, < 0.005 wt.% # OO ##, and remaining amount of iron and inevitable impurities. The @@ (carbon+manganese/6) value of 0.25-0.33%, and (molybdenum+chromium) value of 0.1-0.4%.
5	Seleção de texto de interesse: @@ até elemento (# Mn ##)	@@ 0.025-0.08 wt.% # CC ##, 0.1-0.4 wt.% # Si ##, 1.1-1.6 wt.% # Mn ##
6	Seleção de texto a partir do fim até um marcador (,)	, 1.1-1.6 wt.% # Mn ##
7	Seleção de texto a partir de um valor numérico até um marcador (# Mn ##)	1.1-1.6 wt.%
8	Eliminação de caracteres indesejados (wt.%)	1.1-1.6
9	Padronização da informação (substituir "-"por " ")	1.1 1.6

Fonte: Elaboração do autor. Texto do registro de código GA: 201892684L.

Tabela 3.4 – Sequência geral de tratamentos de textos do campo “Novidade” para a recuperação do teor de manganês (à direita) em um registro de documento de patente.

Etapa	Descrição da etapa	Texto após a etapa
1	Leitura do campo “Novidade”	<i>A high-heat-input structure steel comprises (in %mass) carbon (0.03-0.1), silicon (0.5 or less), manganese (1-2.5), phosphorus (0.04-0.08), sulfur (0.0005-0.004), aluminum (0.003), niobium (0.003-0.04), titanium (0.01-0.08), chromium (1 or less), oxygen (0.003-0.012), nitrogen (0.002-0.01), and remainder of iron and unavoidable impurities. The number density of oxide-containing inclusion containing titanium oxide and/or titanium with particle size of 1 <math>\mu</math> m or less in steel is 300 piece/mm<sup>2</sup> or more.</i>
2	Eliminação de caracteres especiais (#, & e @)	Inalterado (neste caso)
3	Eliminação de termos e expressões prévias aos marcadores	<i>A high-heat-input structure steel comprises (in %mass) carbon (0.03-0.1), silicon (0.5 or less), manganese (1-2.5), phosphorus (0.04-0.08), sulfur (0.0005-0.004), aluminum (0.003), niobium (0.003-0.04), titanium (0.01-0.08), chromium (1 or less), oxygen (0.003-0.012), nitrogen (0.002-0.01), and remainder of iron and unavoidable impurities. The number density of <del>oxide-containing inclusion containing</del> titanium oxide and/or titanium with particle size of 1 <math>\mu</math> m or less in steel is 300 piece/mm<sup>2</sup> or more.</i>
4	Atribuição de marcadores	<i>A high-heat-input structure steel @@ (in %mass) # CC ## (0.03-0.1), # Si ## (0.5 or &lt;), # Mn ## (1-2.5), # PP ## (0.04-0.08), # SS ## (0.0005-0.004), # Al ## (0.003), # Nb ## (0.003-0.04), # Ti ## (0.01-0.08), # Cr ## (1 or less), # OO ## (0.003-0.012), # NN ## (0.002-0.01), and remainder of iron and unavoidable impurities. The number density of titanium oxide and/or titanium with particle size of 1 <math>\mu</math> m or less in steel is 300 piece/mm<sup>2</sup> or more.</i>

5	Seleção de texto de interesse a partir do elemento (# Mn ##) até um marcador (,)	# Mn ## (1-2.5),
6	Seleção de texto a partir de um valor numérico até um marcador (,)	(1-2.5),
7	Eliminação de caracteres indesejados (“ e “)	1-2.5
8	Padronização da informação (substituir “-“por “ “)	1 2.5

Fonte: Elaboração do autor. Texto do registro de código GA: 2013C48913.

Pelo fato de os resumos disponibilizados pela DII se basearem no texto dos documentos originais de patentes que, por sua vez, dependem de como foram escritos por seus inventores, não foi encontrado um padrão de apresentação dos teores de elementos de liga. Assim, diferentes critérios de comandos para início e fim de seleção de textos tiveram que ser adotados, a fim de maximizar a obtenção desses dados e minimizar a verificação ou correção manual de sua apresentação para a montagem da planilha de dados. Esses critérios foram atribuídos na forma de expressões regulares aceitas pelo editor de filtros do VantagePoint mas, de maneira geral podem ser resumidos como descrito nos parágrafos a seguir.

Para os valores que se encontram à esquerda do marcador do elemento químico, a primeira seleção de texto se inicia do marcador de separação de texto “@@” até o elemento de interesse “# Elemento ##”. A seguir, deve-se selecionar o texto no sentido inverso a partir do fim até um marcador que represente uma provável mudança para um segmento de não interesse. Esse critério pode ser o mesmo marcador “@@” no caso do primeiro elemento citado, uma vírgula, a sequência de caracteres “##” que significaria um outro elemento de liga, o início de uma frase (representado como um espaço seguido de uma letra maiúscula ou a sequência de caracteres “and” que representaria um último componente a ser citado.

Para os valores que se encontram à direita do marcador de elemento químico, a primeira seleção de texto se inicia com o próprio marcador # Elemento ## até um critério de parada que represente um segmento de texto que não seja mais de interesse. Como exemplo, o fim da seleção de texto pode ser aplicado pela descrição do teor de um novo elemento (representado por "# "), um novo marcador de separação de texto "@@", pontuações como vírgula, ponto e vírgula, uma nova frase (ponto final seguido de espaço e um caractere maiúsculo) ou uma sequência de caracteres "and".

Em ambos os casos, as etapas seguintes de seleção de texto se iniciam com um caractere numérico ou de um dos marcadores de limitação de quantidade "<" ou ">", que deve ser correspondente ao início do valor de interesse do teor em peso daquele elemento específico. Novamente, diferentes critérios de fim da seleção de textos devem ser aplicados. No caso dos valores à esquerda, como o final da seleção anterior se trata do próprio marcador do elemento, basta não o incluir na etapa seguinte. Já no caso dos valores à direita, os mesmos critérios de seleção da etapa anterior podem ser aplicados, visto representarem também já a mudança para um segmento de texto que provavelmente já não apresentarão os valores esperados.

Em sequência, como também são distintas as formas de apresentação dos valores e diferentes caracteres são utilizados para representar as unidades (como "w%", "wt.%", "weigth%", entre outros) também devem ser eliminados de maneira a padronizar a apresentação para a elaboração da planilha de dados. Dessa maneira, foram adicionadas etapas de limpeza através de tesouros e comandos de encontrar e substituir do editor de filtros do software. Por fim, também foram encontradas diferentes maneiras de separar os valores mínimos e máximos esperados para os teores dos elementos de liga (como "to" ou "-") e, através dos mesmos procedimentos de encontrar e substituir, esses caracteres foram substituídos por um espaço. Dessa maneira, como detalhado no item 3.3 e no Apêndice C, foi possível organizar as informações por colunas e relacionar as informações de todos os elementos químicos buscados com as outras informações de microestrutura, processamento, propriedades e aplicações.

Através desses critérios utilizados, acredita-se ter coberto a grande maioria das formas de apresentação da composição química nos registros de patentes em aços da DII, embora exceções possam ainda ser encontradas. Adicionalmente, diferentes propostas de etapas de seleção e limpeza de textos podem, futuramente, vir a aumentar o número de registros com tais informações recuperadas, bem como se tornarem mais simples para uma diminuição no tempo de processamento do software VantagePoint na geração das listas que alimentarão as planilhas no MS Excel.

### **3.2.2 Propriedades físicas e mecânicas**

Por também apresentarem valores numéricos, os filtros para a busca de valores para as propriedades físicas e mecânicas são semelhantes às buscas dos teores de elementos de liga. Entretanto, dois desafios adicionais foram apresentados para a recuperação e padronização desses segmentos de textos. O primeiro se refere a definição de expressões regulares que descrevem uma propriedade específica, que pode possuir sinônimos e siglas que representam incontáveis maneiras de serem apresentadas no texto dos registros bibliográficos. Já o segundo desafio se refere às unidades e suas variedades de forma de apresentação. Neste trabalho foram desenvolvidos filtros para a recuperação de dados de densidade, limite de escoamento, limite de resistência à tração, módulo de elasticidade, alongamento percentual e dureza (na escala Rockwell C).

Em alguns casos, como a recuperação de limites de resistência à tração que são geralmente expressos em Mega Pascal no caso dos aços, foi necessário criar um tesouro específico para a conversão de trechos de texto similares como “megapascals”, “MPa”, “mega pascal” entre outros fossem padronizados. Para a análise de casos específicos, como aços para dutos de transporte de hidrocarbonetos, também foram criados tesouros para conversão de siglas previstas pela norma API5L que descrevem o tipo de produto pelo seu limite de escoamento em Ksi, para os valores correspondentes em MPa. Os tesouros desenvolvidos para a padronização de unidades se encontram no Apêndice B (Quadro B.2) enquanto a Tabela 3.5 apresenta as expressões regulares e

unidades utilizadas que foram utilizadas como critérios de início ou fim de seleção de textos para cada propriedade. Foram escolhidas as unidades e escalas apresentadas para permitir comparação entre os valores recuperados, embora propriedades como a dureza, por exemplo, podem ser apresentadas em diversas escalas, que podem ser abordadas em atualizações futuras dos filtros elaborados.

Tabela 3.5 – Expressões utilizadas como marcadores e unidades de propriedades físicas e mecânicas.

Propriedade avaliada	Expressões regulares para delimitação de seleção de textos	Unidades e expressões regulares de fim de seleção de textos
Densidade	<i>“density”</i> ; <i>“specific gravity”</i> ; <i>“specific mass”</i>	“g/cm <sup>3</sup> ” e “kg/dm <sup>3</sup> ”
Limite de escoamento	<i>“yield strength”</i> ; “YS”	MPa
Limite de resistência à tração	<i>“tensile strength”</i> ; “UTS”	MPa
Módulo de elasticidade	<i>“Young’s modulus”</i> ; <i>“young modulus”</i> ; <i>“youngs modulus”</i> ; <i>“elastic modulus”</i> ;	GPa
Alongamento percentual	<i>elongation</i>	%
Dureza (Rockwell C)	HRC; Rockwell C	HRC

Fonte: Elaboração do autor.

A mesma lógica de busca de valores numéricos à esquerda ou à direita dos marcadores foi aplicada também para a busca dos valores das propriedades. O mesmo procedimento também foi utilizado nas etapas de limpeza e padronização dos dados para apresentação de um valor nominal ou mínimo em uma coluna e máximo, quando houver, em uma segunda coluna da planilha. Os campos escolhidos para a busca e recuperação dessas informações foram, em ordem de prioridade de apresentação: 1- Vantagem e 2- Novidade, visto que, em testes realizados, estes apresentavam um maior volume de segmentos de textos recuperados com as programações de filtros desenvolvidas. A Tabela 3.6 apresenta, em resumo, uma sequência de tratamentos realizados para a recuperação do limite de escoamento de um aço a partir do campo Vantagem.



Tabela 3.6 – Sequência geral de tratamentos de textos do campo “Vantagem” para a recuperação do Limite de escoamento (à direita) em um registro de documento de patente.

Etapa	Descrição da etapa	Texto após a etapa
1	Leitura do campo Vantagem	<i>The electric main shaft has excellent corrosion resistance, impact absorbing power of more than 45J, yield strength of 650-680 MPa, tensile strength of 810-860 MPa and elongation rate of 12-16%</i>
2	Eliminação de termos e expressões prévias aos marcadores	Inalterado (neste caso)
3	Leitura a partir de uma expressão regular ( <i>yield strength</i> )	<i>yield strength of 650-680 MPa, tensile strength of 810-860 MPa and elongation rate of 12-16%</i>
4	Seleção de texto de interesse desde um caractere numérico até um marcador “MPa”	650-680
5	Padronização da informação (substituir “-” por “ ”)	650 680

Fonte: Elaboração do autor. Texto do registro de código GA: 2012K69143.

Etapas prévias de limpeza de expressões no texto também foram aplicadas de maneira a evitar a recuperação de valores incompatíveis com o real desempenho da proposta de inovação e descrita em seu registro bibliográfico de patente. Por exemplo, uma seção de texto que descreve “*increase the elongation in 20%*”, se aplicada apenas com os conceitos de seleção de texto a partir de “*elongation*” e um valor numérico à esquerda até o critério de parada de encontrar o caractere “%”, poderia retornar o valor que seria interpretado como um alongamento percentual de 20%, entretanto como um falso indício que esse seria o valor encontrado no teste de tração daquele aço com composição química, processamento e microestrutura recuperados no mesmo registro. Dessa maneira, além de um trabalho adicional na criação de tesauros para a versão atual dos filtros, este será um item de atenção contínua para suas futuras

atualizações. Assim, também os analistas de dados devem ser críticos quanto a valores e informações apresentadas antes de relacioná-las em processos de tomada de decisão em seleção de materiais ou análise de indicadores tecnológicos. A atual lista de expressões utilizadas para a exclusão de expressões no texto do registro antes da atribuição de marcadores das propriedades se encontra no Quadro B.1 do Apêndice B.

### 3.2.3 Microestrutura, processos e aplicações

Para a recuperação de informações de interesse em que sua apresentação não é dada na forma numérica, mas sim por palavras ou expressões encontradas de diferentes maneiras em um trecho livre de texto, uma etapa de padronização desses termos é necessária. Como exemplo, podemos deduzir que um aço apresenta a microestrutura perlítica ao ler uma frase contendo a expressão "*ferrite type steel*" ou "*ferritic steel*". Adicionalmente, a primeira expressão também pode estar apresentada em outros registros como "*ferrite-type steel*". Dessa maneira, apesar de não ser necessária a utilização de marcadores para selecionar trechos de textos antes de recuperar informações sobre microestrutura, processos e aplicações quando descritas por palavras em um campo do registro bibliográfico, a principal e uma das primeiras etapas trata de uma padronização desses termos. Novamente, a elaboração de tesouros foi necessária para substituição de termos sinônimos para um mesmo padrão definido.

No caso de processos e aplicações, os códigos da Classificação Internacional de Patentes também se apresentam como fonte de informação a respeito do invento registrado. Entretanto, para a apresentação na base de dados, eles também foram descritos em expressões resumidas, que representam a descrição do código e permitem a comparação na mesma planilha sem a necessidade de recorrer à CIP em todo momento. Tanto para as expressões no texto corrido quanto para os códigos da CIP, foi necessária a utilização de marcadores para a posterior seleção de texto e apresentação. No exemplo dos aços ferríticos, por exemplo, as expressões apresentadas no parágrafo anterior e outras semelhantes foram substituídas pelo termo

“@Ferrita#”. Assim, como a microestrutura de um aço pode apresentar mais de uma fase ou microconstituente, ou o trecho de texto descreva uma transformação que ocorra em um determinado processo, todos os termos iniciados com “@” e finalizados com “#” representariam uma dessas possibilidades.

Logo, de maneira mais resumida, os filtros para a recuperação de fases, processos e aplicações foram desenvolvidos para a leitura do campo de interesse, a padronização de termos pelos tesouros apresentados nos Quadros B.3 a B.7 do Apêndice B<sup>1</sup> e pela posterior leitura dos termos padronizados e identificados pelos marcadores “@” e “#”. Para a identificação da microestrutura foram utilizados os seguintes campos e ordem de prioridade: 1- Título, 2- Novidade e 3- Descrição Detalhada. Para os processos foram utilizados: 1-CIP, 2- Novidade e 3- Descrição detalhada. E, finalmente, para a aplicação foram realizadas buscas de termos e códigos nos campos 1- Uso, 2-Título e 3-CIP. A Tabela 3.7 apresenta, de forma resumida, a sequência de tratamentos para a recuperação de possíveis aplicações de um registro de patente de aço a partir do campo Uso.

Tabela 3.7 – Sequência geral de tratamentos de textos do campo “Uso” para a recuperação de possíveis aplicações em um registro de documento de patente.

Etapa	Descrição da etapa	Texto após a etapa
1	Leitura do campo “Uso”	<i>Steel plate is used for component for pressure vessel and manufacturing pipeline (all claimed) for transporting gas and crude oil, and storage tank.</i>
2	Aplicação de <i>thesaurus</i> de termos para aplicações	<i>Steel plate is used for component for @vasos de pressão# and manufacturing @dutos# (all claimed) for transporting gas and crude oil, and @tanque de armazenagem#.</i>
3	Leitura de termos entre marcadores	<b>@vasos de pressão# @dutos# @tanque de armazenagem#</b>
4	Divisão do texto e separação por vírgulas	vasos de pressão,dutos,tanque de armazenagem

Fonte: Elaboração do autor. Texto do registro de código GA: 2014S06659.

<sup>1</sup> As denominações de processos e aplicações se encontram sem acentuação e sem caracteres como “ç” em vista de incompatibilidade com a versão do VantagePoint utilizada para leitura e posterior transferência dos dados para o MS Excel.

Na atual versão da planilha de dados foram limitados a apresentação de até 3 processos, fases ou microconstituintes e aplicações por registro, em cada um dos campos buscados. Entretanto, diferente do esperado para valores numéricos de propriedades e teores de elementos de liga, é possível que uma mesma aplicação ou processo tenham sido identificados em dois campos distintos, o que geraria duplicatas na apresentação da base de dados. Assim, com o auxílio do MS Excel, foram elaboradas fórmulas para a eliminação de termos iguais dentro de um mesmo campo do registro. Dessa maneira, optou-se também pela apresentação de 1 a 3 processos, fases ou microconstituintes e aplicações, de acordo com as informações recuperadas em todos os três campos utilizados, respeitadas as ordens de prioridade citadas no parágrafo anterior. Os detalhes das fórmulas podem ser consultados no Apêndice C.

### **3.3 Organização da planilha de dados e elaboração dos indicadores**

Cada filtro elaborado no VantagePoint gera uma lista das informações recuperadas que podem ser exportadas para planilhas do MS Excel. Deve-se atentar, inicialmente, para o idioma do editor de planilhas e sua configuração de separadores de decimais e milhar. Como a DII apresenta os resumos em inglês, o separador de decimal, mais comum nas informações trabalhadas é definido por ponto. Cada tipo de informação recuperada foi copiado e colado para uma “planilha de apoio” para a sequência de tratamentos e verificações dos resultados recuperados. Assim, foram elaboradas planilhas denominadas “carbono”, “manganês” etc. para os elementos de liga, uma para cada uma das seis propriedades apresentadas, além das que recuperaram informações de processos, microestrutura e aplicações.

No caso de informações compostas por caracteres numéricos, a lógica geral de todos os filtros desenvolvidos no VantagePoint retorna o padrão “Código GA,informação 1,informação 2”. Os valores numéricos, quando apresentam valores limites de máximo e mínimo são separados por um espaço. No caso de segmentos de textos compostos por palavras, o padrão se torna “Código GA,informação 1,informação 2, ..., última informação recuperada”. A separação por vírgulas e espaços foi adotada para facilitar a estruturação da planilha por

colunas, através do recurso “texto para colunas” do MS Excel. Quando a informação buscada não foi encontrada, o filtro apenas retornou o código GA do registro avaliado, seguido de duas vírgulas o que, nas planilhas, representam um valor vazio.

A substituição dos caracteres “<” e “>” foi realizada também com o auxílio do software MS Excel. Uma combinação de fórmulas foi necessária para casos em que a ordem dos caracteres não estaria adequada para a apresentação na planilha de dados. Como exemplo, um segmento de texto como “*silicon (0.5 or less)*”, após os tratamentos do filtro do elemento silício seria apresentado como “Código GA,,0.5 <”. Entretanto, ao organizar a planilha por colunas: 1- código GA, 2- valor mínimo ou nominal e 3- valor máximo, o valor 0.5 se encontraria na coluna 2, quando deveria estar na coluna 3, representando o valor máximo do teor desse elemento. Para esse caso e semelhantes, como a situação inversa em que o valor numérico se refere ao mínimo em peso daquele elemento, foi necessária uma troca de colunas em uma planilha auxiliar denominada “Ajustador de limites”, com seus detalhes operacionais também apresentados no Apêndice C.

Em cada uma das planilhas de apoio, foram utilizadas fórmulas para a busca da informação de interesse na ordem de prioridade dos campos em que foram buscadas, através de recursos de definição de células vazias e de condicionais pelo uso das fórmulas condicionais do tipo “se”. No caso da composição de um elemento, por exemplo, caso as colunas referentes ao código GA no campo “Novidade” estivessem vazias, seriam consideradas as do campo “Descrição detalhada”. Em caso de novamente vazias, seriam lidas as do campo “Foco Tecnológico”. E, se todas as células se encontrassem vazias, os valores apresentados na planilha principal também se encontrariam como células sem conteúdo. Após a importação de toda a informação recuperada para o MS Excel, foi criada uma planilha principal, denominada “NIT SPD”, que concatena todas as informações de cada um dos códigos GA. As fórmulas elaboradas buscam o código GA nas planilhas de apoio e retornam o dado específico na coluna correspondente das planilhas de apoio.

Durante a organização das planilhas de apoio e após a concatenação dos dados na planilha principal, notaram-se dados que foram apresentados e se mostraram incoerentes para os aços. Como exemplo, diversos registros apresentam teor de carbono acima de 2,14% em peso, ou uma quantidade muito superior à esperada de um outro elemento de liga. Também foram listadas propriedades muito abaixo ou muito acima do que também seria esperado para o desempenho de aços. Ao realizar uma verificação manual com a leitura dos títulos e resumos de registros desses casos, foram encontradas propostas de invenção que não se relacionavam diretamente a uma liga, processo, aplicação ou desenvolvimento relacionado aos aços. Assim, o termo “*steel*” ou o código CIP “C22C-038” podem ter sido atribuídos em um contexto em que o aço é utilizado, mas não o principal tema do pedido de patente. As linhas da planilha foram excluídas manualmente ao longo do desenvolvimento da versão atual, e um tesouro com seus códigos GA foi elaborado para que, em atualizações futuras, esses mesmos registros não sejam apresentados novamente. Assim, a versão atual da base de dados contempla 22.911 registros apresentados.

A verificação manual também permitiu encontrar outros registros em que, embora os valores inconsistentes, foram mantidos pois podem ser úteis a futuros estudos de acordo com a análise desejada. A Tabela 3.8 apresenta o código de cores atribuído nas células referentes ao código GA de cada registro.

Tabela 3.8 – Código de cores atribuído aos registros após verificação manual dos dados recuperados.

Cor	Descrição	Exemplos
Verde	Registro verificado manualmente e valores condizentes com o texto do registro.	
Amarelo	Valores inconsistentes com o esperado para aços	Registro 2014D48763 com limite de escoamento de 3690 MPa e 319% de alongamento
Laranja	Registro referente a aços, porém os dados originalmente recuperados não se referiam à liga em si.	Registro 201717172D em que a dureza de 61HRC se refere ao compósito em que o aço é um dos componentes.

Azul	Propriedades referentes a condições de ensaio ou operações específicas.	Registro 2015399853 em que o limite de escoamento de 40 a 70 MPa se refere a uma espessura de 12 a 120 milímetros.
Vermelho	Registros em que o valor mínimo e máximo de algum elemento ou propriedade está invertido ou reescrito de maneira incorreta.	Registro 2011L86131 com teor de carbono apresentado: 0,03 e 0,02.

Fonte: Elaboração do autor.

Por fim, a elaboração dos indicadores apresentados e discutidos nesta tese foram possíveis a partir de tratamentos disponibilizados pelo MS Excel, como filtro<sup>1</sup> de valores e tabelas dinâmicas. Outros indicadores podem ser elaborados de acordo com as necessidades e nível de proficiência no software. Alguns recursos e instruções são disponibilizados no Apêndice D. Adicionalmente, outros editores de planilhas também podem ser utilizados em futuras atualizações de planilhas de dados semelhantes, desde que compatíveis com o modo que o VantagePoint exporta suas listas, e com adaptações feitas para tal.

---

<sup>1</sup> Nesta tese são apresentados dois conceitos distintos para a palavra “filtro”. No software VantagePoint, os filtros consistem em um conjunto de comandos que permitem selecionar ou modificar um trecho de texto, com sua posterior apresentação e quantificação. No MS Excel, a operação de filtrar dados corresponde a exibir apenas os itens das linhas da planilha que correspondem aos critérios escolhidos.

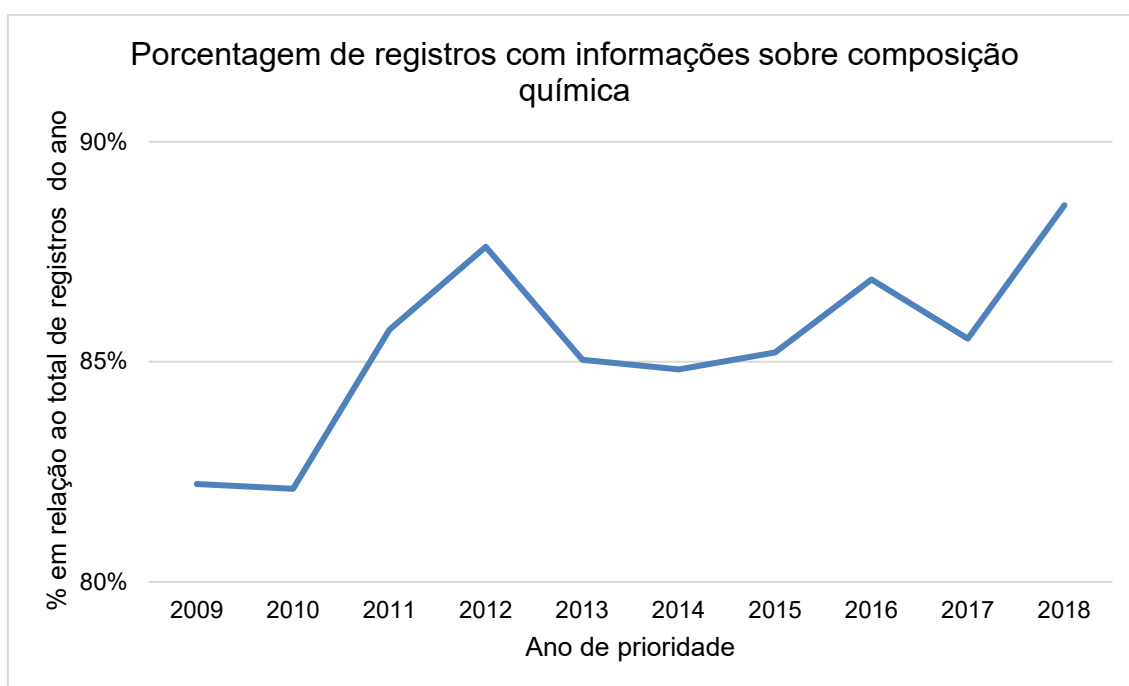




## 4 RESULTADOS E DISCUSSÃO

### 4.1 Avaliação dos dados recuperados e indicadores gerais do patenteamento recente em aços

Dos 22.911 registros bibliográficos de patentes que compõem a planilha avaliada nesta tese foi possível recuperar informações sobre os oito elementos de liga avaliados em 19.627 (85,6%) deles. Por se tratar de segmentos de textos baseados nos textos completos das reivindicações dos documentos originais, o nível de detalhes apresentados ao ponto de conhecer os valores mínimos e máximos da composição de um material é um argumento favorável para a continuidade de estudos de mineração de textos em registros de patentes em uma base de registros bibliográficos em cobertura global. A Figura 4.1 apresenta a evolução desse percentual ao longo dos dez anos avaliados.



Fonte: Elaboração do autor.

Figura 4.1 – Porcentagem de registros em que foi possível recuperar informações sobre a composição química dos elementos avaliados nos registros de patentes em aços com prioridade entre 2009 e 2018 e indexados na *Derwent Innovations Index*.

A tendência geral de aumento no percentual de dados recuperados dos registros ao longo dos dez anos observados é animadora por se tratar de uma evidência de que a *Derwent Innovations Index* preza por seguir com tais detalhes em sua atividade de reescrita de títulos e resumos dos documentos de patentes. Assim, melhorias nos métodos aplicados para a elaboração dos filtros para aços, e uma futura expansão para demais ligas metálicas e classes de materiais, permitirão uma boa cobertura para atividades de elaboração de indicadores para o monitoramento tecnológico e seleção de materiais.

Ao avaliar alguns dos registros pertencentes aos 15% em que não foi possível recuperar a informação sobre o percentual em peso de composição na liga, nota-se que vários registros citam os elementos, mas não detalham suas quantidades. Em vários desses registros ainda foi possível recuperar dados sobre alguma das propriedades buscadas, processamento, microestrutura e aplicações, o que justifica mantê-los na base como propostas de invenções relacionadas aos aços. Esses registros representam 37 diferentes países ou regiões de prioridade, o que não permite concluir, nesta avaliação, que o nível de detalhe dos registros indexados pela DII seja dependente da origem ou do idioma dos documentos originais.

Os filtros elaborados no VantagePoint e o trabalho de organização das informações recuperadas permitiram diferentes níveis de sucesso na busca do máximo detalhamento sobre os desenvolvimentos recentes em aços. As Tabelas 4.1 e 4.2 detalham o número de registros em que foi possível recuperar dados numéricos apresentados na planilha de dados principal elaborada.

Tabela 4.1 – Dados dos elementos de liga recuperados na planilha.

Elemento	Registros	% do total
Carbono	18617	81,26%
Manganês	18247	79,64%
Silício	17706	77,28%
Alumínio	11128	48,57%
Cromo	12814	55,93%
Titânio	8167	35,65%
Nióbio	6663	29,08%
Vanádio	6792	29,65%

Fonte: Elaboração do autor.

Tabela 4.2 – Dados das propriedades recuperadas na planilha.

Propriedade (unidade)	Registros	% do total
LE (MPa)	1277	5,57%
LRT (MPa)	1655	7,22%
Módulo de Young (Gpa)	26	0,11%
Alongamento (%)	569	2,48%
Dureza (HRC)	180	0,79%
Densidade (g/cm <sup>3</sup> )	35	0,15%

Fonte: Elaboração do autor.

O número de registros com informações sobre o percentual em peso de carbono, manganês e silício é próximo ao número geral de registros com alguma informação sobre composição química por se tratar dos elementos de liga mais comuns presentes nos aços. O cromo apresenta um destaque já esperado por caracterizar diversos aços inoxidáveis ou ligas ferrosas com maior resistência à corrosão. Os elementos titânio, nióbio e vanádio, geralmente utilizados como micro ligantes, tiveram seus teores em liga recuperados em cerca de um terço dos registros, o que evidencia sua importância no desenvolvimento recente de aços, seus processos e aplicações. Embora com o uso de um tesouro de limpeza que atualmente contém 55 expressões relacionadas ao alumínio, mas não a sua direta presença nas ligas, sua citação com detalhamento de quantidade em 48% dos registros não é surpreendente. Esse número pode ser justificado pela retomada recente em estudos científicos e patenteamento de aços de baixa densidade [93] e pela atuação do elemento na formação de carbonetos que permitem melhoria nas propriedades mecânicas em aços de microestrutura complexa [85].

A elaboração dos tesouros também foi determinante em uma maior recuperação de valores numéricos das propriedades de Limite de escoamento e de Limite de Resistência à Tração. A substituição de códigos de denominação da norma API 5L pelos seus valores correspondentes dessas propriedades permitiu apresentar esses dados em um percentual superior às demais propriedades, o que abre oportunidades de melhorias em atualizações futuras da base de dados com tratamentos semelhantes que englobem outras normas. Além de possíveis melhorias no desenvolvimento dos filtros no software

bibliométrico para uma recuperação de um maior número dessas informações, o baixo percentual de apresentação dessas propriedades quando comparado à composição química pode estar diretamente relacionado à uma dinâmica comum nas atividades de patenteamento. É importante lembrar que, para a concessão da patente, é necessário a suficiência descritiva do invento, entretanto, não se torna obrigatório o reporte de testes laboratoriais e industriais para provar sua atividade inventiva. Dessa maneira, por estratégia dos próprios titulares em busca da proteção de sua atividade intelectual e industrial, um menor nível de detalhamento dessas vantagens já seria esperado nos documentos originais, o que justifica também sua raridade nos registros bibliográficos.

As Tabelas 4.3, 4.4 e 4.5 apresentam, respectivamente, os processos, microestruturas e aplicações mais citados nos registros apresentados na planilha principal de dados. Como o processo para recuperação dessas informações foi totalmente baseado na elaboração de tesouros para a padronização de códigos CIP e expressões regulares de textos, esses resultados são dependentes da forma em que foram trabalhados esses tesouros na versão apresentada nesta tese. Logo, tanto a denominação quanto os sinônimos ou códigos utilizados podem apresentar alterações em atualizações futuras, também dependentes das escolhas e classificações utilizadas pelos seus desenvolvedores.

Tabela 4.3 – Principais processos recuperados no desenvolvimento dos filtros e apresentados na planilha de dados.

Processo	Registros	% do total
Laminação Controlada	11344	49,51%
Laminação a quente	1566	6,84%
Laminação a frio	1055	4,60%
Pressão reduzida ou vácuo	893	3,90%
Normalização	563	2,46%
Metalurgia do pó	524	2,29%
Endurecimento Superficial	486	2,12%
Desfoforização e Dessulfurização	307	1,34%
Refino Térmico ou <i>Verguten</i>	298	1,30%
Nitruração	284	1,24%

Fonte: Elaboração do autor. Total de registros que apresentam ao menos um processo: 14792 (65%).

Tabela 4.4 – Principais microestruturas apresentadas na planilha de dados.

Fase ou microconstituente	Número de registros	% do total
Austenita	1048	4,57%
Ferrita	736	3,21%
Martensita	621	2,71%
Duplex	618	2,70%
Austenita retida	523	2,28%
Martensita revenida	318	1,39%
Bainita	273	1,19%
Perlita	48	0,21%

Fonte: Elaboração do autor. Total de registros que apresentam ao menos uma fase ou microconstituente: 3728 (16%).

Tabela 4.5 – Principais aplicações para aços apresentadas na planilha de dados.

Aplicação	Número de registros	% do total
Construção civil	1029	4,49%
Dutos	774	3,38%
Automóveis	698	3,05%
Mancais e rolamentos	619	2,70%
Transformadores	619	2,70%
Moldes e matrizes	484	2,11%
Aços elétricos	457	1,99%
Molas	390	1,70%
Ferramentas	376	1,64%
Equipamentos Industriais	329	1,44%

Fonte: Elaboração do autor. Total de registros que apresentam ao menos uma aplicação: 7971 (35%).

Em relação ao processamento, as 24 denominações indicadas nos tesouros de expressões regulares e de códigos da CIP foram apresentadas 20.280 vezes nos registros da planilha, já eliminadas as duplicatas entre diferentes campos dos registros bibliográficos. As futuras atualizações dos filtros e planilha de dados poderão aumentar esse número de denominações e de sucesso na recuperação, bem como possibilitar um maior nível de detalhes para casos de subdivisões da CIP ou do dicionário de termos. Como exemplo, o destaque para o processo de laminação controlada se deve à atribuição desse termo aos registros classificados pelo código “C21D-008” que, em sua descrição em português pelo INPI trata de “Modificação das propriedades físicas por deformação combinada com, ou seguida de, tratamento térmico”. A opção pela denominação de “Laminação Controlada” advém do conhecimento de que os

processos termomecânicos de deformação a quente com o intervalo entre passes têm se tornado fundamentais para a melhoria da resistência à tração em aços, principalmente quando combinadas com o resfriamento acelerado [94], quando possível sua aplicação em escala industrial. Em especial, o código da CIP mais detalhado “C21D-008/02” que trata “para produzir chapas grossas ou tiras para estampagem profunda” foi o subgrupo mais presente além dos derivados do código “C22C-038” utilizado na expressão de busca. Entretanto, essa mesma tendência já foi observada em estudos anteriores de monitoramento tecnológico do desenvolvimento de aços a partir de documentos de patentes [10,55]. Logo, o avanço no detalhamento dos processos de produção seria também facilitado pela criação de novos códigos da CIP mais específicos e próximos da tecnologia incorporada aos inventos atuais.

Por outro lado, a não existência de códigos da CIP para microestruturas dos materiais pode justificar o menor percentual de informação recuperada na versão atual da base de dados, baseada somente em termos que as descrevem. Adicionalmente, essa informação também pode ser estrategicamente não citada nos documentos originais, a exemplo das suas propriedades. Com o conhecimento atual da metalurgia, a microestrutura pode ser estimada com os detalhes da composição química e condições de processamento citadas na patente. Entretanto, uma busca por “*austenitic steels*” (ou expressões próximas) em uma base de documentos de patentes também não apresentaria todas as invenções relacionadas a aços austeníticos, o que dificultaria o monitoramento por parte de concorrentes, por exemplo. Já para fins de análise dos desenvolvimentos recentes, e do conhecimento para o embasamento de planejamento de atividades de pesquisa e desenvolvimento, melhorias futuras no processo de mineração de textos e nos tesouros utilizados em relação aos atuais serão ainda mais proveitosas.

Em termos de aplicações, a atualização dos tesouros também pode ser explorada à exaustão em busca de mais sucesso na recuperação e em seus detalhamentos. As atuais 46 denominações utilizadas que retornaram 9.853 apresentações na versão atual da planilha de dados em termos ou códigos da CIP poderiam ser aprimoradas com avanços na utilização do processamento de

linguagem natural nos textos de documentos completos ou registros de patentes, e a incorporação deles nos tesouros aplicados. Por outro lado, os dados atuais já reforçam a versatilidade de aplicações em aços que, mesmo aparentemente limitados no desenvolvimento de composições de novas ligas totalmente originais, ainda apresentam tal volume de depósitos de patentes que envolvam seus processos de produção e beneficiamento, e a melhoria de suas propriedades.

#### **4.2 Avanços no detalhamento de indicadores para atividades de monitoramento tecnológico**

Uma primeira contribuição possível para o uso da planilha de dados se refere a uma etapa antes mesmo da elaboração dos indicadores: a busca por registros bibliográficos de patentes de uma classe de materiais específica. A Tabela 2.2 apresentou o resultado de diversos testes e análises de seus desenvolvedores e retornou 486 registros bibliográficos que, após mais uma etapa de verificação em software bibliométrico e limpeza para maior precisão, resultou em 426 registros [74]. Da literatura consultada em aços de baixa densidade, os aços duplex de matriz predominantemente ferrítica apresentam composições típicas que variam entre 3 e 7% de alumínio e 2 a 12% de manganês [85]. Assim, a Tabela 4.6 compara o número de registros recuperados, por ano, no estudo anterior, e a partir do número de registros que apresentam valores de alumínio acima de 2% e de manganês acima de 3% na planilha de dados, nos anos que compreendem ambos os escopos temporais.

Tabela 4.6 – Comparação do número de registros em aços de baixa densidade.

Ano de prioridade	Estudo [74]	Dados na planilha
2009	20	3
2010	20	1
2011	19	6
2012	24	7
2013	34	14
2014	42	14
2015	55	12
2016	57	29
Total	271	86

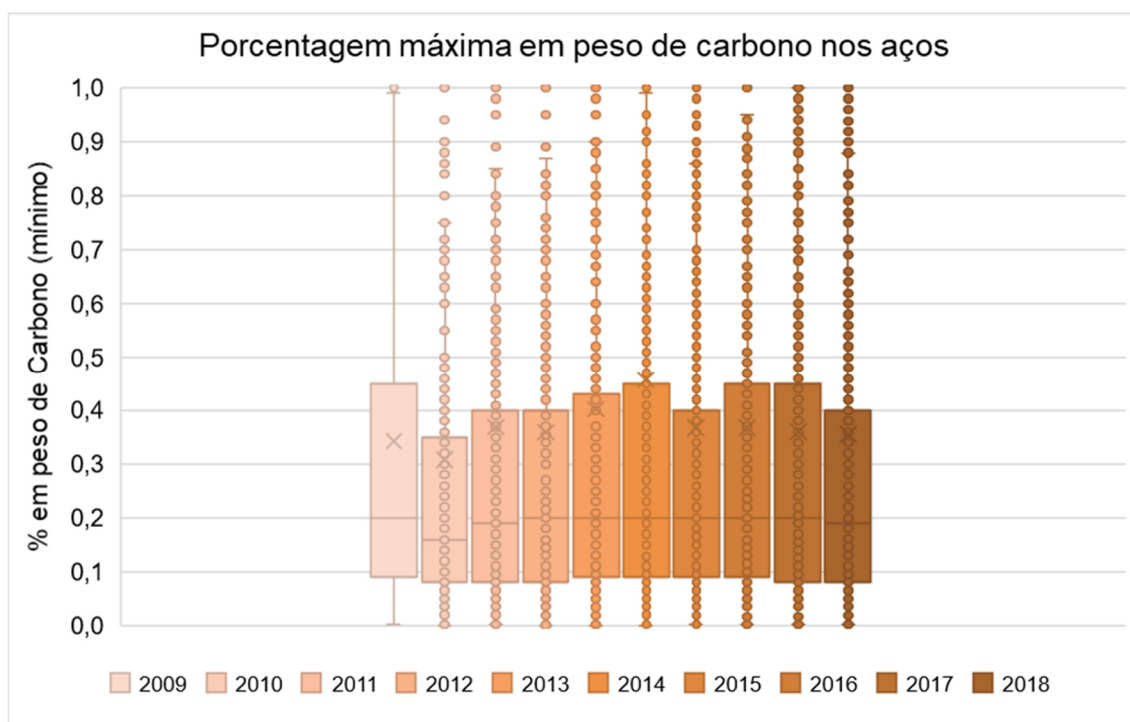
Fonte: Elaboração do autor.

Embora o número de registros recuperados a partir da planilha seja consideravelmente menor no total, eles puderam ser obtidos em um tempo muito menor que o desenvolvimento de uma nova expressão de busca, análise, importação para o software e limpeza. Cabe lembrar que esse resultado também é dependente da recuperação de dados dos teores de elementos químicos avaliados que, por sua vez, depende da disponibilidade de sua presença nos registros e do método empregado nesta primeira versão que pode, futuramente, ser aprimorado. Ainda assim, esses 86 resultados podem servir como uma amostra “padrão ouro” para que os resultados de testes de expressões de busca na base de dados de patentes possam ser comparados com os obtidos na planilha, para sua maior precisão. Adicionalmente, podem ser utilizados para a leitura e identificação de termos chave e códigos da CIP que auxiliem a realização de buscas futuras. Por outro lado, outros estudos de caso em aços, que dependam de outros elementos de liga, também podem apresentar maior sucesso que este avaliado.

A busca ativa por dados sobre a composição química dos elementos de liga em aços em registros bibliográficos de patentes permitiu ampliar o escopo de análise de elementos de ligas em aços não cobertos pela definição de um código específico na CIP. Embora faça parte da definição clássica de aços, o carbono é citado em apenas dois códigos da CIP em sua família de ligas (C22C-



038/36<sup>1</sup> e /56<sup>2</sup>), e com um baixo nível de detalhes por classificarem apenas registros com uma quantidade acima de 1,7% em peso. A Figura 4.2 apresenta a distribuição dos teores percentuais máximos de carbono citados em 17.709 registros de patentes entre os anos de 2009 e 2018 e recuperados pela planilha desenvolvida.



Fonte: Elaboração do autor.

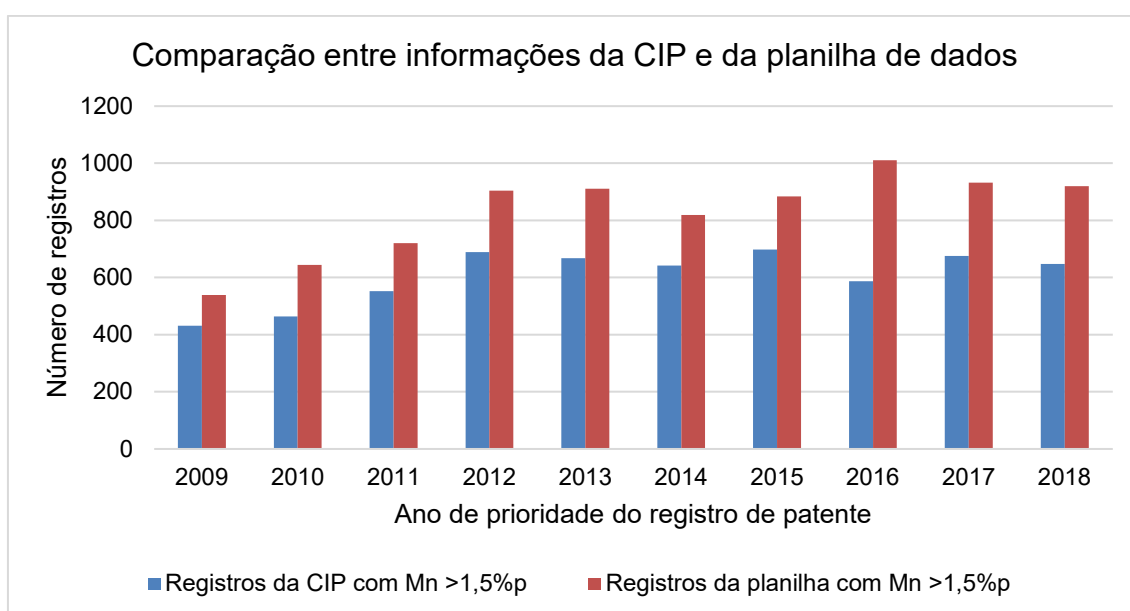
Figura 4.2 – Diagrama de caixas dos valores máximos de carbono citados nos registros de patentes em aço entre 2009 e 2018.

Nota-se que os desenvolvimentos recentes na tradicional indústria siderúrgica têm focado em produtos de baixo teor em peso de carbono. Mais de 75% dos pedidos de patentes nos últimos 10 anos, em que foi possível recuperar esses dados, se referem a ligas com menos de 0,5% em peso, com uma ligeira tendência recente de uma diminuição ainda mais expressiva deste valor. Esse tipo de indicador pode ser elaborado com o auxílio da planilha aqui desenvolvida

<sup>1</sup> Da definição da CIP versão 2020.1 em português: “Ligas ferrosas, p. ex. ligas de aço / contendo cromo / com mais de 1,7% em peso de carbono”.

<sup>2</sup> Ligas ferrosas, p. ex. ligas de aço / contendo cromo / com níquel / com mais de 1,7% em peso de carbono”.

para casos específicos de processos de produção, aplicações ou da análise de alguma propriedade física ou mecânica de interesse. Além disso, um outro benefício no uso da planilha para análise de elementos de liga também é justificado pela flexibilidade de filtrar os registros com valores acima ou abaixo de um determinado valor. Como exemplo, a Figura 4.3 apresenta um comparativo entre o número de registros da amostra final classificados nos códigos “C22C-038/38”<sup>1</sup> e “C22C-038/58”<sup>2</sup>, e o número de registros com teor máximo de manganês acima de 1,5% entre aqueles em que foi possível recuperar tais dados na planilha.



Fonte: Elaboração do autor.

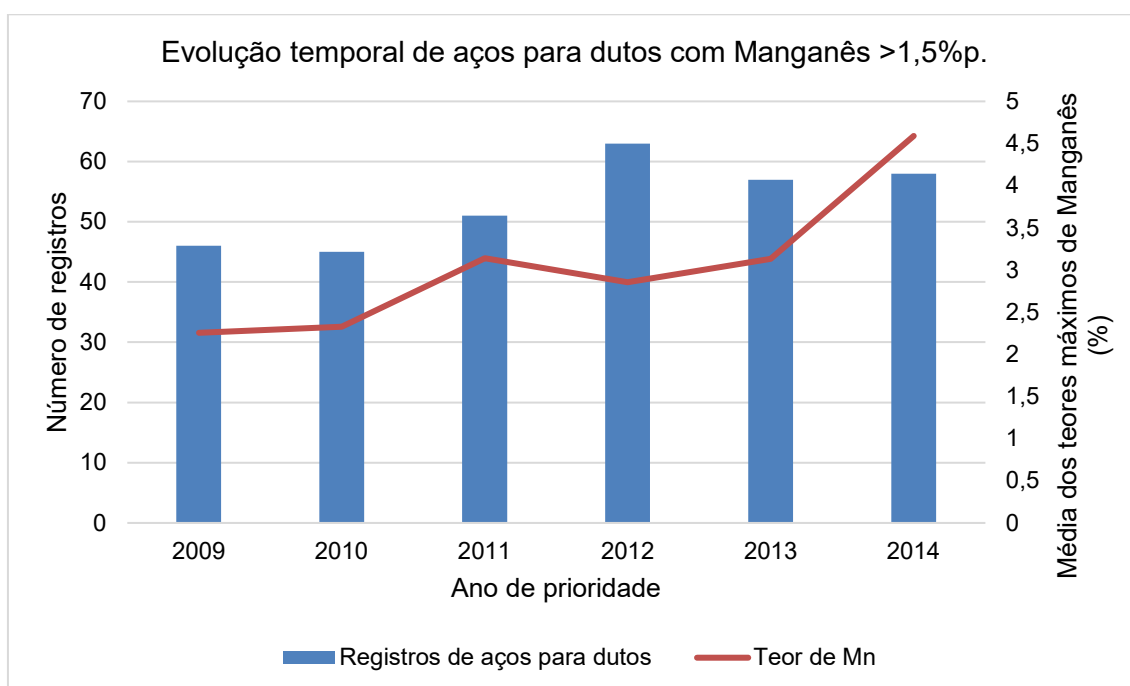
Figura 4.3 – Número de registros com informação de depósitos de patentes com manganês acima de 1,5% em peso, obtidos por códigos da CIP e pela planilha de dados.

Além de incluir registros que possuem acima de 1,5% em peso de manganês, sem a necessidade de também conterem cromo e níquel como no caso dos códigos CIP, esse filtro poderia ser aplicado para qualquer valor e

<sup>1</sup> Da definição da CIP versão 2020.1 em português: “Ligas ferrosas, p. ex. ligas de aço / contendo cromo / com mais de 1,5% em peso de manganês”.

<sup>2</sup> “Ligas ferrosas, p. ex. ligas de aço / contendo cromo / com níquel / com mais de 1,5% em peso de manganês”.

elementos liga de interesse com o uso da planilha. Entretanto, sabe-se que a simples contagem de registros com a presença de um determinado elemento químico citado no título ou resumo poderia ser realizada através de buscas nas próprias bases de registros de patentes ou agrupamento de termos e contagem em um software bibliométrico. Por outro lado, com a apresentação de seus teores em peso como elementos de liga, os indicadores podem trazer mais detalhes sobre a dinâmica de desenvolvimento em um setor de mercado específico. Como exemplo, a Figura 4.4 apresenta um detalhamento adicional, quando comparada à Figura 2.12, com a média dos valores máximos de manganês em registros de patentes de aços para dutos de transporte de hidrocarbonetos.



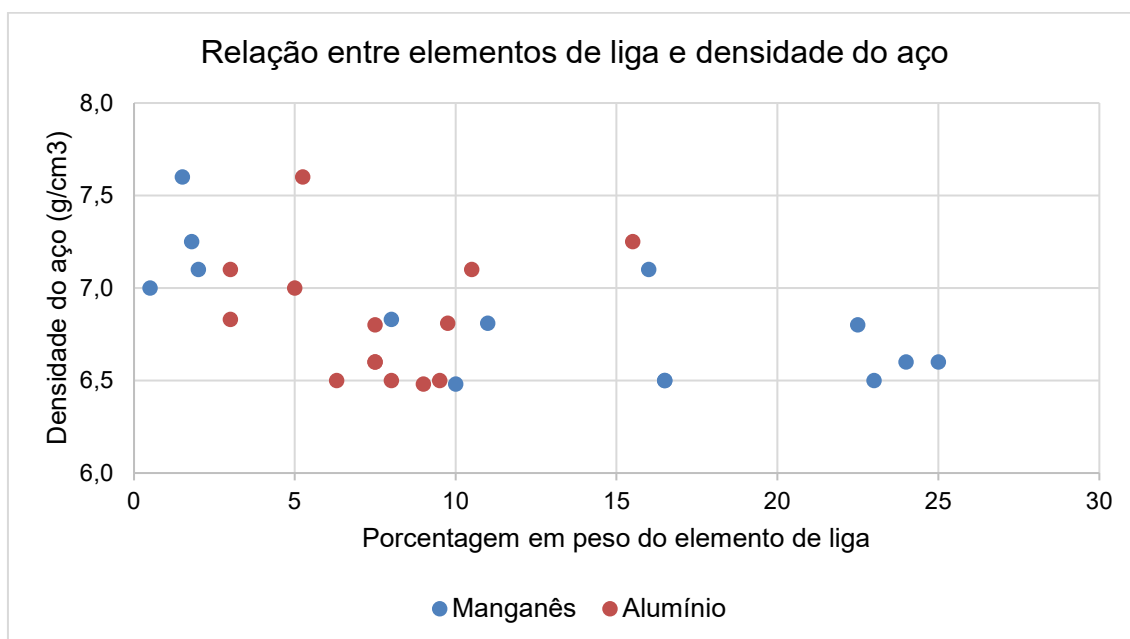
Fonte: Elaboração do autor. Dados de [89] aplicados à planilha.

Figura 4.4 – Número de registros de patentes de aços para dutos de transporte de hidrocarbonetos entre 2009 e 2014 e média dos teores máximos de manganês.

Nota-se que não apenas o número de patentes com esse elemento de liga acima de 1,5% em peso teve aumento no período avaliado, mas que várias das invenções propõem limites máximos maiores, em média e para essa

aplicação, a cada ano. Em um mercado de tal competição em que o número de pedidos de patentes com novas composições químicas antecede até mesmo sua permissão em revisão de normas, tal detalhamento é fundamental para melhor compreender a dinâmica de novos desenvolvimentos propostos.

Com a nova possibilidade de trabalhar com valores numéricos das propriedades de materiais citadas, os indicadores elaborados a partir de registros de patentes também podem contribuir com o auxílio da compreensão de conceitos e tendências abordados nas publicações científicas em temas de desenvolvimento tecnológico recente, como os aços de baixa densidade. Como exemplo, a Figura 4.5 apresenta dados de 14 registros apresentados na planilha desenvolvida em que foi possível recuperar os dados de densidade (valor mínimo citado), e valores médios (entre os mínimos e máximos citados) de alumínio e manganês.



Fonte: Elaboração do autor.

Figura 4.5 – Relação entre a densidade de aços e o percentual médio em peso de alumínio e manganês a partir de dados apresentados na planilha.

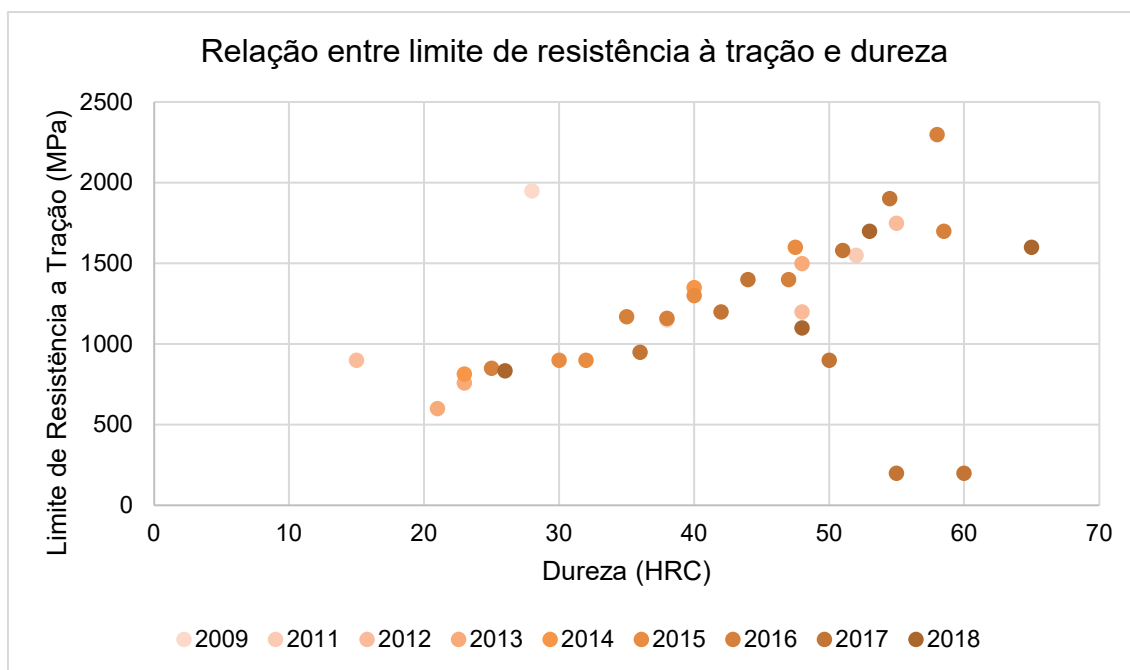
Esse indicador corrobora com a literatura de que a adição de alumínio e manganês nos aços diminui a densidade das ligas ferrosas, desconsiderados efeitos de outros elementos e da densidade da fase contendo majoritariamente

ferro (7,85% para a ferrita e 8,15% austenita) [85]. Adicionalmente, artigos científicos também citam que os desenvolvimentos recentes de aços de baixa densidade exploram a adição de até 12% de alumínio visto que, já nesse valor, os efeitos no módulo elástico da liga são prejudicados [85,95]. Além disso, a adição de manganês em valores superiores a 10% é comumente associada a tenacificação de aços austeníticos pelos mecanismos de plasticidade induzida, em especial por transformação de fase (TRIP, da sigla em inglês) ou por formação de maclas (TWIP, da sigla em inglês) [85].

Logo, o conhecimento da composição química de novas ligas propostas e de suas propriedades físicas e mecânicas estimadas, através de seus registros bibliográficos de registros de patentes, é mais uma fonte de informação para a análise dos avanços científicos e tecnológicos em materiais. Adicionalmente, patentes recentes tratam da apresentação de possíveis soluções que ainda não são produzidas em escala comercial, porém que podem vir a ser em um futuro próximo. Assim, além de orientação para atividades de pesquisa e desenvolvimento acadêmicas e industriais, a utilização dessas informações em processos de seleção de materiais também pode ser considerada.

#### 4.3 Contribuições para a Seleção de Materiais

Em adição às opções que podem ainda não constar em bases comerciais, a planilha de dados permite relacionar os dados recuperados para análise de tendências também em uma variável adicional, como o tempo. Assim, além de comparar a evolução da composição química nos inventos propostos relacionados a aços, pode-se também avaliar como suas propriedades evoluem em um período específico, o que trata de uma contribuição potencial em atividades de ensino e pesquisa que podem ser futuramente exploradas. A Figura 4.6 apresenta um diagrama de comparação entre o limite de resistência à tração e dureza entre os 33 registros apresentados na planilha que disponibiliza ambos os valores, com a escala de tempo do ano de prioridade dos pedidos de patentes.



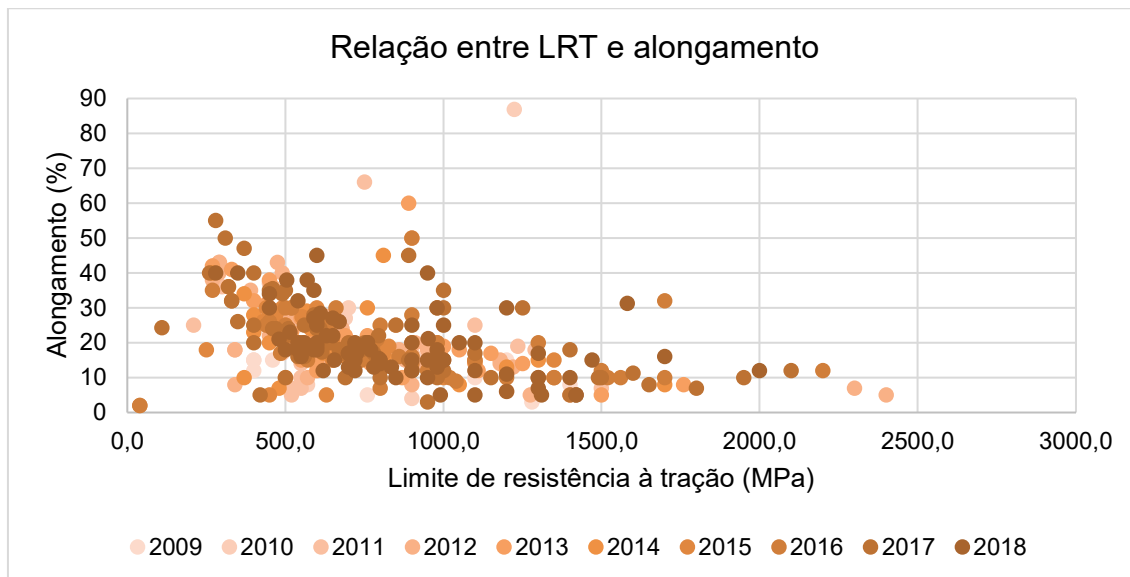
Fonte: Elaboração do autor.

Figura 4.6 – Relação entre limite de resistência à tração e dureza a partir de dados apresentados na planilha.

Do conhecimento em ciência dos materiais, sabe-se que as propriedades avaliadas tendem a seguir uma tendência linear. Entretanto, é possível observar que existem propostas de desenvolvimento nos últimos dez anos que buscam priorizar a resistência à tração ou a dureza. Adicionalmente, é conveniente lembrar que outros dados podem não ter sido recuperados com o método empregado, não constem nos registros dos documentos ou mesmo nos textos completos. Também é possível que tais eventos sequer estejam sendo patenteados como forma de manter o segredo industrial por parte de seus desenvolvedores.

No conjunto de dados filtrados para a elaboração da Figura 4.6 foram citados os processos de laminação controlada, têmpera isotérmica, normalização, pressão reduzida ou vácuo, refino térmico, e nitruração. Entre as aplicações figuraram molas, transformadores, tubos rígidos, trilhos, moldes e matrizes e tubos rígidos. Logo, procuras mais específicas por processos e aplicações podem refinar a análise de acordo com o interesse de pesquisa. O mesmo procedimento pode ser realizado entre propriedades concorrentes, como

limite de resistência à tração e alongamento, como apresentado na Figura 4.7, em que foi possível utilizar dados de 396 registros entre os anos de 2009 e 2018.



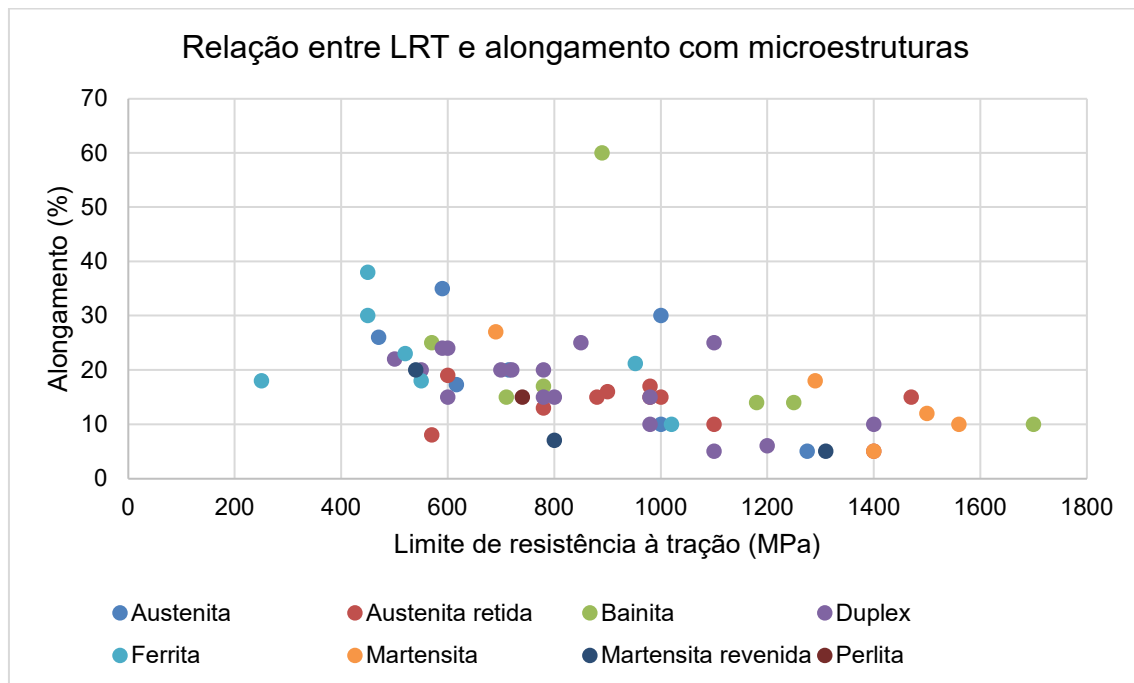
Fonte: Elaboração do autor.

Figura 4.7 – Relação entre limite de resistência à tração e alongamento a partir de dados apresentados na planilha.

Comparada às Figuras 2.10 e 2.11, nota-se que os desenvolvimentos recentes em aços também formam uma “curva da banana” com um notável acúmulo de pontos entre 500 e 1000 MPa de limite de resistência à tração e 10 a 30% de alongamento, que provavelmente figurarão como as principais opções comerciais em futuro próximo. Entretanto, propostas de invenções de materiais que combinem uma alta resistência e alongamento são também presentes, o que fornece insumos para pesquisadores em aços para avaliar como estão atingindo esses valores em termos de composição química e processamento. Novos processos, como o tratamento térmico de têmpera e partição, têm permitido combinar propriedades atrativas para diferentes aplicações [96]. A mesma comparação de propriedades pode ser realizada com o acréscimo de outras variáveis de análise. A Figura 4.8<sup>1</sup>, por exemplo, apresenta dados de 58

<sup>1</sup> Para efeito de elaboração deste indicador, apenas a primeira coluna da planilha referente a microestrutura da planilha foi considerada.

registros em que também foi possível recuperar dados sobre os microconstituintes desses aços.



Fonte: Elaboração do autor.

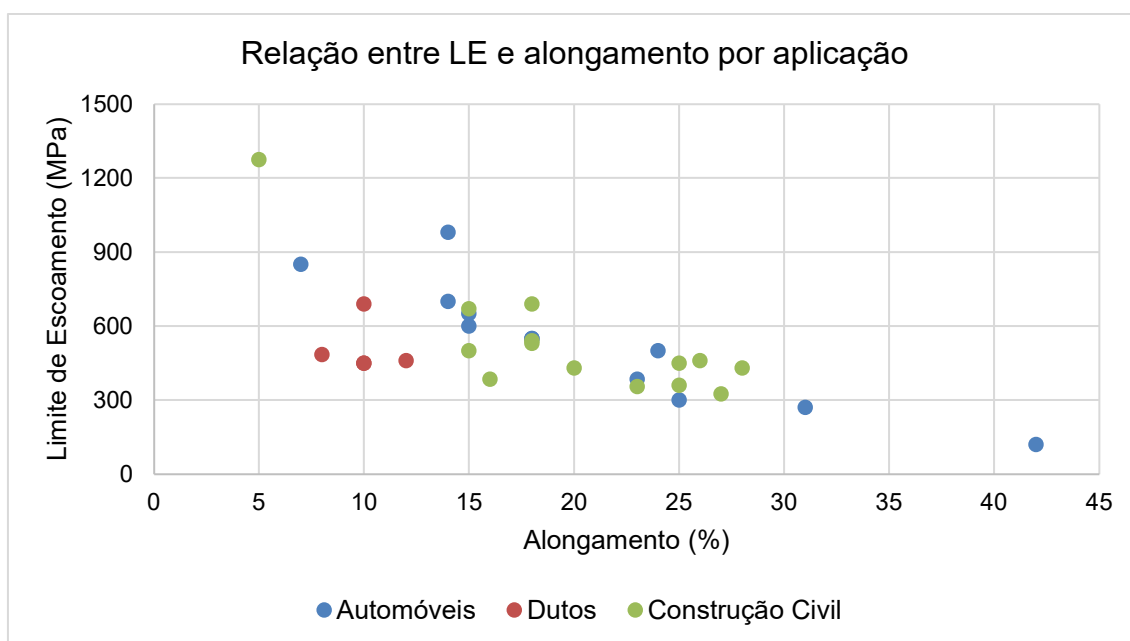
Figura 4.8 – Relação entre limite de resistência à tração e alongamento, com detalhes de microestruturas, a partir de dados apresentados na planilha.

A análise da figura corrobora com o conhecimento de aços em que a microestrutura predominantemente ferrítica compreende limites de resistência menores em prol de um alongamento percentual maior, enquanto maiores resistências à tração em detrimento de uma menor ductilidade como características dos aços martensíticos. Entretanto, nota-se que a região que apresentava a maior parte dos inventos na Figura 4.6 se repete com uma grande diversidade de fases e microconstituintes distintos. Provavelmente, trata-se de desenvolvimentos de microestruturas multifásicas e/ou complexas, ou de avanços nas condições de processamento que permitem o trabalho de outras variáveis que podem vir a ser apresentadas em futuras atualizações de filtros e tesouros, como a precipitação de carbonetos ou controle de tamanho de grão.

Adicionalmente, a organização dos dados em uma planilha permite que pesquisadores acadêmicos e industriais, possam explorar possíveis aplicações



de seus estudos e inventos, sejam eles relacionados a composição química ou processos. Como exemplo, a Figura 4.9 apresenta um diagrama composto por 31 dados de limite de escoamento e de alongamento para as três principais aplicações recuperadas e apresentadas na atual versão da planilha.



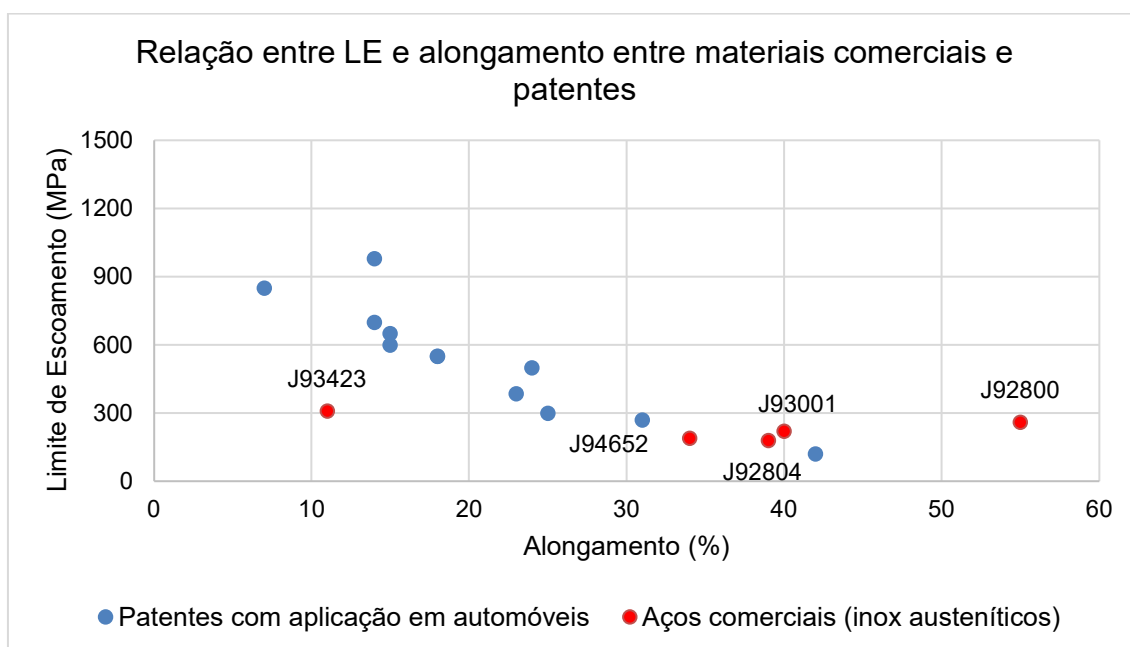
Fonte: Elaboração do autor.

Figura 4.9 – Relação entre limite de escoamento e alongamento em registros com aplicação requisitada nos mercados de automóveis, dutos e construção civil.

Apesar de poucos dados, para a situação apresentada e em sua versão atual, confirma-se a versatilidade de propriedades dos aços para o mercado automobilístico, desde os que priorizam ductilidade aos que prezam pela resistência à tração, como apresentado na Figura 2.10. Também é notado um foco em desenvolvimentos recentes de aços com alongamento acima de 15% para aplicações voltadas à construção civil, que pode orientar a seleção de materiais para uma nova tendência ainda não perceptível em bases de dados que contenham apenas materiais comerciais. Os desenvolvimentos aços para dutos ocupam valores intermediários, o que pode indicar que o aumento do limite de escoamento, como previsto nas revisões das normas, possa ainda não ser o

destaque dos desenvolvimentos mais recentes e que estejam sendo patenteados.

A organização de dados em uma planilha, aberta para edições e atualizações além das informações colhidas pelos filtros elaborados, tornou possível algumas comparações entre dados recuperados dos registros de patentes e materiais comerciais. A Figura 4.10 foi elaborada a partir da Figura 4.9, com a adição de dados de alguns materiais comerciais, a título de exemplo, identificados pelos seus códigos ASTM UNS obtidos a partir de uma base de dados gratuita [36].



Fonte: Elaboração do autor.

Figura 4.10 – Relação entre limite de escoamento e alongamento em registros de documentos de patentes e aços comerciais.

O mesmo procedimento permite que pesquisadores utilizem os dados de seus experimentos para comparar com as opções comerciais ou novas opções que têm sido patenteadas. A inserção desses dados permitirá, por exemplo, a avaliação de opções semelhantes em composição química, fases, processos utilizados e aplicações requeridas, de acordo com a disponibilidade que se apresentam. Assim, espera-se que o método desenvolvido, neste primeiro protótipo até suas versões mais atualizadas, possa contribuir para que esforços

repetitivos de pesquisa e desenvolvimento sejam evitados, para que novos inventos em materiais sejam estimulados e que, principalmente por parte de pesquisadores acadêmicos, possam ser patenteados em busca de aplicações para solução de desafios técnicos em materiais.



## 5 CONCLUSÕES

A elaboração de uma planilha de dados de materiais a partir da mineração de textos de registros bibliográficos de documentos de patentes se mostrou viável, embora com muitas oportunidades de desenvolvimentos e aperfeiçoamentos futuros. A comparação entre os resumos apresentados pelos registros bibliográficos indexados pela *Derwent Innovations Index* e seus respectivos textos das reivindicações de documentos originais mostrou um padrão semelhante na apresentação de informações, embora não idênticos e em quantidade de detalhes. Mesmo assim, a utilização de marcadores para seleção de segmentos de textos permitiu o desenvolvimento de filtros em software bibliométrico que recuperou informações sobre a composição química, microestruturas, processos, aplicações e de seis diferentes propriedades de aços com pedidos de patentes registrados nos últimos dez anos.

Foi possível apresentar informações sobre percentuais em peso de pelo menos um de oito elementos de liga avaliados em 85% dos 22.911 registros que compõem a primeira versão da planilha de dados. Esse percentual pode ser maior com a inclusão de mais elementos de liga em atualizações futuras. Já para as propriedades, em geral, a recuperação foi abaixo de 10%. Uma alternativa para aumentar esse percentual, no futuro, é aproveitar de valores descritos em normas que, ao serem substituídos nos textos dos registros, serão recuperados com maior facilidade nos filtros bibliométricos.

A busca por dados obtidos a partir da mineração de textos de caracteres numéricos, como a composição química e propriedades, permitiu elaborar indicadores que os relacionam, e que podem ser explorados de acordo com outros interesses futuros. Já a busca por processos, aplicações e microestruturas, que partiu de seleção de textos de caracteres alfabéticos após o uso de tesouros de agrupamentos, disponibiliza análises de monitoramento tecnológico e de seleção de materiais com um maior nível de detalhes, comparado ao uso de processamento de linguagem natural e códigos da CIP para agrupamento e contagem de registros.

Para atividades de monitoramento tecnológico destacam-se contribuições como a possibilidade de indicadores quantitativos de elementos de liga que não

são cobertos por um código específico da CIP. No caso do teor de carbono em aços, por exemplo, foi possível avaliar que o foco dos últimos dez anos se concentra na siderurgia de baixo carbono, o que seria inviável de confirmar com códigos da CIP que descrevem sua quantidade apenas acima de 1,7% em peso. Adicionalmente, o número de registros apresentados a partir de dados da planilha com valores de manganês acima de 1,5% em peso retornou um número maior que documentos classificados com a mesma informação de códigos da CIP. Logo, os indicadores elaborados possuem um nível maior de detalhes e seu uso na avaliação de tecnologias específicas se torna ainda mais apurado. No estudo de caso de aços para dutos de transporte de hidrocarbonetos, foi possível avaliar a evolução dos teores máximos de elementos previstos em normas. Já como observado no cenário dos aços de baixa densidade, torna-se possível avaliar os desenvolvimentos ao nível do efeito dos elementos de liga nas propriedades, o que abre uma nova oportunidade de análise complementar à literatura científica.

No campo da seleção de materiais, além de informações complementares a opções disponíveis nas bases de dados já existentes, foi possível adicionar uma terceira variável nos diagramas de comparação de propriedades de materiais. Isso permitiu avaliar a dinâmica temporal dos últimos dez anos nos valores de limite de resistência à tração e alongamento dos aços, com propostas que visam superar o desempenho apresentado na literatura. Adicionalmente, a mesma análise com a informação recuperada sobre a microestrutura citada nessas invenções complementa tanto a compreensão sobre os desenvolvimentos recentes, quanto se apresenta como uma ferramenta gráfica adicional para auxílio à tomada de decisão. Por fim, a possibilidade de comparar diferentes aplicações em um mesmo diagrama de propriedades pode abrir oportunidades de novos desenvolvimentos para mercados específicos.

Ainda que em uma primeira versão, a organização dos dados em uma única planilha fornece um banco de informações de composição química, processos, aplicações e propriedades que pode orientar pesquisadores a avaliar o cenário de desenvolvimentos recentes em suas atuais ou futuras pesquisas. Em especial, conhecer aplicações sobre inventos semelhantes pode incentivar

pesquisadores a buscar um potencial comercial de seus inventos e estimular o uso da patente como mecanismo de proteção intelectual e de retorno do investimento colocado. Assim, tanto para estudos de monitoramento tecnológico quanto de seleção de materiais, incluindo a identificação de aplicações, o uso de registros bibliográficos de patentes como fonte de informação se mostra ainda mais factível, visto que o nível de detalhes foi aprimorado, assim como novas possibilidades de análises que relacionam os dados recuperados. Novos avanços em mineração de textos permitirão, no futuro, que esses dados sejam ainda em maior número, com menos etapas necessárias, e com maior confiabilidade.





## 6 SUGESTÕES PARA FUTUROS TRABALHOS

A primeira sugestão de trabalhos futuros é continuar o desenvolvimento de tesouros e filtros bibliométricos, ou outras ferramentas de mineração de textos, que permitam aumentar a recuperação de informação de patentes em aços e elaborar planilhas de dados semelhantes para outras classes de materiais. Para o aperfeiçoamento da versão desenvolvida nesta tese, e que também se aplicaria a esses novos desenvolvimentos, algumas sugestões e caminhos possíveis são listados.

- Aplicar ou aperfeiçoar a lógica de recuperação dos dados numéricos para outros elementos de liga e propriedades de materiais.

- Inserir linhas de índice de mérito nos diagramas de relação entre propriedades.

- Desenvolver um método que torne possível a análise de registros referentes somente a patentes concedidas.

- Acrescentar os titulares das patentes na planilha, para que a busca por possíveis parceiros de pesquisa e licenciamento de tecnologias seja facilitada.

- Adaptar a metodologia para outros sistemas de recuperação e organização de dados, de preferência em softwares livres.

- Realizar a adaptação dos filtros de software bibliométrico para bases de textos completos e comparar os resultados obtidos.

- Explorar o potencial da ferramenta para atividades de ensino em engenharia.



## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] TOTAL MATERIA. **The world's most comprehensive materials database**. Disponível em: <https://www.totalmateria.com/>. Acesso em: 20 out. 2020.
- [2] ASHBY, M. F.; BRÉCHET, Y. J. M.; CEBON, D.; SALVO, L. Selection strategies for materials and processes. **Materials and Design**, vol. 25, p. 51–67, 2004. DOI 10.1016/S0261-3069(03)00159-6.
- [3] CHATTERJEE, P.; ATHAWALE, V. M.; CHAKRABORTY, S. Materials selection using complex proportional assessment and evaluation of mixed data methods. **Materials and Design**, vol. 32, p. 851–60, 2011. DOI 10.1016/j.matdes.2010.07.010.
- [4] CHATTERJEE, P.; CHAKRABORTY, S. Material selection using preferential ranking methods. **Materials and Design**, vol. 35, p. 384–93, 2012. DOI 10.1016/j.matdes.2011.09.027.
- [5] SANTOS, C. V.; OLIVEIRA, B. S.; MILANEZ, D. H.; LEIVA, D. R.; FARIA, L. I. L. Contribuições da Bibliometria para a Seleção de Materiais no Desenvolvimento de Produtos. *In: Encontro Brasileiro de Bibliometria e Cientometria: A Ciência em Rede*, 6. 2018, Rio de Janeiro **Anais [...]**. p. 453-463.
- [6] ALBIÑANA, J. C.; VILA, C. A framework for concurrent material and process selection during conceptual product design stages. **Materials and Design**, vol. 41, p. 433–46, 2012. DOI 10.1016/j.matdes.2012.05.016.
- [7] RAMALHETE, P. S.; SENOS, A. M. R.; AGUIAR, C. Digital tools for material selection in product design. **Materials and Design**, vol. 31, p. 2275–87, 2010. DOI 10.1016/j.matdes.2009.12.013.
- [8] MILANEZ, D. H.; FARIA, L. I. L.; AMARAL, R. M.; GREGOLIN, J. A. R. Claim-based patent indicators: A novel approach to analyze patent content and monitor technological advances. **World Patent Information**, vol. 50, p. 64–72, 2017. DOI 10.1016/j.wpi.2017.08.008.
- [9] MILANEZ, D. H.; AMARAL, R. M. ; FARIA, L. I. L. De; GREGOLIN, J. A. R. Technological indicators of nanocellulose advances obtained from data and text mining applied to patent documents. **Materials Research-Ibero-american Journal of Materials**, vol. 17, p. 1513–22, 2014. DOI 10.1590/1516-

1439.266314.

[10] OLIVEIRA, B. S.; MILANEZ, D. H.; LEIVA, D. R. Indicadores sobre desenvolvimentos recentes em aços a partir da análise de documentos de patente. *In: Congresso anual da ABM, 73.*, 2018, São Paulo. **Anais [...]**. São Paulo: Associação Brasileira de Metalurgia Materiais e Mineração, 2018. p. 1-10.

[11] FARAG, M. M. **Materials & process selection for engineering design**. Boca Raton: CRC Press, 2008. 432 p. ISBN 9781420063080.

[12] ASHBY, M. F. **Seleção de materiais no projeto mecânico**. 3a. ed. São Paulo: Campos, 2010. 603 p. ISBN 9780750661683.

[13] KUMAR, R.; JAGADISH; RAY, A. Selection of Material for Optimal Design Using Multi-criteria Decision Making. **Procedia Materials Science**, vol. 6, p. 590–6, 2014. DOI 10.1016/j.mspro.2014.07.073.

[14] BRIFCANI, N.; DAY, R.; WALKER, D.; HUGHES, S.; BALL, K.; PRICE, D. A Review of Cutting-edge Techniques for Material Selection. *In: Second International Conference on Advanced Composite Materials and Technologies for Aerospace Applications, 2.*, Wrexham - Reino Unido. **Anais [...]**. Glyndwr University, 2012. p. 58–64.

[15] EDWARDS, K. L. Selecting materials for optimum use in engineering components. **Materials and Design**, vol. 26, p. 469–73, 2005. DOI 10.1016/j.matdes.2004.07.004.

[16] JAHAN, A.; ISMAIL, M. Y.; SAPUAN, S. M.; MUSTAPHA, F. Material screening and choosing methods - A review. **Materials and Design**, vol. 31, p. 696–705, 2010. DOI 10.1016/j.matdes.2009.08.013.

[17] ZAID, A. I. O. Material Selection for Design, Manufacturing and Application. **Key Engineering Materials**, vol. 510–511, p. 386–93, 2012. DOI 10.4028/www.scientific.net/KEM.510-511.386.

[18] NICOGUARO. **Material-comparison-strength-vs-density\_plain**.

Disponível em: [https://commons.wikimedia.org/wiki/File:Material-comparison--strength-vs-density\\_plain.svg](https://commons.wikimedia.org/wiki/File:Material-comparison--strength-vs-density_plain.svg). Acesso em: 5 out. 2020.

[19] JAHAN, A.; EDWARDS, K. L. A state-of-the-art survey on the influence of normalization techniques in ranking: Improving the materials selection process in engineering design. **Materials and Design**, vol. 65, p. 335–42, 2015. DOI

10.1016/j.matdes.2014.09.022.

[20] LANDRU, D.; BRÉCHET, Y.; ASHBY, M. F. Finding applications for materials. **Advanced Engineering Materials**, vol. 4, p. 343–9, 2002. DOI 10.1002/1527-2648(20020605)4:6<343::AID-ADEM343>3.0.CO;2-V.

[21] SALIMON A, A.; BRÉCHET, Y.; ASHBY, M. F.; GREER, A. L. Selection of applications for a material. **Advanced Engineering Materials**, vol. 6, p. 249–65, 2004. DOI 10.1002/adem.200400002.

[22] SALIMON, A.; BRÉCHET, Y.; ASHBY, M. F.; GREER, A. L. Potential applications for steel and titanium metal foams. **Journal of Materials Science**, vol. 40, p. 5793–9, 2005. DOI 10.1007/s10853-005-4993-x.

[23] ALLWOOD, J. M.; KING, G. P. F.; DUFLOU, J. A structured search for applications of the incremental sheet-forming process by product segmentation. **Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture**, vol. 219, p. 239–44, 2005. DOI 10.1243/095440505X8145.

[24] VOSviewer. **Welcome to VOSviewer**. Disponível em: <https://www.vosviewer.com/>. Acesso em: 20 out. 2020.

[25] SANTOS, J.; GOUVEIA, R. M.; SILVA, F. J. G. Designing a new sustainable approach to the change for lightweight materials in structural components used in truck industry. **Journal of Cleaner Production**, vol. 164, p. 115–23, 2017. DOI 10.1016/j.jclepro.2017.06.174.

[26] PU, Y.; MA, F.; ZHANG, J.; YANG, M. Optimal Lightweight Material Selection for Automobile Applications Considering Multi-Perspective Indices. **IEEE Access**, vol. 6, p. 8591–8, 2018. DOI 10.1109/ACCESS.2018.2804904.

[27] SANTOS, C. V. dos; LEIVA, D. R.; COSTA, F. R.; GREGOLIN, J. A. R. Materials Selection for Sustainable Executive Aircraft Interiors. **Materials Research-Ibero-american Journal of Materials**, vol. 19, p. 339–52, 2016. DOI 10.1590/1980-5373-MR-2015-0290.

[28] ALWOOD, J. M.; CULLEN, J. M. **Sustainable materials: With both eyes open**. 1. ed. Cambridge: UIT Cambridge, 2011. 384 p. ISBN 978-1-906-86005-9.

[29] SINGH, S.; OLUGU, E. U.; MUSA, S. N.; MAHAT, A. B.; WONG, K. Y.

Strategy selection for sustainable manufacturing with integrated AHP-VIKOR method under interval-valued fuzzy environment. **International Journal of Advanced Manufacturing Technology**, vol. 84, p. 547–63, 2016. DOI 10.1007/s00170-015-7553-9.

[30] GOVINDAN, K.; MADAN SHANKAR, K.; KANNAN, D. Sustainable material selection for construction industry - A hybrid multi criteria decision making approach. **Renewable and Sustainable Energy Reviews**, vol. 55, p. 1274–88, 2016. DOI 10.1016/j.rser.2015.07.100.

[31] PEREIRA, B. C. F.; OLIVEIRA, B. S.; LEIVA, D. R. Análise bibliométrica da produção científica em métodos de tomada de decisão para seleção de materiais. *In*: Congresso Brasileiro de Engenharia e Ciências dos Materiais, 23., 2018, Natal-RN. **Anais [...]. Metallum**, 2018. p. 9144-9155.

[32] MANDAVIA, M.; VIJAYRAGHAVAN, K. **Global steel market to reach \$865.5 billion by 2020: MarketLine**. 2017. Disponível em: <https://economictimes.indiatimes.com/industry/indl-goods/svs/steel/global-steel-market-to-reach-865-5-billion-by-2020-marketline/articleshow/57560156.cms>. Acesso em: 6 nov. 2018.

[33] ANSYS. **Granta Selector**. Disponível em: [www.grantadesign.com/products/ces/](http://www.grantadesign.com/products/ces/). Acesso em: 20 out. 2020.

[34] MATWEB. **Material property data**. Disponível em: <http://www.matweb.com/>. Acesso em: 20 out. 2020.

[35] MAKE IT FROM. **Material Properties Database**. Disponível em: <https://www.makeitfrom.com/>. Acesso em: 20 out. 2020.

[36] CUSTOMPARTNET. **Custom Part Net**. Disponível em: <http://www.custompartnet.com/>. Acesso em: 20 out. 2020.

[37] KNOVEL. **Engineering Data and Technical References to Solve Engineering Problems**. Disponível em: <https://app.knovel.com/web/index.v>. Acesso em: 20 out. 2020

[38] KEELER, S.; KIMCHI, M.; J. MOONEY, P. **Advanced High-Strength Steels Application Guidelines Version 6.0**: 2017. 314 p. Disponível em: <https://www.worldautosteel.org/projects/advanced-high-strength-steel-application-guidelines/>. Acesso em: 20 out. 2020.

- [39] PIQUEIRA, J. R. C.; OLIVEIRA, S. M. Engenharia De Sistemas Complexos. **Estudos Avançados**, vol. 31, p. 249–55, 2017. DOI 10.1590/s0103-40142017.3191018.
- [40] SILVA, L. C. S.; KOVALESKI, J. L.; GAIA, S.; FRANCISCO, A C. Informação Tecnológica: Identificando tecnologias, vantagens e aplicações através Do banco nacional e Internacional de patentes. **Holos**, vol. 29, p. 139–50, 2013.
- [41] MOGEE, M. E. Patents and Technology Intelligence. (Eds.) **Keeping abreast of science and technology: technical intelligence for business**, Columbus: Battelle Press, p. 295–335, 1997.
- [42] JUNGSMANN, D. de M.; BONETTI, E. A. **Inovação e propriedade intelectual. Guia para o Docente**. 2010. 97 p. ISBN 9788575193891.
- [43] INPI - Instituto Nacional da Propriedade Industrial. **Busca de Patentes**. Disponível em: <http://www.inpi.gov.br/menu-servicos/informacao/busca-de-patentes>. Acesso em: 1 ago. 2018.
- [44] OKUBO, Y. Bibliometric Indicators and Analysis of Research Systems Methods and Examples. **OECD Science, Technology and Industry Working Papers 1997/01**, 1997.
- [45] FARIA, L. I. L.; GREGOLIN, J. A. R.; SANTOS, R. N. M. Informação Tecnológica e Seleção de Materiais. **International Journal of Information Science for Decision Making (Online)**, vol. 1, p. 27–42, 1998.
- [46] MARTINO, J. P. A review of selected recent advances in technological forecasting. **Technological Forecasting and Social Change**, vol. 70, p. 719–33, 2003. DOI 10.1016/S0040-1625(02)00375-X.
- [47] PORTER, A. L.; CUNNINGHAM, S. W. **Tech mining. Exploiting New Technologies for Competitive Advantage**. New Jersey: Wiley-Interscience. 2005. ISBN 9786468600.
- [48] ERNST, H. The Use of Patent Data for Technological Forecasting: The Diffusion of CNC-Technology in the Machine Tool Industry. **Small Business Economics**, vol. 9, p. 361–81, 1997. DOI 10.1023/A:1007921808138.
- [49] WALTMAN, L.; VAN ECK, N. J.; NOYONS, E. C. M. A unified approach to mapping and clustering of bibliometric networks. **Journal of Informetrics**, vol. 4, p. 629–35, out. 2010. DOI 10.1016/j.joi.2010.07.002.

- [50] MADANI, F.; WEBER, C. The evolution of patent mining: Applying bibliometrics analysis and keyword network analysis. **World Patent Information**, vol. 46, p. 32–48, 2016. DOI 10.1016/j.wpi.2016.05.008.
- [51] ABBAS, A.; ZHANG, L.; KHAN, S. U. A literature review on the state-of-the-art in patent analysis. **World Patent Information**, vol. 37, p. 3–13, 2014. DOI 10.1016/j.wpi.2013.12.006.
- [52] WIPO - World Intellectual Property Organization. **International Patent Classification (IPC) Official Publication**. 2016. Disponível em: <http://www.wipo.int/classifications/ipc/en/>. Acesso em: 5 fev. 2017.
- [53] ESPACENET. **Classificação Internacional de Patentes**. Disponível em: [https://lp.espacenet.com/help?locale=pt\\_LP&method=handleHelpTopic&topic=ipc](https://lp.espacenet.com/help?locale=pt_LP&method=handleHelpTopic&topic=ipc). Acesso em 03 nov. 2020.
- [54] JENKINS, G. Function and application - Is there a difference? **World Patent Information**, vol. 11, p. 121–4, 1989. DOI 10.1016/0172-2190(89)90140-3.
- [55] OLIVEIRA, B. S.; ISHIKAWA, T. T.; FARIA, L. I. L. De; LEIVA, D. R. Monitoramento Tecnológico em ligas ferrosas utilizadas em tubos na indústria petrolífera a partir da análise de documentos de patentes. *In*: Congresso anual da ABM, 70., 2015, Rio de Janeiro. **Anais [...]**. Associação Brasileira de Metalurgia Materiais e Mineração, 2015. p. 331-339.
- [56] CHANCHETTI, L. F.; OVIEDO DIAZ, S. M.; MILANEZ, D. H.; LEIVA, D. R.; DE FARIA, L. I. L.; ISHIKAWA, T. T. Technological forecasting of hydrogen storage materials using patent indicators. **International Journal of Hydrogen Energy**, vol. 41, p. 18301–10, 2016. DOI 10.1016/j.ijhydene.2016.08.137.
- [57] OLIVEIRA, B. S.; MILANEZ, D. H.; ISHIKAWA, T. T.; FARIA, L. I. L. ; LEIVA, D. R. Indicadores sobre o desenvolvimento de tecnologias em conexões para tubos utilizados em exploração e produção de óleo e gás. *In*: Congresso Nacional de Engenharia de petróleo, gás natural e biocombustíveis, 1., 2015, Campina Grande. **Anais [...]**. Editora Realize, 2015. 8 p.
- [58] MILANEZ, D. H.; NOYONS, E.; DE FARIA, L. I. L. A delineating procedure to retrieve relevant publication data in research areas: the case of nanocellulose. **Scientometrics**, vol. 107, 2016. DOI 10.1007/s11192-016-1922-5.
- [59] MILANEZ, D. H.; FARIA, L. I. L.; LEIVA, D. R.; KIMINAMI, C. S.; BOTTA, W.



J. Assessing technological developments in amorphous/glassy metallic alloys using patent indicators. **Journal of Alloys and Compounds**, vol. 716, p. 330–5, 2017. DOI 10.1016/j.jallcom.2017.05.105.

[60] DE OLIVEIRA, B. S.; MILANEZ, D. H.; LEIVA, D. R.; DE FARIA, L. I. L.; BOTTA, W. J.; KIMINAMI, C. S. Thermal spraying processes and amorphous alloys: Macro-indicators of patent activity. **Materials Research-Ibero-american Journal of Materials**, vol. 20, 2017. DOI 10.1590/1980-5373-MR-2017-0416.

[61] ALBALA, M. **Making Sense of Big Data in the Petabyte Age**. Bullet Cognizant 20-20 Insights Executive Summary. Disponível em: <https://www.cognizant.com/InsightsWhitepapers/Making-Sense-of-Big-Data-in-the-Petabyte-Age.pdf>. Acesso em 20. out. 2020.

[62] CHOUDHARY, A. K.; HARDING, J. A.; TIWARI, M. K. Data mining in manufacturing: A review based on the kind of knowledge. **Journal of Intelligent Manufacturing**, vol. 20, p. 501–21, 2009. DOI 10.1007/s10845-008-0145-x.

[63] GROBELNIK, M.; MLADENIC, D.; MILIC-FRAYLING, N. Text mining as integration of several related research areas. **ACM SIGKDD Explorations Newsletter**, vol. 2, p. 99–102, 2000. DOI 10.1145/380995.381051.

[64] SHEREMETYEVA, S. Natural language analysis of patent claims. **Proceedings of the ACL-2003 workshop on Patent corpus processing - Volume 20**, p. 66–73, 2003. DOI <http://dx.doi.org/10.3115/1119303.1119311>.

[65] NOH, H.; JO, Y.; LEE, S. Keyword selection and processing strategy for applying text mining to patent analysis. **Expert Systems with Applications**, vol. 42, p. 4348–60, 2015. DOI 10.1016/j.eswa.2015.01.050.

[66] NIEMANN, H.; MOEHRLE, M. G.; FRISCHKORN, J. Use of a new patent text-mining and visualization method for identifying patenting patterns over time: Concept, method and test application. **Technological Forecasting and Social Change**, vol. 115, p. 210–20, 2017. DOI 10.1016/j.techfore.2016.10.004.

[67] MILANEZ, D. H. **Elaboração de Indicadores de Ciência e Tecnologia para o monitoramento de avanços em nanocelulose**. 2015. 187 f. Dissertação (Mestrado) - Curso de Ciência e Engenharia de Materiais, Universidade Federal de São Carlos, São Carlos, 2015.

[68] CLARIVATE ANALYTICS. **Derwent Innovations Index on Web of**

**Science**. Disponível em: <https://clarivate.com/webofscicencegroup/solutions/webofscicence-derwent-innovations-index>. Acesso em: 20 out. 2020.

[69] CAPES - Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior.

**Periódicos CAPES**. Disponível em: <http://www-periodicos-capes-gov-br.ez31.periodicos.capes.gov.br/>. Acesso em: 20 out. 2020.

[70] GOOGLE. **Google Patents**. Disponível em: [patents.google.com](http://patents.google.com). Acesso em 20 out. 2020.

[71] ESPACENET. **Espacenet Patent Search**. Disponível em: <https://www.epo.org/searching-for-patents/technical/espacenet.html>. Acesso em 20 out. 2020.

[72] DANG, J.; MOTOHASHI, K. Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality. **China Economic Review**, vol. 35, p. 137–55, 2015. DOI 10.1016/j.chieco.2015.03.012.

[73] FISCH, C. O.; BLOCK, J. H.; SANDNER, P. G. Chinese university patents: quantity, quality, and the role of subsidy programs. **Journal of Technology Transfer**, vol. 41, p. 60–84, 2016. DOI 10.1007/s10961-014-9383-6.

[74] SILVA, R. F.; OLIVEIRA, B. S.; LEIVA, D. R. Indicadores de patentes em aços de baixa densidade e aços de alto módulo elástico. *In*: Encontro Nacional dos Estudantes de Engenharias Metalúrgica, de Materiais e Mineração. 19. 2019. São Paulo. **Anais [...]**. Associação Brasileira de Metalurgia, Materiais e Mineração. p. 98–108, 2019. DOI 10.5151/2594-4711-33584.

[75] MATSUMOTO, A.; NAKATA, H.; SHUNSUKE, T. **Hot-rolled steel sheet for coiled tubing**. Titular: Jfe Steel. JP n. JP6384635-B1. Depósito: 25 jan. 2017. Concessão: 05 set. 2018.

[76] BONOTTO, M. E. K. K.; FERREIRA, G. I. S.; FLORES, H. F.; BITTENCOURT, I. C.; DIAS, M. L.; VAN DER LAAN, R. H. **Glossário da área de organização e tratamento da informação**. Disponível em: <http://www.ufrgs.br/fabico/documentos-graduacao-e-comgrads/glossario-de-ciencia-da-informacao>. Acesso em: 31 jul. de 2018.

[77] TSENG, Y. H.; LIN, C. J.; LIN, Y. I. Text mining techniques for patent analysis. **Information Processing and Management**, vol. 43, p. 1216–47, 2007. DOI 10.1016/j.ipm.2006.11.011.

[78] MAYYAS, A.; QATTAWI, A.; OMAR, M.; SHAN, D. Design for sustainability

in automotive industry: A comprehensive review. **Renewable and Sustainable Energy Reviews**, vol. 16, p. 1845–62, 2012. DOI 10.1016/j.rser.2012.01.012.

[79] AYDIN, H.; ESSADIQI, E.; JUNG, I. H.; YUE, S. Development of 3rd generation AHSS with medium Mn content alloying compositions. **Materials Science and Engineering A**, vol. 564, p. 501–8, 2013. DOI 10.1016/j.msea.2012.11.113.

[80] RUTH HEUSS; MÜLLER, N.; SINTERN, W. Van; STARKE, A.; TSCHIESNER, A. **Lightweight, heavy impact - How carbon fiber and other lightweight materials will develop across industries and specifically in automotive**. McKinsey & Company. Visual Media Europe. 2012. 24 p.

[81] KIM, H.; SUH, D. W.; KIM, N. J. Fe-Al-Mn-C lightweight structural alloys: A review on the microstructures and mechanical properties. **Science and Technology of Advanced Materials**, vol. 14, 2013. DOI 10.1088/1468-6996/14/1/014205.

[82] RANA, R.; LAHAYE, C.; RAY, R. K. Overview of Lightweight Ferrous Materials: Strategies and Promises. **Jom**, vol. 66, p. 1734–46, 2014. DOI 10.1007/s11837-014-1126-5.

[83] KUZIYAK, R.; KAWALLA, R.; WAENGLER, S. Advanced high strength steels for automotive industry. **Archives of Civil and Mechanical Engineering**, Committee of Civil and Mechanical Engineering, vol. 8, p. 103–17, 2008. DOI 10.1016/S1644-9665(12)60197-6.

[84] WORLD AUTO STEEL. **Advanced High-Strength Steel (AHSS) Definitions**. 2020. Disponível em: <https://www.worldautosteel.org/steel-basics/automotive-advanced-high-strength-steel-ahss-definitions/>. Acesso em 20 out. 2020.

[85] CHEN, S.; RANA, R.; HALDAR, A.; RAY, R. K. Current state of Fe-Mn-Al-C low density steels. **Progress in Materials Science**, vol. 89, p. 345–91, 2017. ISSN 00796425. DOI 10.1016/j.pmatsci.2017.05.002.

[86] RENNÓ, M. ; LEMGRUBER, N. O ciclo de vida de um empreendimento de dutos. *In*: FREIRE, J.L. F. (Org.). **Engenharia de Dutos**. Rio de Janeiro: Associação Brasileira de Engenharia de Ciências Mecânicas, 2009. p. 1.1-1.20.

[87] BOTT, I. ; SICILIANO, F.; LEMOS, P. F. .; RIOS, P. R. . Aços para Dutos. *In*:

FREIRE, J. L. F. (Org.). **Engenharia de Dutos**. Rio de Janeiro: Associação Brasileira de Engenharia de Ciências Mecânicas, 2009. p.4.1-4.18.

[88] AMERICAN PETROLEUM INSTITUTE. **API 5L: Specification for Line Pipe**. 45 ed. Washington Dc: API Publishing Services, 2012.

[89] OLIVEIRA, B. S. **Indicadores de patentes para o monitoramento tecnológico em aços para dutos utilizados em transportes de hidrocarbonetos**. 2016. 93 f. Dissertação (Mestrado) - Curso de Ciência e Engenharia de Materiais, Universidade Federal de São Carlos, São Carlos, 2016

[90] MILANEZ, D. H.; MILANEZ, M. G.; AMARAL, R. M. Do; FARIA, L. I. L. De; GREGOLIN, J. A. R. The Earliest Priority Selector for compiling patent indicators. **14th International Society of Scientometrics and Informetrics Conference**, vol. 6, p. 1877–80, 2013.

[91] OCDE - Organisation for economic co-operation and development. **The Measurement of Scientific and Technological Activities: Using Patent Data as Science and Technology Indicators**. 1994. 108 p. (n. 94). ISBN 9789264065574.

[92] OCDE - Organisation for economic co-operation and development. **OECD Patent Statistics Manual**., 2009. 162 p. ISBN 9789264054127.

[93] ANDRADE, M. A.; OLIVEIRA, B.S.; MILANEZ, D. H.; LEIVA, D. R.; ZEPON, G. *In: Materials Science & Technology*. 2018. Anais [...]. Columbus, OH, Estados Unidos da América, 2018. p. 341-347.

[94] HILLENBRAND, H.; LIESSEM, A.; BIERMANN, K.; HECKMANN, C. J.; SCHWINR, V. Development of grade X120 pipe material for high - pressure gas transportation lines. *In: International Conference on Pipeline Technology*. 4. 2004. **Anais [...]** Ostend - Bélgica. p.12.

[95] RANA, R.; LIU, C.; RAY, R. K. Low-density low-carbon Fe-Al ferritic steels. **Scripta Materialia**, vol. 68, p. 354–9, 2013. DOI 10.1016/j.scriptamat.2012.10.004.

[96] SPEER, J. G.; RIZZO ASSUNÇÃO, F. C.; MATLOCK, D. K.; EDMONDS, D. V. The “quenching and partitioning” process: Background and recent progress. **Materials Research**, vol. 8, p. 417–23, 2005. DOI 10.1590/S1516-14392005000400010.

## APÊNDICE A LISTA DE MARCADORES E TESAuros DE COMPOSIÇÃO

### QUÍMICA

A seguir, são apresentados os tesauros elaborados e utilizados nos filtros desenvolvidos no software VantagePoint para a recuperação de informação nos registros de patentes em aços.

Quadro A.1 – Marcadores utilizados para a seleção de texto nos filtros de recuperação das informações de teores de elementos químicos em aços nos registros de patentes.

Comprise   comprising   contains   containing   includes   include   including   consist   consisting   composed of   following ratio   involves using   prepared using   having composition   made from   following chemical components   formed by   following composition   prepared by   by taking   by preparing   following chemical composition   formed of   has components of   steel has   steel having   with weight percentages
---

Fonte: Elaboração do autor.

Quadro A.2 – Marcadores utilizados para os elementos de ligas em aços utilizados na elaboração da planilha.

carbono	(C c)arbon
manganês	(M m)anganese   (^  \ )(Mn)( - )
silício	(S s)ilicon   (^  \ )(Si)( - )
alumínio	(A a)luminum   (A a)luminium   (^  \ )(Al)( - )
cromo	(C c)hromium   100 1 (^  \ )(Cr)( - )
titânio	(T t)itanium   (^  \ )(Ti)( - )
nióbio	(N n)iobium   (^  \ )(Nb)( - )
vanádio	(V v)anadium

Fonte: Elaboração do autor.

Quadro A.3 – Expressões utilizadas para eliminação prévia à atribuição de marcadores.

Marcadores gerais de composição química	iron-niobium-carbon alloy comprises   raw material containing   raw materials containing   raw material comprising   raw materials comprising   comprises oxide particles   welding material comprises   chrome-base alloy steel is composed   Ferrochrome contains   Dephosphorizer further comprises   molten iron contains   molten iron comprises   molten iron comprising   chrome-molybdenum-nickel-tungsten alloy contains   scale comprises   wear-resistant layer   slag requirement comprises   degrees C   ceramic layer comprises   work layer contains   scrap material comprises   string involves   \, and includes   composition comprised (in ppm)   complement containing   complement consisting   ratio of volume   ratio of elements   inclusion containing   impurities compris   impurities     impurities cont   impurities consist   impurities has   further cont
---	---

	<p>         additionally cont   further compris   further I   structure compris   structure cont   rest I   rest consist   rest compris   rest cont   coating comprises   coating comprising   cermet layer comprises   has a structure compris   has a matrix structure compris   has a tensile   retained austenite compris   has yield   raw materials comprises   comprises raw materials   rare earth element compris   rare earth element cont   rare earth element I   rare earth element consist   rare earth element has   rare earth compris   rare earth cont   rare earth I   rare earth consist   rare earth has   one element including   also I   also cont   also has   cladding powder contains   also compris   comprises iron   contains iron   including iron   has iron   includes iron   comprising iron   comprises Fe   contains Fe   including Fe   has Fe   includes Fe   comprising Fe   iron I   composed of iron   remainder cont   remainder compris   remainder I   remainder cons   remnants cons   remnants cont   remnants I   remnants compris   contains the following components as impurities   comprises the following components as impurities   includes the following components as impurities   contains following substances as impurities   has composite   molten bath cont   one element cont   trace elements cont   has carbon equivalent   has carbon saturation   has a processing   comprises fine precipitate containing   solution containing   impurity cont   additionally includes   resin has   oxide particles compris   oxide layer compris   solution compris   fraction of the steel compris   includes a ferrite   when this steel comprises   includes pearlite   comprises prior-austenite   comprises rare-earth   comprises oxide   plate contains 1 area%   precipitate comprising   has surface comprising   comprises zinc   residual Fe including   impurities at following ratio   has iron-hafnium   has microstructure compris   where mixture includes   impurity has   impurity cont   impurity cons   impurity compris   impurity I   iron-based   oxide layer (A) comprises   rest at compliance to following ratios   such that the content of   The product of content   by the following relationships: at the ratio   at compliance to following ratios   comprises sulfide   comprised (in ppm) of   scrap contains   structure including   INDEPENDENT CLAIM is included   including typical steel   contains oxide   inclusions comprising   ore contains   ash contains   contains following substances as admistures   impurities it contains   has cold brittleness sensitivity   comprising ferrite   soleplate, comprises   magnesium chloride   Metallurgy roller material comprises   Refining slag comprises   comprised (in ppm)   has Md30   Cross-iron alloy comprises   further contain   further compris   further I   inoculant involves mixing   deoxidizer I   includes a structure cont   optionally compris   silicon zirconium alloy particle comprises   , including [0-9]   , excluding [0-9]   , not equal to [0-9]   , not including [0-9]   (including [0-9]   (excluding [0-9]   (not including [0-9]   self-lubricating material comprises   calcining manganese steel alloy comprising   interface layer comprises   ean gas zone comprises   hot metal component comprising   hot metal comprises   hot metal composition   coating layer compris   containing (in mass ppm   consisting of (in mass ppm   condition containing   raw material content in 440 grade steel compris   formed ingot I   smelting process contain   comprises (i) smelting   molten steel sample containing   high-magnesium lime comprises   comprises (i) taking   floating seal in the casting process cont   cementite comprising   molten iron compris   containing molten iron   comprises (in pts.wt.   substrate contains   deoxidizing agent compris   sulfur particles consist   ferro chrome comprises   oxides (per mm<sup>2</sup>) contain   depth consist of   film contains   contains oxides   oxygen compris   oxygen 98letrôn   modifier compris   modifier cont   modifier I   modifier consist   The raw material elements have following relationship including   ink comprise   inserting alloy wire compris   alloy wire compris   coating is compris   composite metamorphic particles compris   contains       </p>
--	---

	<p>fine precipitate   slag composition cont   has structure cont   alloy raw material of the metal shelf is compos   inclusion compris   composite inclusion cont   Slab steel material contain   comprises \in parts weight   comprises a core layer   ferrosilicon alloy compris   boron iron alloy compris   for alloying steel conta   comprises a chromium passivation layer   passivation layer comprises a base material including   smelting low alloy steel compris   insulation film compris   inorganic component cont   composite iron alloy cont   coolant compris   core wire compris   refining agent compris   ferrite single phase and I   precipitates cont   direction compris   modifying agent compris   rare earth magnesium alloy powder compris   rare earth alloy composite core wire compris   solution composition compris   alloy particles compos   substance compris   comprise substance   microstructure cont   outer rings contain   99letrônico steel has   surface cont   welding wire\, comprises \i\   dust comprises   aluminate slag comprises   oxide dispersion-strengthened ferrite steel has   Ferroboron-based alloy-containing durable material comprises   unprocessed steel comprises   hromium passivation layer comprises a base material including   42-chromium-molybdenum alloy comprising   additive compris   steelmaking steel comprises   insulation film contain   semi-steel component compris   containing molten iron   calcium-based composite particles comprises   zirconium iron comprises   composite deoxidizer comprises  deoxidizing agent comprises   desulfurizing agent comprises   mother liquor containing</p>
Carbono	<p>nanocarbon   total carbon present   equivalent carbon   carbon equivalent   carbon equivalence   carbon potential   carbon scrap   carbonate   carbonized   carbon index   carbon disulfide   carbon fiber   carbon-free   free carbon   slag-line brick comprises   center portion contains   carbon content of surface layer   carbonitriding   99letrônicos99   99letrônicos9999   carbon-based   carbon steel at ratio   carbon compound   carbon monoxide   carbonyl   water carbon   carbon ferrochrome   until carbon content   carbon dioxide   carbon distribution ratio   C02   carbon concentration   ^CO\$   carbon oxide   carbon powder   carbon sulfide   carbon black   carbon agent   carbon distribution   carbon concentration distribution   carbonyl   decarbonization   carbonate   carbon segregation   carbon saturation   trace elements containing carbon, molybdenum, boron and cobalt   carbon concentration ratio   ratio of content \T\ with respect to amount of carbon   content of vanadium and carbon is   drawing \13\ carbon   high-carbon steel board has the following structure   iron-niobium-carbon alloy comprises   low carbon steel plate \5\   amount of carbon in steel sheet   non-carbon non-hydrogen element contains   carbon in the weight ratio   The Rockwell hard of high-carbon   carbon compound   metal carbon   titanium\carbon   4\carbon + nitrogen   10\carbon + nitrogen   sum of carbon and manganese\6   excluding carbon   niobium and carbon linked by relationship   carbon activity   carbon ferrochrome comprises   cementite comprising   carbon compound \C80\   ferrochromium mother solution contains   tapping carbon   decarbonizing   iron-niobium-carbon alloy comprises   \0.07 + titanium\4\ wt.% carbon   ccarbon divide by   carbon aluminum   carbon manganese</p>
Manganês	<p>manganese powder   manganese dioxide   manganese deoxidize   manganese/chromium   manganese concentration   oxide of manganese   manganese in martensite   solution of manganese   65-manganese-type   manganese ore   phosphorus-manganese   manganese oxide   manganese \II\ oxide   manganese\II\ sulfide   manganese sulfide   Mn sulfide   ferromanganese   ferromanganese iron comprises   manganese alloy   silicon-manganese   silicon manganese   manganese sulfate   manganese/6   manganese ore   manganese/sulfur   manganese nitride   manganese carbide   manganese carbonate   chromium-manganese-</p>

	nickel-nitride   sulfur-manganese ratio   sulfur with respect to manganese   manganese metal   manganese segregation portion of the microstructure comprises   The area ratio of nitride including silicon and manganese   titanium-calcium-manganese-oxygen-sulfur composition compris   sum of carbon and manganese $\sqrt{6}$   manganese and silicon in a weight ratio   manganese and aluminum in the steel   chromium-manganese-titanium waste steel   heating high-silicon high-manganese   aluminum-magnesium-titanium-calcium-manganese-oxygen-sulfur inclusions comprises   silicon-manganese alloy comprises   manganese divide by   carbon manganese   silicon-manganese alloy comprises
Silício	silicon oxide   silicon dioxide   silicon carbide   silicon nitride   silicon powder   Si oxide   SiO <sub>2</sub>   silicon alloy   oxide of silicon   silicon concentration   silica   silicone   ferrosilicon   metal silicon   siliconizing   desiliconization   bulk silicon   nano-silicon   silicon monoxide   siliconized   siliconizing   silicon segregation   silicon-aluminum-calcium deoxidizing agent comprises   The area ratio of nitride including silicon and manganese   silicon compound   ferrosilicon alloy cont   silicon-containing   calcium-based composite particles compris   ferrosilicon powder compris   manganese and silicon in a weight ratio   of silicon and aluminum   ratio of silicon to oxygen   heating high-silicon high-manganese   silicon-manganese alloy comprises   aluminum-silicon-calcium-barium alloy that comprises   silicon divide by   silicon-manganese alloy comprises   ferro-silicon powder
Alumínio	acid melting aluminum   100let aluminum   aluminum ore   aluminum 100letr   aluminum oxide   aluminum trioxide   aluminum nitride   aluminum carbide   aluminum sulfide   aluminum monosulfide   aluminum sulfate   Al <sub>2</sub> O <sub>3</sub>   alumina   aluminum alloy   aluminum alloy comprises   ferroaluminum   aluminum titanate   aluminum hydroxide   aluminum carbonate   aluminum phosphate   aluminum deoxidation   aluminum alloy in composite comprises   deoxidizing aluminum   aluminum killed steel   aluminum block containing   material contains 100 or more aluminum   deoxidizing agent comprising aluminum   Aluminum ingot   adding aluminum   deoxidizing by adding aluminum   transparent aluminum   Aluminium oxynitride   aluminum aluminum powder ball   adding 1% aluminum of the liquid steel   alloy wire comprises   ferroaluminum alloy contains   silicon-aluminum-calcium deoxidizing agent comprises   aluminum pellet   aluminum line   aluminum-boron-chromium   sum of aluminum and titanium   of silicon and aluminum   aluminum-free deoxidizer   pure aluminum   aluminum carbonate   manganese and aluminum in the steel   Aluminum wire   aluminum-magnesium-titanium-calcium-manganese-oxygen-sulfur inclusions comprises   aluminum-silicon-calcium-barium alloy that comprises   aluminum divide by   Als/Al <sub>t</sub>   aluminum-iron alloy comprises   deoxidant is aluminum   carbon aluminum   purifying agent comprises   aluminumoxide



Cromo	chromium equivalent   equivalent chromium   ferrochromium   ferrochrome   chromium oxide   chromium oxide   Cr2O3   chromium nitride   chromium(III) oxide   chromium hydroxide   chromium powder   chromium carbide   chromium alloy   oxide of chromium   chromium concentration   chromium type   chromium ore   chromium-based carbide   chromium segregation   to chromium/iron   chromium cast iron scrap   excluding chromium   chrome residues   chromium-manganese-nickel-nitride   chromium concentration   aluminum-boron-chromium   CR rolled   42chromium-molybdenum   40 Cr steel   85 chromium   high-chromium-based alloy powder comprises   molybdenum+vanadium+chromium)/5   chromium, molybdenum and vanadium/5   ferrochrome alloy containing   chromium content of the chromium-rich phase   niobium is replaced with 1 pts.wt. of vanadium, nickel, molybdenum and chromium   Cr-containing Nb   carbon ferrochrome comprises   chromium ore comprises   oxide layer containing chromium   chromium-nickel-molybdenum-strontium complex   chromium-manganese-titanium waste steel   ferrochromium mother solution contains   chromium divide by   ferrochrome alloy containing   12-chromium-nickel   20 chromium alloy steel
Titânio	titanium oxide   titanium dioxide   titanium carbide   titanium boron   titanium alloy   TiC   TiB2   titanium powder   oxide of titanium   titanium diboride   titanium ore   titanium nitride   ferrotitanium   titanium segregation   niobium (Nb) and titanium (Ti)   niobium-titanium carbonitrides   nano titanium   titanium-4(carbon+nitrogen)   titanium/carbon   deposition rate of titanium   titanium(III) oxide   titanium-type composite inclusion   titanium having a mass ratio   titanium/nitrogen   effective titanium   nitrides of titanium   titanium is not more than 10(carbon + nitrogen)   titanium pigment   titanium iron   effective value of titanium   sum of aluminum and titanium   titanium-calcium-oxygen inclusions   titanium-calcium-manganese-oxygen-sulfur composition compris   titanium-iron   ferrotitanium powder contains   chromium-manganese-titanium waste steel   aluminum-magnesium-titanium-calcium-manganese-oxygen-sulfur inclusions comprises   titanium compound by oxide   parts of titanium in relation to   titanium divide by   titanium white powder   titanium iron comprises   ferrotitanium comprising
nióbio	niobium oxide   niobium carbide   NbC   niobium alloy   niobium powder   niobium concentration   niobium type   ferroniobium   niobium carbonitride   niobium segregation   niobium-containing carbide   niobium containing carbide   niobium nitride   niobium precipitate   iron-niobium-carbon alloy comprises   sum (B value) of nitrogen and niobium   niobium in the from of precipitate   Vanadium is partially or completely substituted with niobium   niobium is replaced with 1 pts.wt. of vanadium, nickel, molybdenum and chromium   with 5/9 niobium (Nb)   vanadium+5/9 niobium   niobium (Nb) and titanium (Ti)   Nb observed   ((V)+(Nb))/2   Cr-containing Nb   molybdenum, vanadium and niobium   niobium-titanium carbonitrides   ferro columbium containing niobium   temperature of niobium   niobium and carbon linked by relationship   nitrogen/niobium   iron-niobium-carbon alloy comprises   niobium divide by
vanádio	vanadium oxide   vanadium carbide   vanadium alloy   vanadium powder   vanadium concentration   vanadium type   ferrovanadium   vanadium nitride   vanadium slag   vanadium precipitate   content of vanadium and carbon is   molybdenum+vanadium+chromium)/5   Vanadium is partially or completely substituted with niobium   niobium is replaced with 1 pts.wt. of vanadium, nickel, molybdenum and chromium     ((V)+(Nb))/2   molybdenum, vanadium and niobium   vanadium+5/9 niobium   contents of V and N satisfy   vanadium steel with   (V gamma   (V alpha   ^eV\$   vanadium nitrogen alloy   sum of chromium, molybdenum and vanadium/5   vanadium-nitrogen alloy   vanadium divide by   vanadium precipitation amount

Quadro A.4 – Marcadores utilizados para limites superiores e inferiores dos elementos de liga e propriedades.

<	or less than   not greater than   less than   or less   no more than   not more than   up to   lesser than   up to   maximum   at most   traces to   at least
>	more than   more than   not less than   or more   minimum

Fonte: Elaboração do autor.

**APÊNDICE B LISTA DE MARCADORES E TESAuros DE PROCESSOS,  
MICROESTRUTURAS, PROPRIEDADES E APLICAÇÕES**

Quadro B.1 – Expressões utilizadas para eliminação prévia à atribuição de marcadores das propriedades.

Dureza	hardness deviation   hardness increased by   hardness difference is within   hardness difference less than   hardness steel can be raised   hardness deviation   softening rate of less than 13 HRC
Alongamento	elongation rate   yield elongation   yield point elongation   yielding point elongation   yielding elongation   creep elongation   strip elongation   distributed elongation   elongation Ae   leveling elongation   specific elongation   drawing elongation   specific elongation   stretch elongation   0\2% elongation   0\5% elongation   0\1% elongation   MPa%   MPa %   Mpax%   Mpax %   MPa\.%   elongation increased   high elongation   yield strength-elongation balance   high-elongation   bending-stretching machine is performed with elongation   relation between break elongation   A80   calculated as elongation   local elongation   ratio of elongation   plastic elongation strength   sigma 4   delta 3   delta 5   A50   at maximum load   having ratio of strength and elongation of
Densidade	piece/cm3   dislocation density   density of precipitate   loose density   density of oxide particles   average number density of   density of micro   number density
Limite de escoamento e limite de resistência à tração	(YS/UTS) ratio   ratio of yield strength   yield strength/tensile strength ratio   anisotropy of the 0.2% yield strength   ratio of yield strength and tensile strength   tensile strength increased   yield strength increased   high-temperature tensile strength   difference of tensile strength   yield strength variation   tensile strength variation   tensile strength unevenness   yield strength-elongation balance   increases tensile strength for about   yield strength controlled   yield strength of the same steel coil differs by less than   compressive yield strength   compressive yield strength   film having tensile strength   winding-up tensile strength   difference between the maximum value and the minimum value among the tensile strength

Fonte: Elaboração do autor.

Quadro B.2 – Substituições de denominações da norma API 5L para valores de Limite de Escoamento e Limite de Resistência à Tração.

Denominação	LE	LRT
X42 ou X 42	290 MPa	655 MPa
X46 ou X 46	320 MPa	655 MPa
X52 ou X 52	360 MPa	760 MPa
X56 ou X 56	390 MPa	760 MPa
X60 ou X 60	415 MPa	760 MPa
X65 ou X 65	450 MPa	760 MPa
X70 ou X 70	485 MPa	760 MPa
X80 ou X 80	555 MPa	825 MPa
X90 ou X 90	625 MPa	915 MPa
X100 ou X 100	690 MPa	990 MPa
X120 ou X 120	830 MPa	1145 MPa
80ksi ou 80 ksi	552 MPa	552 MPa
110ksi ou 110 ksi	758 MPa	758 MPa
125ksi ou 125 ksi	862 MPa	862 MPa

Fonte: Elaboração do autor.

Quadro B.3 – Expressões utilizadas no tesauro de termos de processos.

**@Laminacao controlada#	100 1 gas circulation
100 1 controlled.rolling	**@Nitruracao#
**@Laminacao a quente#	100 1 Nitriding
100 1 hot.rolling	**@Carburizacao#
100 1 hot.rolled	100 1 Carburising
**@Laminacao a frio#	**@Carbo nitruracao#
100 1 cold.rolling	100 1 Carbo.nitriding
100 1 cold.rolled	**@Desoxidacao#
**@Endurecimento Superficial#	100 1 deoxidising
100 1 surface hardening	**@Desforizacao e Desulfurizacao#
100 1 surface hardened	100 1 Dephosphoris
**@Tempera Interrompida#	100 1 Desulfuris
100 1 interrupted quenching	**@Descarburizacao#
100 1 interrupting quenching	100 1 Decarburis
100 1 interrupted quenched	**@Refino Termico ou Verguten#
**@Tempera Isotermica#	100 1 heat refin
100 1 isothermal quenching	100 1 Verguten
100 1 isothermal quenched	**@Estampagem profunda#
**@Martempera#	100 1 deep.draw
100 1 martempering	**@Refusao de eletroescoria#
100 1 martempered	100 1 Electroslag remelting
**@Normalizacao#	**@Endurecimento por precipitacao#
100 1 normalizing	100 1 hardening by precipitation
100 1 normalized	100 1 precipitation hardening
**@Alivio de Tensoes#	**@Trabalho a quente#
100 1 stress.relieving	100 1 hot.worked
100 1 stress.relieved	100 1 hot.working
**@Metalurgia do po#	100 1 hot.press
100 1 powder metallurgy	100 1 hot.stamp
100 1 metallic powder	**@Trabalho a frio#
100 1 powder metallurgically	100 1 cold working
100 1 powder steel	100 1 cold worked
**@Estampagem profunda#	100 1 shot.peening
100 1 deep.draw	100 1 burnishing
**@Pressao reduzida ou 105letr#	100 1 cold expansion
100 1 vacuum	
100 1 reduced pressure	
**@Atmosfera controlada#	
100 1 inert gas	
100 1 controlled atmosphere	
100 1 pulverulent material	

Fonte: Elaboração do autor.

## Quadro B.4 – Códigos da CIP utilizados no tesouro de processos.

**@Laminacao Controlada#	100 1 C21C-001/02
100 1 C21D-008/00	**@Descarburizacao#
100 1 C21D-008/02	100 1 C21C-007/068
100 1 C21D-008/04	**@Refino Termico ou Verguten#
100 1 C21D-008/06	100 1 C21D-001/25
100 1 C21D-008/08	**@Laminacao a quente#
100 1 C21D-008/10	100 1 B21B-001/26
100 1 C21D-008/12	**@Laminacao a frio#
**@Endurecimento Superficial#	100 1 B21B-001/28
100 1 C21D-001/06	**@Estampagem profunda#
100 1 C21D-001/08	100 1 C21D-009/48
100 1 C21D-001/09	**@Atmosfera controlada#
100 1 C21D-001/10	100 1 C21D-001/74
**@Tempera Interrompida#	100 1 ^C21D-001/76\$
100 1 C21D-001/19	100 1 ^C21D-001/767\$
**@Tempera Isotermica#	**@Nitruacao#
100 1 C21D-001/20	100 1 C23C-008/24
**@Martempera#	100 1 C23C-008/26
100 1 C21D-001/22	**@Carburizacao#
**@Normalizacao#	100 1 C23C-008/20
100 1 C21D-001/28	100 1 C23C-008/22
**@Alivio de Tensoes#	**@Carbo nitruacao#
100 1 C21D-001/30	100 1 C23C-008/30
**@Pressao reduzida ou 106letr#	100 1 C23C-008/32
100 1 C21C-007/10	**@Trabalho a quente#
100 1 ^C21D-001/773\$	100 1 C21D-007/13
**@Desoxidacao#	**@Trabalho a frio#
100 1 ^C21C-007/06\$	100 1 C21D-007/02
**@Metalurgia do po#	100 1 C21D-007/04
100 1 C22C-033/02	100 1 C21D-007/06
100 1 B22F-001/00	100 1 C21D-007/08
100 1 B22F-001/02	100 1 C21D-007/10
100 1 B22F-003/24	100 1 C21D-007/12
**@Desfoforizacao e Desulfurizacao#	**@Refusao de eletroescoria#
100 1 C21C-007/064	100 1 C22B-009/18

Fonte: Elaboração do autor.

Quadro B.5 – Expressões utilizadas no tesouro de termos de aplicações.

**@molas#	**@Eletrodomestico#	100 1 teeth for card-clothing
100 1 spring	100 1 household appliance	**@Laminas de patins#
**@trilhos#	100 1 domestic appliance	100 1 blades for skates
100 1 rails	100 1 domestic electrical appliance	**@Mancais e rolamentos#
**@Aneis#	100 1 kitchen equipment	100 1 bearing
100 1 ^ring\$	**@barcos e navios#	**@Transformadores#
**@Canos de espingarda#	100 1 ^ship\$	100 1 transformer
100 1 shotgun barrel	100 1 ^ships\$	100 1 iron core
**@Tubos 107letrôn#	100 1 shipbuilding	**@Automovel#
100 1 rigid pipe	100 1 ^boat\$	100 1 ^motor vehicle\$
**@Dutos#	100 1 ^boats\$	100 1 ^motor-vehicle\$
100 1 pipeline	**@trocador de calor#	100 1 ^motor vehicles
100 1 pipe.line	100 1 heat exchanger	100 1 automobile
100 1 linepipe	**@virabrequim e cames#	**@Automovel 107letrôni#
100 1 line.pipe	100 1 crankshaft	101 1 electric vehicle
100 1 API 5L	100 1 camshaft	**@Automovel hibrido#
100 1 X42	**@estruturas marinhas#	100 1 hybrid vehicle
100 1 X 42	100 1 marine structure	**@Autopecas#
100 1 X46	100 1 marine structural	100 1 autoparts
100 1 X 46	**@motor a combustao#	100 1 motor parts
100 1 X52	100 1 combustion engine	100 1 ^motor-vehicle component\$
100 1 X 52	100 1 combustion motor	100 1 motor vehicle component\$
100 1 X56	**@celula combustivel#	100 1 automobile components
100 1 X 56	100 1 fuel cell	100 1 vehicle body parts
100 1 X60	**@roda de engrenagem#	**@Equipamentos Industriais#
100 1 X 60	100 1 gearwheel	100 1 industrial machine
100 1 X65	100 1 gear wheel	**@Construcao civil#
100 1 X 65	**@motor 107letrôni#	100 1 bridge
100 1 X70	100 1 eletric motor	100 1 civil engineering
100 1 X 70	**@tanque de armazenagem#	100 1 civil construction
100 1 X80	100 1 storage tank	100 1 building
100 1 X 80	**@processamento quimico#	**@Maquina de 107letrônico#
100 1 X90	100 1 chemical process	100 1 construction machine
100 1 X 90	**@mineracao#	**@Vazos de 107letrôn#
100 1 X100	100 1 mining	100 1 pressure vessel
100 1 X 100	**@caldeira#	**@Evaporadores#
100 1 X120	**@Eixo simples#	100 1 boiler
100 1 X 120	100 1 plain shaft	**@Moldes e matrizes#
**@Tubos 107letrônico#	**@Aros e pneus#	100 1 mold steel
100 1 flexible pipe	100 1 tyre	100 1 die steel
**@Obuzes explosivos#	100 1 ^rim\$	100 1 cold press die
100 1 explosive shell	**@Ferramentas de corte#	100 1 cold forming
**@Ferramentas de corte manuais#	100 1 knife	**@Laminador#
100 1 knife	100 1 scythe	100 1 steel roller
100 1 scythe	100 1 scissor	100 1 Material for hot rolling
100 1 scissor	100 1 hand cutting tool	**@Ferramentas#
100 1 hand.cut.tool	100 1 milling cutter	100 1 high speed steel
**@Cabos#	100 1 machining cutter	100 1 tool steel
100 1 cable	100 1 saw blade	100 1 steel tool
100 1 cables	**@Chapas de blindagem#	100 1 mahining tool
**@lmas ou corpos 107letrônico#	100 1 armour plate	100 1 machine tool
100 1 ^magnets\$	**@Aguilhas e dentes de bicos de cardas#	**@Acos 107letrônico#
100 1 magnetic body	100 1 needle	100 1 electrical steel
100 1 magnetic bodies		
**@Eletrodos de solda#		
100 1 welding electrode		
100 1 welding wire		

Fonte: Elaboração do autor.

Quadro B.6 – Códigos da CIP utilizados no tesauro de aplicações.

**@Molas#	100 1 C21D-009/22	**@Vigas#
100 1 C21D-009/02	100 1 C21D-009/24	100 1 B23K-101/28
**@Trilhos#	**@Fios e tiras#	**@Eixos simples#
100 1 C21D-009/04	100 1 C21D-009/52	100 1 C21D-009/28
100 1 C21D-009/06	100 1 B23K-101/32	**@Esferas e roletes#
100 1 B23K-101/26	**@Imas ou corpos	100 1 C21D-009/36
**@Canos de	108letrônico#	**@Eixos de manivela e
espingarda#	100 1 H01F-001/[0-9]*	comes#
100 1 C21D-009/10	**@Eletrodos de solda#	100 1 C21D-009/30
**@Canos de artilharia#	100 1 B23K-035/[0-9]*	**@Aros e pneus#
100 1 C21D-009/12	**@Rodas de	100 1 C21D-009/34
**@Tubos 108letrôn#	engrenagem#	**@Chapas de
100 1 C21D-009/14	100 1 C21D-009/32	blindagem#
100 1 F16L-009/02	**@Paineis de chapa#	100 1 C21D-009/42
100 1 F16L-009/04	100 1 B23K-101/18	**@Agulhas e dentes de
100 1 F16L-009/06	**@Trocadores de calor#	bicos de cardas#
100 1 B23K-101/06	100 1 B23K-101/14	100 1 C21D-009/26
100 1 B23K-101/10	**@Vasos de pressão#	**@Laminas de patins#
100 1 B23K-101/04	100 1 B23K-101/12	100 1 C21D-009/20
100 1 B23K-101/08	**@Tiras ou folhas#	**@Mancais e
100 1 ^B23C-037/08\$	100 1 B23K-101/16	rolamentos#
100 1 B23C-037/083	**@Dispositivos	100 1 F16C-033/3[0-9]*
**@Obuzes explosivos#	108letrônicos#	100 1 F16C-033/4[0-9]*
100 1 C21D-009/16	100 1 B23K-101/36	100 1 F16C-033/5[0-9]*
**@Ferramentas de	100 1 B23K-101/38	100 1 F16C-033/6[0-9]*
corte#	**@Estruturas alveolares#	100 1 F16C-033/7[0-9]*
100 1 C21D-009/18	100 1 B23K-101/02	100 1 F16C-033/8[0-9]*

Fonte: Elaboração do autor.



Quadro B.7 – Expressões utilizadas no tesauro de termos de microestrutura.

**@Martensita#	**@Perlita#
100 1 martensitic	100 1 pearlitic
100 1 martensite.type	100 1 pearlite.type
100 1 martensite.steel	100 1 pearlite.steel
**@Austenita#	**@Duplex#
100 1 austenitic	100 1 Austenite.ferrite
100 1 austenite.stainless steel	100 1 ferrite.austenite
100 1 austenite.type	100 1 duplex cast steel
100 1 austenite.steel	100 1 duplex steel
**@Ferrita#	100 1 duplex.stainless.steel
100 1 ferritic	100 1 dual.phase
100 1 ferrite.stainless steel	100 1 duplex structure
100 1 ferrite.type	100 1 two.phase
100 1 ferrite.based	**@Martensita revenida#
100 1 ferrite.base	100 1 tempered.martensite
100 1 ferrite.steel	**@Austenita retida#
**@Bainita#	100 1 retained.austenite
100 1 bainitic	100 1 residual.austenite
100 1 bainite.steel	
100 1 bainite.type	

Fonte: Elaboração do autor.



## **APÊNDICE C INSTRUÇÕES DETALHADAS DE ATUALIZAÇÃO DOS FILTROS E DA PLANILHA DE DADOS**

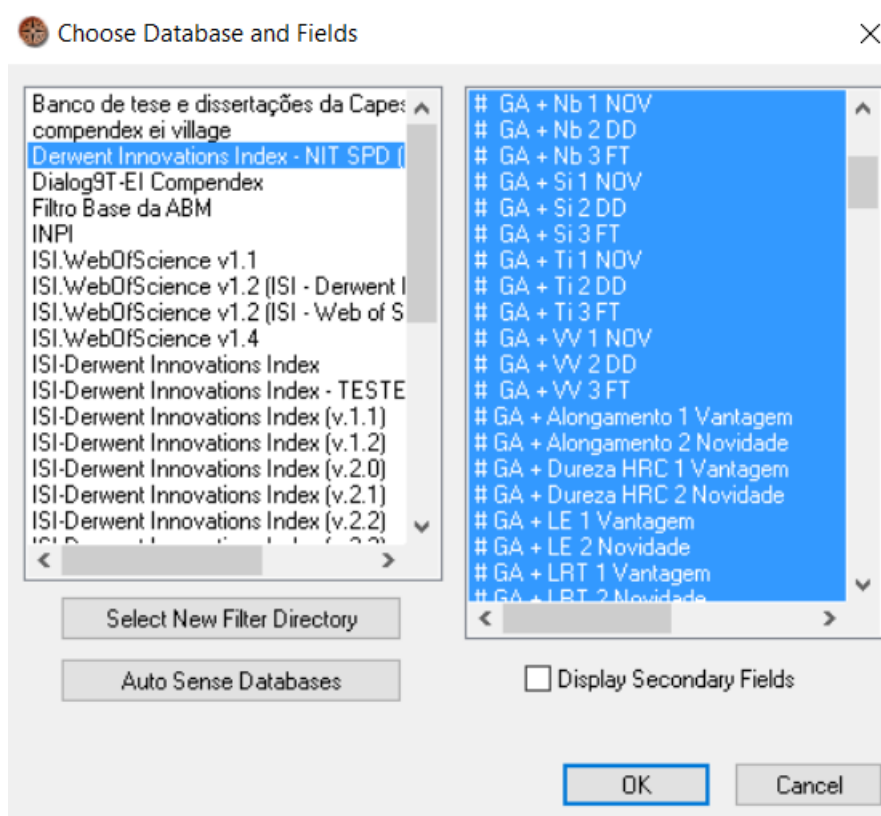
Este apêndice tem como objetivo acrescentar detalhes de operacionalização para futuras replicações do método utilizado no desenvolvimento desta tese, descritos na Metodologia. A intenção é de que seu conteúdo se torne um manual para que novas atualizações sejam realizadas por pesquisadores do NIT/Materiais e demais interessados, com domínio nos softwares VantagePoint e de um editor de planilhas, a exemplo do MSExcel, utilizado na pesquisa.

### **Importação dos registros no Software VantagePoint**

A versão do filtro desenvolvido nesta tese foi denominada “Derwent Innovations Index – NIT SPD (1.0)”, a partir da versão 5.0 do software VantagePoint. As futuras atualizações devem se atentar para que os tesouros estejam dentro das pastas correspondentes para suas leituras, caso contrário, mensagens de erro serão apresentadas. Por padrão foram definidas as seguintes pastas, considerando sua operação no Windows.

- Filtro: “Derwent Innovations Index – NIT SPD (1.0)”
- Local do filtro: c:/VantagePoint/Import Filters
- Local dos tesouros: c:/VantagePoint/Thesaurus

Os filtros para os campos necessários para alimentar a planilha de dados estão identificados com “#” à frente de sua denominação. Outros campos, de filtros e versões anteriores elaborados por pesquisas no NIT/Materiais foram mantidos, de forma a permitir a elaboração de indicadores também a partir desses campos. A Figura C.1 apresenta a tela inicial de importação de novos registros e para a escolha dos campos selecionados, já após a seleção dos arquivos do formato “.txt” recuperados da DII.



Fonte: Elaboração do autor.

Figura C.1 – Tela de escolha de filtros no software VantagePoint e de seus campos para a elaboração e atualização da planilha de dados.

Após a importação, são geradas as listas dos campos escolhidos e disponibilizados para consulta e futura transferência para planilhas do MS Excel. Para o trabalho a partir de registros já importados em outras versões de filtros da DII, a mudança pode ser realizada a partir do seguinte caminho:

- 1 – Acesso ao menu “File” do VantagePoint 5.0
- 2- Acesso à opção “Dataset properties”
- 3 – Acesso à opção “Change Dataset configurations”
- 4 – Acessar “replace database”
- 5 – Escolher o filtro “Derwent Innovations Index – NIT SPD (1.0)”

Os procedimentos apresentados nas próximas seções correspondem aos utilizados na elaboração da primeira versão da planilha, através da versão “Office 365” e sua atualização corrente em outubro de 2020.

### Migração de dados numéricos

Esta seção se refere à migração dos valores numéricos percentuais de elementos de liga e das propriedades. A Figura C.2 apresenta a lista elaborada para os teores de carbono obtidos a partir do campo “Novidade” dos registros. A operação de selecionar é simples, entretanto, devido ao volume de dados, seja esperada a demora de alguns segundos após o comando de “copiar” para o processamento por parte do VantagePoint (v.5.0).

The screenshot shows a software interface with a list of 22911 items on the left and a table of records on the right. The table has two columns: '# Records' and '# Instances'. The data in the table is as follows:

	# Records	# Instances
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1
10	1	1
11	1	1
12	1	1
13	1	1
14	1	1
15	1	1
16	1	1
17	1	1
18	1	1
19	1	1
20	1	1
21	1	1
22	1	1
23	1	1
24	1	1
25	1	1

Fonte: Elaboração do autor.

Figura C.2 – Exemplo de lista gerada no VantagePoint para os teores de carbono no campo novidade, com sua seleção total e comando de copiar.

A migração para o Excel pode ser executada normalmente a partir de uma nova planilha aberta e da função “colar”. Entretanto, sugere-se que a partir desse momento sejam utilizadas planilhas que não apresentam o valor “0” em células vazias. Esse procedimento pode ser executado a partir do seguinte caminho:

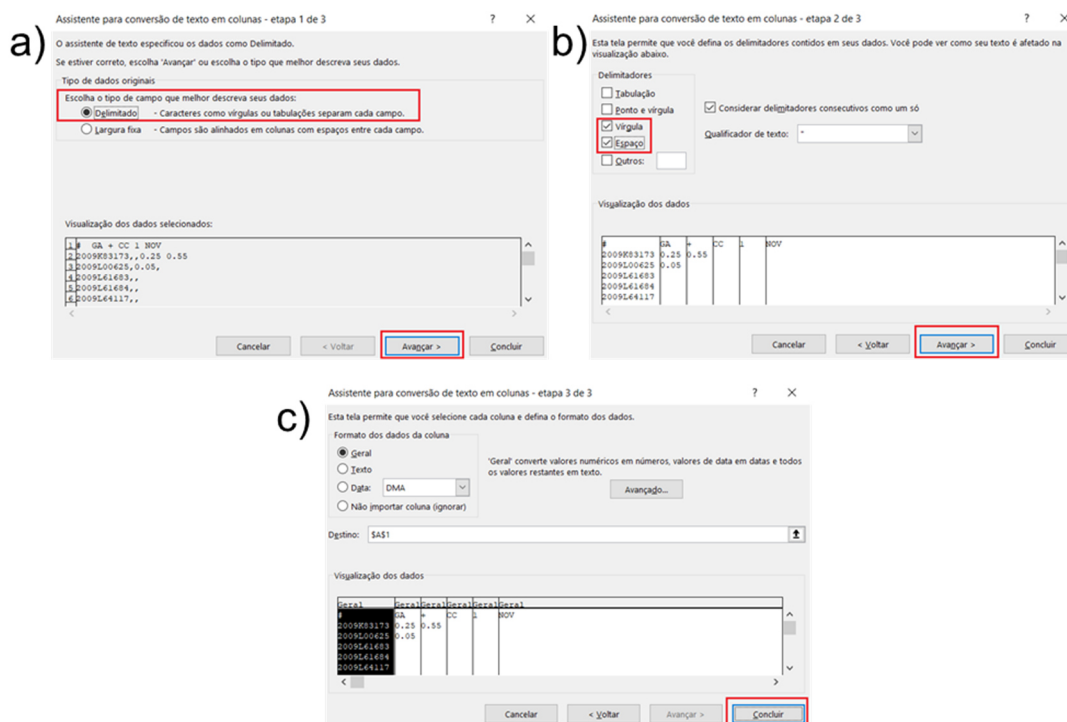
- Passo 1: Acessar o menu “Arquivo” do MSEXcel.
- Passo 2: Acessar o campo “Opções”
- Passo 3: Acessar o menu “Avançado”
- Passo 4: Buscar a opção “Mostrar um zero nas células cujo valor é zero” e certificar que esteja desmarcada.
- Passo 5: Clicar em “OK”.

A Figura C.3 apresenta a planilha de apoio após a transferência do VantagePoint. Sugere-se excluir as três primeiras colunas e trabalhar com as colunas a partir da listagem dos códigos GA. A próxima etapa consiste em dividir o texto a partir de suas vírgulas e espaços para que os valores numéricos (composição química ou valores das propriedades) sejam alocados corretamente nas colunas seguintes. Para tal, deve-se selecionar toda a primeira coluna e, a seguir, utilizar o comando “Texto para colunas” disponível na aba “Dados” do menu principal. A Figura C.4 apresenta a sequência de procedimentos e opções a serem escolhidas, enquanto a Figura C.5 mostra como deve ser apresentada a nova configuração da planilha após esse procedimento.

	A	B	C	D	E	F
1		# Records	# Instance	#	GA + CC 1 NOV	
2		1	1	1	2009K83173,,0.25 0.55	
3		2	1	1	2009L00625,0.05,	
4		3	1	1	2009L61683,,	
5		4	1	1	2009L61684,,	
6		5	1	1	2009L64117,,	
7		6	1	1	2009L64118,0.1,	
8		7	1	1	2009L64119,0.30 0.45,	
9		8	1	1	2009M15411,,	
10		9	1	1	2009M15818,< 0.12,	
11		10	1	1	2009M15819,,	
12		11	1	1	2009M15821,2.2 2.8,	
13		12	1	1	2009M15823,0.35 0.65,	
14		13	1	1	2009M15825,< 0.03,	
15		14	1	1	2009M15827,0.11 0.17,	
16		15	1	1	2009M15829,0.16 0.25,	
17		16	1	1	2009M35243,0.2 0.5,	
18		17	1	1	2009M35246,0.3 1.0,	
19		18	1	1	2009M41456,0.360 0.440,	
20		19	1	1	2009M41457,0.520 0.730,	
21		20	1	1	2009M56641,0.04 0.10,	

Fonte: Elaboração do autor.

Figura C.3 – Exemplo de lista após copiada e colada para o MSEXcel com destaque para as colunas a serem excluídas.



Fonte: Elaboração do autor.

Figura C.4 – Ilustração de procedimentos no menu "Texto para colunas".

2	2009K83173	0.25	0.55
3	2009L00625	0.05	
4	2009L61683		
5	2009L61684		
6	2009L64117		
7	2009L64118	0.1	
8	2009L64119	0.30	0.45
9	2009M15411		
10	2009M15818	<	0.12
11	2009M15819		
12	2009M15821	2.2	2.8
13	2009M15823	0.35	0.65
14	2009M15825	<	0.03
15	2009M15827	0.11	0.17
16	2009M15829	0.16	0.25
17	2009M35243	0.2	0.5
18	2009M35246	0.3	1.0
19	2009M41456	0.360	0.440
20	2009M41457	0.520	0.730
21	2009M56641	0.04	0.10

Fonte: Elaboração do autor.

Figura C.5 – Disposição das colunas na planilha de apoio após comando “Texto para colunas” para dados numéricos.

Pelo fato de os registros não apresentarem sempre o mesmo padrão de disponibilização das informações numéricas, e a primeira versão dos filtros desenvolvidos não cobrirem todas as exceções, algumas inconsistências ainda são apresentadas com valores além das colunas que seriam esperadas “B” e “C”. Assim, é necessária a verificação e exclusão manual desses valores. Sugere-se, como apresentado na Figura C.6, realizar essa verificação a partir de filtros de dados na coluna “F” (Menu Dados ->Filtros). As colunas “D” e “E” podem ser verificadas depois, visto que, em alguns casos, os valores que apresentam são provenientes dos caracteres “<” e “>” que foram utilizados para atribuição de limites máximos e mínimos de cada elemento.



	A	B	C	D	E	F	G	H	I	J	K
1											
2	2009K83173	0.25	0.55								
3	2009L00625	0.05									
4	2009L61683										
5	2009L61684										
6	2009L64117										
7	2009L64118	0.1									
8	2009L64119	0.30	0.45								
9	2009M15411										
10	2009M15818	<	0.12								
11	2009M15819										
12	2009M15821	2.2	2.8								
13	2009M15823	0.35	0.65								
14	2009M15825	<	0.03								
15	2009M15827	0.11	0.17								
16	2009M15829	0.16	0.25								
17	2009M35243	0.2	0.5								
18	2009M35246	0.3	1.0								
19	2009M41456	0.360	0.440								
20	2009M41457	0.520	0.730								
21	2009M56641	0.04	0.10								
22	2009M56647	0.05	0.16								
23	2009M56648	0.50	0.57								
24	2009M56650	3.0	5.5								
25	2009M56651	0.25	0.35								
26	2009M56652	0.25	0.35								
27	2009M56653										
28	2009M56654	0.05	0.10								
29	2009M81797	0.1	0.2								

Verificar!

Fonte: Elaboração do autor.

Figura C.6 – Exemplo de verificação manual de valores apresentados além da coluna “D”.

Após essa etapa de verificação, é esperado que sejam apresentados dados apenas nas colunas “A”, “B” e “C”. A partir desse momento, a importação dessas colunas deve ser realizada para a planilha de apoio “Ajustador limites”. O procedimento com esses casos será tratado no item seguinte.

### Uso da planilha de apoio “Ajustador de Limites”

Esta planilha de apoio foi desenvolvida para migrar, automaticamente, os valores máximos e mínimos dos elementos de liga para a coluna correspondente “máximo” ou “mínimo” quando acompanhadas dos caracteres “<” e “>”. Suas fórmulas se baseiam em condicionais “se” para que a mudança de colunas seja, ou não, realizada. Sua operação é bastante simples, bastando “copiar” e “colar

valores” nas colunas à esquerda, e buscar os resultados das colunas identificadas à direita, como identificado e demonstrado na Figura C.7. A migração para a próxima planilha também deve ser realizada a partir de “copiar” e “colar valores”, e recomenda-se verificar se a planilha de destino se encontra com a opção “Mostrar um zero nas células cujo valor é zero” esteja também desmarcada.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	COLE AQUI (ORIGEM)				PROCESSOS COM FORMULAS! NÃO ALTERAR								RESULTADOS			
2	# GA + CC 1 NOV															
3	2009K83173	0.25	0.55		0.25	0.55		0.25	0.55		0.25	0.55		2009K83173	0.25	0.55
4	2009L00625	0.05			0.05			0.05			0.05			2009L00625	0.05	
5	2009L61683													2009L61683		
6	2009L61684													2009L61684		
7	2009L64117													2009L64117		
8	2009L64118	0.1			0.1			0.1			0.1			2009L64118	0.1	
9	2009L64119	0.30	0.45		0.30	0.45		0.30	0.45		0.30	0.45		2009L64119	0.30	0.45
10	2009M15411													2009M15411		
11	2009M15818	<	0.12			0.12		<	0.12			0.12		2009M15818		0.12
12	2009M15819													2009M15819		
13	2009M15821	2.2	2.8		2.2	2.8		2.2	2.8		2.2	2.8		2009M15821	2.2	2.8
14	2009M15823	0.35	0.65		0.35	0.65		0.35	0.65		0.35	0.65		2009M15823	0.35	0.65
15	2009M15825	<	0.03			0.03		<	0.03			0.03		2009M15825		0.03
16	2009M15827	0.11	0.17		0.11	0.17		0.11	0.17		0.11	0.17		2009M15827	0.11	0.17
17	2009M15829	0.16	0.25		0.16	0.25		0.16	0.25		0.16	0.25		2009M15829	0.16	0.25
18	2009M35243	0.2	0.5		0.2	0.5		0.2	0.5		0.2	0.5		2009M35243	0.2	0.5
19	2009M35246	0.3	1.0		0.3	1.0		0.3	1.0		0.3	1.0		2009M35246	0.3	1.0
20	2009M41456	0.360	0.440		0.360	0.440		0.360	0.440		0.360	0.440		2009M41456	0.360	0.440
21	2009M41457	0.520	0.730		0.520	0.730		0.520	0.730		0.520	0.730		2009M41457	0.520	0.730
22	2009M56641	0.04	0.10		0.04	0.10		0.04	0.10		0.04	0.10		2009M56641	0.04	0.10
23	2009M56647	0.05	0.16		0.05	0.16		0.05	0.16		0.05	0.16		2009M56647	0.05	0.16
24	2009M56648	0.50	0.57		0.50	0.57		0.50	0.57		0.50	0.57		2009M56648	0.50	0.57
25	2009M56650	3.0	5.5		3.0	5.5		3.0	5.5		3.0	5.5		2009M56650	3.0	5.5
26	2009M56651	0.25	0.35		0.25	0.35		0.25	0.35		0.25	0.35		2009M56651	0.25	0.35
27	2009M56652	0.25	0.35		0.25	0.35		0.25	0.35		0.25	0.35		2009M56652	0.25	0.35
28	2009M56653													2009M56653		
29	2009M56654	0.05	0.10		0.05	0.10		0.05	0.10		0.05	0.10		2009M56654	0.05	0.10

Fonte: Elaboração do autor.

Figura C.7 – Exemplo de funcionamento da planilha de apoio Ajustador de Limites.

### Migração de dados de caracteres alfabéticos

Esta seção se refere à migração de dados do VantagePoint para o Excel para as listas de microestrutura, processos e aplicações. As etapas de geração das listas no VantagePoint, “cópia” e “cola” para o MSEXcel são os idênticos aos dos dados numéricos. A principal diferença se dá a partir da etapa “Texto para colunas” em que o caractere separador (Figura C.5b) deve ser apenas a vírgula. Diferente dos dados numéricos, para esses campos são 4 as colunas a serem trabalhadas: “A” (código GA), “B”, “C” e “D”, como demonstrado na Figura C.8.

	A	B	C	D
1	# GA+Processos 1 CIP			
2	2009K83173			
3	2009L00625			
4	2009L61683			
5	2009L61684			
6	2009L64117			
7	2009L64118			
8	2009L64119			
9	2009M15411			
10	2009M15818			
11	2009M15819			
12	2009M15821	Metalurgia do po		
13	2009M15823			
14	2009M15825			
15	2009M15827			
16	2009M15829	Laminacao Controlada		
17	2009M35243			
18	2009M35246			
19	2009M41456			
20	2009M41457			
21	2009M56641	Laminacao Controlada		
22	2009M56647	Desfoforizacao e Desulfurizacao	Pressao reduzida ou vacuo	Laminacao Controlada
23	2009M56648	Desfoforizacao e Desulfurizacao	Pressao reduzida ou vacuo	
24	2009M56650			
25	2009M56651	Desfoforizacao e Desulfurizacao	Pressao reduzida ou vacuo	Laminacao Controlada

Fonte: Elaboração do autor.

Figura C.8 – Disposição da planilha após o comando “Texto para colunas” para dados de caracteres alfabéticos.

### Compilação geral das planilhas auxiliares e elaboração da planilha principal

Antes da migração para o arquivo que contempla todas as planilhas auxiliares e a planilha de dados principal recomenda-se que, para cada tipo de dado a ser avaliado (elementos de liga, propriedades, microestruturas, processos e aplicações), seja criada uma planilha de apoio para compilar os dados de seus diferentes filtros, provenientes dos diferentes campos do registro bibliográfico. Como exemplo, a Figura C.9 apresenta os dados dos percentuais de carbono a partir do seu retorno nos filtros e campos de “Novidade”, “Descrição Detalhada” e “Foco Tecnológico”.

	A	B	C	D	E	F	G	H	I
1	Novelty			DD			FT		
2	2009K83173	0,2500	0,5500	2009K83173			2009K83173		
3	2009L00625		0,0500	2009L00625			2009L00625		
4	2009L61683			2009L61683			2009L61683		
5	2009L61684			2009L61684			2009L61684		
6	2009L64117			2009L64117	0,1100	0,1400	2009L64117		
7	2009L64118	0,1000		2009L64118			2009L64118		
8	2009L64119	0,3000	0,4500	2009L64119			2009L64119		
9	2009M15411			2009M15411			2009M15411		
10	2009M15818		0,1200	2009M15818			2009M15818		
11	2009M15819			2009M15819			2009M15819		
12	2009M15821	2,2000	2,8000	2009M15821			2009M15821		
13	2009M15823	0,3500	0,6500	2009M15823			2009M15823		
14	2009M15825		0,0300	2009M15825			2009M15825		
15	2009M15827	0,1100	0,1700	2009M15827			2009M15827		
16	2009M15829	0,1600	0,2500	2009M15829			2009M15829		
17	2009M35243	0,2000	0,5000	2009M35243			2009M35243		
18	2009M35246	0,3000	1,0000	2009M35246			2009M35246		
19	2009M41456	0,3600	0,4400	2009M41456			2009M41456		
20	2009M41457	0,5200	0,7300	2009M41457			2009M41457		
21	2009M56641	0,0400	0,1000	2009M56641			2009M56641		
22	2009M56647	0,0500	0,1600	2009M56647			2009M56647		
23	2009M56648	0,5000	0,5700	2009M56648			2009M56648		
24	2009M56650	3,0000	5,5000	2009M56650			2009M56650		
25	2009M56651	0,2500	0,3500	2009M56651			2009M56651		

Fonte: Elaboração do autor.

Figura C.9 – Planilha de apoio para compilar os dados recuperados de teor de carbono dos três diferentes campos.

Nessa planilha de apoio, devem ser realizadas as seguintes operações:

1 – Substituição dos pontos “.” por vírgulas “,” (separadores decimais nos sistemas estadunidense e brasileiro). Pode ser realizada pela seleção de todas as colunas e, em seguida, acessar o menu superior “página inicial” e o comando “substituir”.

2 – Atribuição do formato “valores” para as colunas que contém os dados numéricos. Essa etapa pode ser realizada pela seleção das colunas (somente com os dados, não as com os códigos GA), clicar com o botão direito, acesso a opção “Formatar Células” e, por fim, a escolha do formato “Número”.

A seguir, todas as colunas podem ser copiadas e coladas na aba correspondente do arquivo “NIT SPD\_Full”, que compila todas as planilhas de apoio e a planilha principal “NIT SPD”. Deve-se atentar que todas as planilhas já

contêm fórmulas, entretanto, de acordo com o volume de registros recuperados em atualizações futuras, estas devem ser adaptadas para o número de linhas correspondentes. As fórmulas estão condicionadas para a análise de “verdadeiro” e “falso” para leitura de colunas vazias, seguidas de fórmulas condicionais “se” para a apresentação dos dados recuperados nos filtros correspondentes, com a ordem de prioridade estabelecida entre eles (da esquerda para a direita) conforme demonstrado na Figura C.10.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	#GA+Carbono	(NOV)		#GA+Carbono	(DD)		#GA+Carbono	FT							
2	2016130785			2016130785	1,7000	2,5000	2016130785			VERDADEIRO	VERDADEIRO	FALSO	FALSO	1,7000	2,5000
3	2016740136			2016740136			2016740136	0,8000	1,6000	VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO	0,8000	1,6000
4	201704657Q		0,0120	201704657Q			201704657Q			VERDADEIRO	FALSO	VERDADEIRO	VERDADEIRO	0,0120	
5	2017307780			2017307780			2017307780	0,6500	0,7800	VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO	0,6500	0,7800
6	201758206N			201758206N			201758206N	0,1100	0,1600	VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO	0,1100	0,1600
7	201774102D			201774102D		0,0700	201774102D			VERDADEIRO	VERDADEIRO	VERDADEIRO	FALSO		0,0700
8	201954013H			201954013H			201954013H	0,0300	0,1000	VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO	0,0300	0,1000
9	2013E99823			2013E99823			2013E99823			VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO		
10	2013M37212			2013M37212			2013M37212			VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO		
11	2013P18855			2013P18855			2013P18855		0,0250	VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO		0,0250
12	2013T72237			2013T72237			2013T72237			VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO		
13	2013X07891			2013X07891			2013X07891			VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO		
14	2015119537			2015119537			2015119537	0,3890	0,4500	FALSO	FALSO	VERDADEIRO	VERDADEIRO		
15	2018260305			2018260305	0,0300	0,1000	2018260305			VERDADEIRO	VERDADEIRO	FALSO	FALSO	0,0300	0,1000
16	201856070H			201856070H			201856070H	0,4500	0,7000	VERDADEIRO	VERDADEIRO	VERDADEIRO	VERDADEIRO	0,4500	0,7000

Fonte: Elaboração do autor.

Figura C.10 –Planilha de apoio para compilar os dados recuperados de teor de carbono de três diferentes campos.

Por fim, deve-se atentar que a planilha principal “NIT SPD” é totalmente composta por fórmulas do tipo “PROCV”, como destacado na Figura C.11. Assim, o primeiro passo de uma nova atualização deve ser colar os códigos GA referentes aos registros importados na primeira coluna e verificar se as fórmulas correspondem aos valores recuperados nas planilhas de apoio.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Código GA	C (%p)		Mn (%p)		Si (%p)		Al (%p)		Cr (%p)		Ti (%p)		Nb (%p)		V (%p)		Processo 1
2		mín	máx	mín	máx	mín	máx	mín	máx	mín	máx	mín	máx	mín	máx	mín	máx	
3	2009K83173	0,2500	0,5500	1,2000	1,6000	0,1500	0,4500			0,2000	0,9000							
4	2009L00625		0,0500	10,0000	14,0000	1,0000				2,8000	3,0000							
5	2009L61683																	
6	2009L61684																	
7	2009L64117	0,1100	0,1400	0,8000	1,0000	0,2000	0,4000		0,0200	9,1000	10,0000			0,0400	0,0800	0,1800	0,2500	
8	2009L64118		0,1000							0,5000				0,1460	0,1500	0,5000		
9	2009L64119	0,3000	0,4500	0,5500	1,0000	0,4000	1,0000			2,7000	3,5000					0,6000	1,2000	
10	2009M15411																	
11	2009M15818		0,1200			0,1500	0,7500			0,3000	1,2500							
12	2009M15819																	
13	2009M15821	2,2000	2,8000	0,2000	0,9000	1,3000				4,0000	5,6000			0,2000	5,0000	5,0000	21,0000	Metalurgia do po

Fonte: Elaboração do autor.

Figura C.11 – Detalhes da organização da planilha principal “NIT SPD” e de suas fórmulas “PROCV”.



## APÊNDICE D INSTRUÇÕES DETALHADAS DE ELABORAÇÃO DE INDICADORES

Este apêndice tem como objetivo apresentar detalhes de operações realizadas no MS Excel (versão Office 365 em outubro de 2020) para elaboração dos indicadores, como descritos na Metodologia e apresentados no item de Resultados e Discussões. Além disso, apresenta recomendações de procedimentos a serem seguidos no tratamento dos dados recuperados, e de buscas dos textos originais dos registros de interesse na *Derwent Innovations Index*. A intenção é de que seu conteúdo se torne um guia para que novos indicadores sejam elaborados e analisados por pesquisadores do NIT/Materiais e demais interessados em elaborar uma planilha semelhante, ou em melhorias dos métodos utilizados neste trabalho.

Recomenda-se que sejam mantidas algumas configurações básicas em toda a operação, como formato das células e suas configurações em “número” e “textos”. Também é recomendado que se mantenha as células vazias sem a apresentação de nenhum valor, o que indica que o dado correspondente não foi recuperado. Caso, ao abrir a planilha, sejam apresentados zeros no lugar das células sem valor, a alteração pode ser realizada a partir do seguinte caminho:

- Passo 1: Acessar o menu “Arquivo” do MSExcel.
- Passo 2: Acessar o campo “Opções”
- Passo 3: Acessar o menu “Avançado”
- Passo 4: Buscar a opção “Mostrar um zero nas células cujo valor é zero” e certificar que esteja desmarcada.
- Passo 5: Clicar em “OK”.

### **Recomendações para filtros em colunas e classificação de dados no MSExcel**

Por padrão, a planilha deve apresentar, em cada coluna da Linha 2, a opção de filtros de textos, cores ou valores conforme demonstrados na Figura D.1. Caso não estejam disponíveis, os filtros podem ser acionados a partir da

seleção de toda a linha 2, e acionamento do botão “filtro” no menu superior “Dados”.

The image shows the Microsoft Excel interface. The 'Dados' ribbon is active, and the 'Filtro' button is highlighted with a red box. Below the ribbon, a data table is displayed with columns for 'Código GA', 'C (%p)', 'Mn (%p)', 'Si (%p)', 'Al (%p)', 'Cr (%p)', 'Ti (%p)', 'Nb (%p)', and 'V (%p)'. A dropdown menu is open for the 'Código GA' column, showing various filtering options such as 'Classificar de A a Z', 'Classificar de Z a A', 'Classificar por Cor', 'Modo de Exibição Planilha', 'Limpar Filtro de "Processo 1"', 'Filtrar por Cor', 'Filtros de Texto', and 'Pesquisar'. The 'Pesquisar' section includes a search bar and a list of checkboxes for filtering criteria, such as '(Selecionar Tudo)', 'Alívio de tensões', 'Atmosfera controlada', 'Carbo nitração', 'Carburização', 'DesCarburização', 'Desforforização e Desulfurização', 'Desoxidación', and 'Endurecimento por precipitação'.

Código GA	C (%p)	Mn (%p)	Si (%p)	Al (%p)	Cr (%p)	Ti (%p)	Nb (%p)	V (%p)
2009K83173	0,2500	0,5500	1,2000	1,6000	0,1500	0,4500	0,2000	0,9000
2009L00625	0,0500	0,0500	10,0000	14,0000	1,0000			
2009L61683								
2009L61684								
2009L64117	0,1100	0,1400	0,8000	1,0000	0,2000	0,4000	0,0200	9,1000
2009L64118	0,1000							10,0000
2009L64119	0,3000	0,4500	0,5500	1,0000	0,4000	1,0000		2,7000
2009M15411								3,5000
2009M15818	0,1200				0,1500	0,7500		
2009M15819							0,3000	1,2500
2009M15821	2,2000	2,8000	0,2000	0,9000	1,3000		4,0000	5,6000
2009M15823	0,3500	0,6500	0,8000	1,0000	0,2000	0,5000	5,0000	7,0000
2009M15825	0,0300	0,1000	1,0000	1,0000	0,1000	1,0000	14,0000	22,0000
2009M15827	0,1100	0,1700	0,2500	0,5500	0,0500	0,0200	0,0800	
2009M15829	0,1600	0,2500	1,2000	1,6000	0,4000	1,0000	1,5000	
2009M35243	0,2000	0,5000	1,8000	3,8000	0,8000	2,8000		0,0400
2009M35246	0,3000	1,0000	13,0000	22,0000	3,0000	7,0000		0,2500
2009M41456	0,3600	0,4400	1,1400	1,5600	0,4500	0,8500	0,0100	0,0300
2009M41457	0,5200	0,7300	0,4000	0,9000	0,8500	0,0500	0,2500	0,0100
2009M55641	0,0400	0,1000	1,5000	2,5000	0,6000	2,0000	0,0400	0,2000

Fonte: Elaboração do autor.

Figura D.1 – Exemplo de apresentação de filtros nas colunas do MS Excel para operação da planilha de dados.

Os filtros permitem que sejam apresentadas somente as linhas, daquela coluna, que correspondem aos textos, cores ou faixa de valores selecionados. A partir dessa seleção, as linhas que não correspondem aos critérios colocados se tornarão ocultas. Logo, para a elaboração de gráficos ou tabelas a partir de dados filtrados, recomenda-se criar uma cópia da planilha principal “NIT\_SPD”, como o exemplo desta tese, através da opção “mover ou copiar” apresentada ao clicar com o botão direito no nome da planilha na parte inferior da janela conforme demonstrado na Figura D.2. O procedimento também pode ser elaborado de outras formas, de acordo com o interesse e da proficiência no software como, por exemplo, copiar apenas uma parte de interesse da planilha.



15	2009M15825		0		1,0000
16	2009M15827	0,1100	0		0,0500
17	2009M15829	0,1600	0		
18	2009M35243	0,2000	0		2,8000
19	2009M35246	0,3000	1		7,0000
20	2009M41456	0,3600	0		0,8500
21	2009M41457	0,5200	0		0,8500
22	2009M56641	0,0400	0		2,0000
23	2009M56647	0,0500	0		0,5000
24	2009M56648	0,5000	0		0,4000
25	2009M56650	3,0000	5		2,0000
26	2009M56651	0,2500	0		2,0000
27	2009M56652	0,2500	0		2,0000
28	2009M56653				
29	2009M56654	0,0500	0		0,0400

				Inserir...	
				Excluir	
				Renomear	
				Mover ou Copiar...	
				Exibir Código	
				Proteger Planilha...	
				Cor da Guia	>
				Ocultar	
				Reexibir...	
				Selecionar Todas as Planilhas	

				Capa	NIT_SPD	
--	--	--	--	------	---------	--

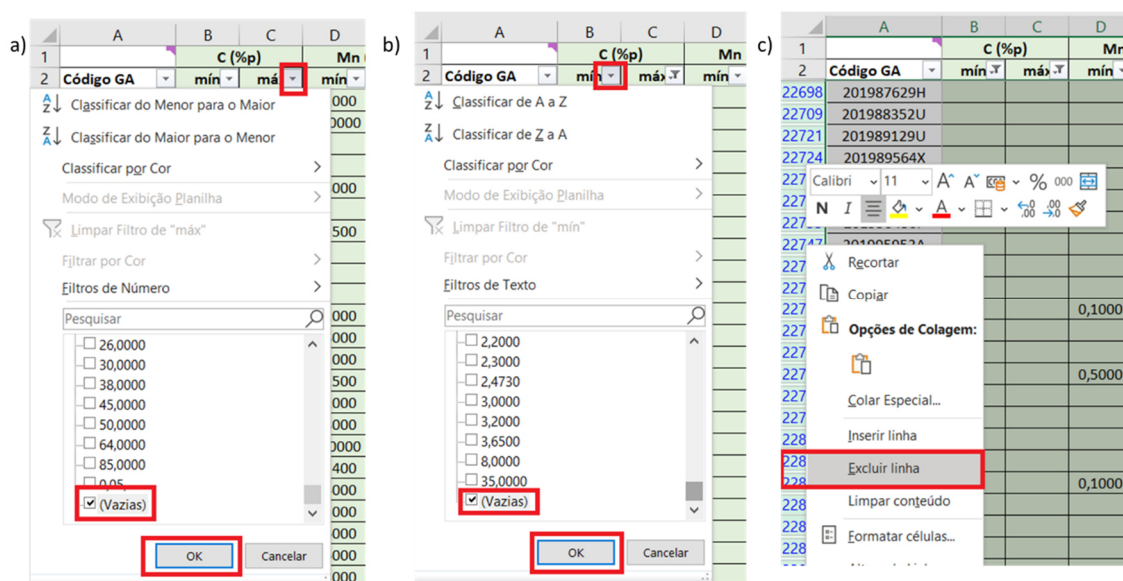
Fonte: Elaboração do autor.

Figura D.2 – Etapas para criar uma cópia da planilha principal para elaboração de gráficos e tabelas.

Os filtros por cores foram disponibilizados na primeira coluna “Código GA”. A atribuição das cores pode ser consultada no item 3.3 da tese. A inclusão ou exclusão dos dados desses registros nos indicadores fica a cargo do interesse das futuras pesquisas.

### Recomendações para elaboração de gráficos

Para a elaboração de gráficos que comparam dados de duas colunas distintas, recomenda-se excluir previamente as linhas com células vazias, ou seja, que não recuperaram dados nessas duas colunas de interesse. Esse procedimento pode ser elaborado a partir de filtros, em ambas as colunas, apenas das células vazias. A seguir, pode-se selecionar todas as linhas filtradas, selecioná-las e acionar a opção “Excluir linhas”. A Figura D.3 apresenta uma sequência de etapas que podem ser seguidas para esse tipo de procedimento, a partir dos dados das colunas que contém os teores de Carbono. A seguir, pode-se retirar os filtros e buscar o gráfico de interesse a partir do menu superior “Inserir” -> “Gráficos”.



Fonte: Elaboração do autor.

Figura D.3 – Etapas para exclusão de linhas que não apresentam valores: a) seleção de células vazias na coluna de máximo de carbono; b) seleção de células vazias na coluna de mínimo de carbono; c) exclusão das linhas vazias.

### Recomendações sobre tabelas dinâmicas

As tabelas dinâmicas podem ser úteis para contagem geral de dados, cálculos de somas e médias, e elaboração de tabelas que relacionam mais de duas colunas da planilha de dados. Sua elaboração pode ser feita a partir do menu “Inserir” -> “Tabela Dinâmica”. A tela a seguir pedirá para selecionar as células de interesse, geralmente as colunas que se deseja relacionar. No caso das composições químicas e propriedades, em que os dados são apresentados em mínimo e máximo, sugere-se renomear as colunas de interesse para melhor identificação. A Figura D.4 apresenta um exemplo de Tabela Dinâmica elaborada a partir de valores de média de quantidade máxima de carbono “Cmax”, processos e aplicações selecionadas.

Rótulos de Linha	Esferas e roletes	Moldes e matrizes	Total Geral
Desfoforização e Desulfurização	0,85	0,63	0,68
Endurecimento por precipitação		0,98	0,98
Endurecimento Superficial		0,19	0,19
Laminação a frio	1,20		1,20
Laminação a quente		1,60	1,60
Laminação Controlada	0,70	0,09	0,29
Martêmpera	0,71	0,46	0,47
Metalurgia do pó	1,10	0,20	0,50
Nitruração		1,52	1,52
Normalização	0,20	0,73	0,55
Pressão reduzida ou vácuo	0,80	0,41	0,46
Refino Térmico ou Verguten	0,25	0,47	0,46
Refusão de eletroescória		0,46	0,46
Têmpera interrompida		0,68	0,68
Têmpera Isotérmica	2,00	0,45	0,45
Trabalho a frio		0,35	0,68
Trabalho a quente		1,00	1,00
<b>Total Geral</b>	<b>0,844727273</b>	<b>0,569673913</b>	<b>0,605437352</b>

Fonte: Elaboração do autor.

Figura D.4 – Exemplo de tabela dinâmica a partir de dados de carbono (média dos valores máximos), processos e aplicações selecionadas.

### Buscas por código GA na *Derwent Innovations Index*

Caso seja necessário os registros na *Derwent Innovations Index* para verificação de mais detalhes além dos dados colocados na planilha, a busca por seus códigos GA pode ser elaborada a partir da busca avançada da DII, como demonstrado na Figura D.5. A busca deve ser realizada a partir da expressão “GA=(‘Código 1’ OR ‘Código 2’ OR ... OR ‘Código N’)”, e seus resultados serão apresentados como os resultados de uma pesquisa desses códigos na plataforma. Para mais detalhes da busca avançada, podem ser consultados os tutoriais e sistemas de ajuda da própria base.

a) Pesquisa Básica Pesquisa de patente citada Pesquisa avançada

Utilize rótulos de campo, operadores booleanos, parênteses e resultados de consultas para criar sua consulta. Os resultados aparecerão na Tabela do histórico de busca, na parte inferior da página. [\(Saiba mais sobre a Pesquisa avançada\)](#)

*Exemplo:* TS=(nanotub\* AND carbon) NOT AU=Smalley RE  
#1 NOT #2 [mais exemplos](#) | [visualizar o tutorial](#)

GA=(2009L64117 OR 2009M15821 OR 2009M15829)

b) **Histórico de pesquisa:**

Resultados	Resultados	
		<input type="button" value="Salvar histórico/Criar alerta"/> <input type="button" value="Abrir histórico salvo"/>
# 1	3	GA=(2009L64117 OR 2009M15821 OR 2009M15829) <i>Índices=CDerwent, EDerwent, MDerwent Tempo estipulado=Todos os anos</i>

Fonte: Elaboração do autor.

Figura D.5 – Exemplo de busca avançada por códigos GA na *Derwent Innovations Index*: a) modelo de expressão de busca; b) resultados apresentados.

## ANEXO A CÓDIGOS DA CLASSIFICAÇÃO INTERNACIONAL DE PATENTES REFERENTES A AÇOS

Quadro A.5 - Códigos da Classificação Internacional de Patentes referentes a ligas ferrosas, em sua versão em português disponibilizada pelo INPI.

<b>Código</b>	<b>Descrição</b>
C22C 38/00	Ligas ferrosas, p. ex. ligas de aço
C22C 38/02	· contendo silício
C22C 38/04	· contendo manganês
C22C 38/06	· contendo alumínio
C22C 38/08	· contendo níquel
C22C 38/10	· contendo cobalto
C22C 38/12	· contendo tungstênio, tântalo, molibdênio, vanádio ou nióbio
C22C 38/14	· contendo titânio ou zircônio
C22C 38/16	· contendo cobre
C22C 38/18	· contendo cromo
C22C 38/20	· · com cobre
C22C 38/22	· · com molibdênio ou tungstênio
C22C 38/24	· · com vanádio
C22C 38/26	· · com nióbio ou tântalo
C22C 38/28	· · com titânio ou zircônio
C22C 38/30	· · com cobalto
C22C 38/32	· · com boro
C22C 38/34	· · com mais de 1,5% em peso de silício
C22C 38/36	· · com mais de 1,7% em peso de carbono
C22C 38/38	· · com mais de 1,5% em peso de manganês
C22C 38/40	· · com níquel
C22C 38/42	· · · com cobre
C22C 38/44	· · · com molibdênio ou tungstênio
C22C 38/46	· · · com vanádio
C22C 38/48	· · · com nióbio ou tântalo
C22C 38/50	· · · com titânio ou zircônio
C22C 38/52	· · · com cobalto
C22C 38/54	· · · com boro
C22C 38/56	· · · com mais de 1,7% em peso de carbono
C22C 38/58	· · · com mais de 1,5% em peso de manganês
C22C 38/60	· contendo chumbo, selênio, telúrio ou antimônio, ou mais de 0,04% em peso de enxofre