ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

TECHNOLOGICAL CONVERGENCE IDENTIFICATION MODEL (TCIM) FOR R&D&I ACTIVITIES -

MODELO DE IDENTIFICAÇÃO DE CONVERGÊNCIA TECNOLÓGICA (MICT) VOLTADA À ATIVIDADES DE PD&I –

BARTHOLOMEO OLIVEIRA BARCELOS
https://orcid.org/0000-0003-1178-1212/ barthobarcelos@gmail.com
Universidade Federal de Santa Catarina: Florianópolis, Santa Catarina, BR

ALEXANDRE LEOPOLDO GONÇALVES
https://orcid.org/0000-0002-6583-2807/ a.l.goncalves@ufsc.br
Universidade Federal de Santa Catarina: Florianópolis, Santa Catarina, BR

LIA CAETANO BASTOS
https://orcid.org/0000-0002-8822-7634/ lia.c.bastos@ufsc.br
Universidade Federal de Santa Catarina: Florianópolis, Santa Catarina, BR

Recebido em: 07/09/2025
Aprovado em: 30/12/2025
Publicado em: 30/12/2025

RESUMO

Os avanços tecnológicos têm acelerado a ascensão de tecnologias emergentes, com essas rápidas mudanças as organizações precisam identificar novas oportunidades de inovação. Nesse contexto, a *Technology Convergence* (TC) emerge como um fator crítico, que integra tecnologias distintas para atender às complexas demandas da sociedade e do mercado competitivo. O desenvolvimento de pesquisas focadas em identificar tecnologias emergentes é vital para responder eficazmente às forças disruptivas e inovar em negócios existentes. Nesse sentido, o objetivo do estudo é propor um modelo voltado à identificação de TC como suporte à tomada de decisão de gestores em atividades de Pesquisa, Desenvolvimento e Inovação (PD&I). O método utilizado foi a implementação do modelo para identificar convergências tecnológicas a partir da análise de patentes, integrando *Knowledge Graphs* (KG), tecnologias semânticas em *Natural Language*

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

TECHNOLOGICAL CONVERGENCE IDENTIFICATION MODEL (TCIM) FOR R&D&I ACTIVITIES -

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Processing (NLP) e Artificial Neural Networks (ANN) baseadas em arquiteturas *Transformers*. Os resultados preliminares indicam que a integração de KGs, NLP e ANNs representa uma possível solução, com viabilidade para identificar padrões de convergência a partir de dados de patentes, auxiliando gestores na tomada de decisões em atividades de PD&I.

Palavras-chave: análise de patente; convergência tecnológica; grafo de conhecimento; processamento de linguagem natural; redes neurais artificiais.

.

ABSTRACT

Technological advancements have accelerated the emergence of new technologies, and with these rapid changes, organizations must identify new innovation opportunities. In this context, Technology Convergence (TC) emerges as a critical factor, integrating distinct technologies to meet the complex demands of society and the competitive market. The development of research focused on identifying emerging technologies is vital to effectively respond to disruptive forces and innovate in existing businesses. To this end, the objective of this study is to propose a model aimed at identifying TC to support managers' decision-making in Research, Development, and Innovation (R&D&I) activities. The method employed was the implementation of the model to identify technological convergences from patent analysis, integrating Knowledge Graphs (KG), semantic technologies in Natural Language Processing (NLP), and Artificial Neural Networks (ANN) based on Transformer architectures. Preliminary results indicate that the integration of KGs, NLP, and ANNs represents a possible solution, demonstrating viability for identifying convergence patterns from patent data and assisting managers in decision-making during R&D&I activities.

Keywords: patent analysis; technology convergence; knowledge graph; natural language processing; artificial neural networks.

1 INTRODUCTION

Technological advancement has been a driving force for global development, creating opportunities and challenges, particularly in the context of innovative technologies that rely on increasing computational power (WIPO, 2023a). In this scenario, Technology Convergence (TC) is characterized by the fusion and overlapping of technologies to meet complex market demands (Kim; Kim, 2012). Innovations resulting from TC enable the creation of hybrid technologies, business diversification, and increased organizational competitiveness (Lee, 2023; Wang; Lee, 2023). To explore this phenomenon, patent databases are widely used, being recognized as fundamental sources of innovation and technological change (Érdi et al., 2013; Afifuddin; Seo, 2024).

This lays the foundation for the field of Patent Analysis (PA). PA plays an essential role in understanding and identifying opportunities related to TC, providing support for strategic decision-making (Afifuddin; Seo, 2024). Despite this, existing research methods in TC primarily focus on

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

qualitative descriptions of the global evolution of technologies and technological fields, lacking quantitative methods that reveal the details of these technological changes (He; Shi; Tan, 2022). Organizations face significant challenges, such as the difficulty in anticipating future convergent technological fields, which impacts the development and maintenance of innovation strategies (Lee; Hong; Kim, 2021). This uncertainty highlights the need to integrate advanced computational resources, such as text mining techniques, Natural Language Processing (NLP), and Artificial Neural Networks (ANN), for a deeper and broader analysis of emerging trends and patterns (Song; Luan; Liang, 2023; Afifuddin; Seo, 2024).

However, it is essential to propose methodologies, models, or frameworks capable of assisting in the identification and anticipation of multi-technological convergences. This involves not only the study of technology pairs in specific fields, but also a more comprehensive analysis of multisectoral interrelationships, offering a more complete, comprehensive, and applied perspective (Kim; Lee, 2017; Lee; Hong; Kim, 2021; Wang; Hsu, 2023).

Considering the challenges identified, this study addresses the following research question: "How can managers be assisted in making decisions to enhance the identification of convergent technological fields in R&D&I activities?" Consequently, the general objective is established as: to propose a model aimed at identifying technological convergence as support for managerial decision-making in R&D&I activities.

The following document presents the constructs that underpin the research proposal, followed by related works and the instantiation of the proposed model, concluding with the study's final considerations.

2 THEORETICAL BACKGROUND

This section presents the main concepts and definitions that underpin the study, addressing topics such as: Technology Convergence (TC), Patent Analysis (PA), Natural Language Processing (NLP) and Large Language Models (LLMs), and Knowledge Representation (KR) and Knowledge Graph (KG).

2.1 Technology Convergence (TC)

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

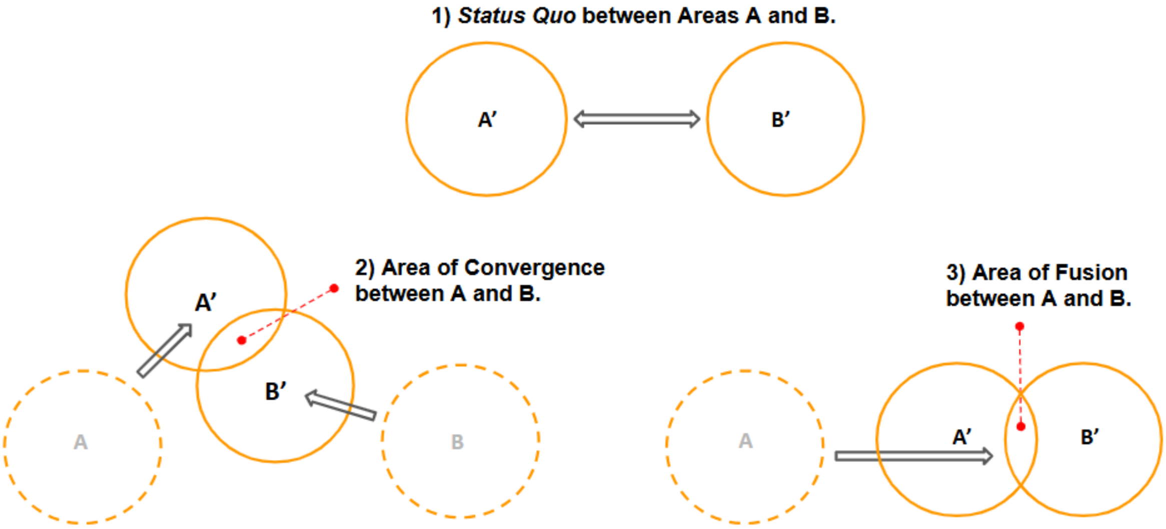
TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

TECHNOLOGICAL CONVERGENCE IDENTIFICATION MODEL (TCIM) FOR R&D&I ACTIVITIES -

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Technology Convergence is a core characteristic of the 21st-century innovation system (Jeong; Kim; Choi, 2015). Considered one of the cornerstones of technological innovation, technology convergence is an emergent phenomenon located at the intersection between previously unrelated technological fields (Wambsganss et al., 2023). It is also understood as the "blurring of boundaries" between sectors or technological fields, resulting in the emergence of new technologies (Nguyen; Moehrle, 2023). For Ardito, Natalicchio, and Petruzzeli (2023), TC represents the phenomenon where a core technology, initially developed in a specific domain, integrates new technologies in distinct application domains. In a convergence scenario, the distance between fields or areas will gradually decrease until a substitutive or additional area of convergence is formed (Figure 1) (Curran; Bröring; Leker, 2010).

Figure 1 - Convergence Process



Source: Curran and Leker (2011)

The Convergence Process designed by Curran and Leker (2011) is corroborated and qualified by Nguyen and Moehrle (2023), who address the process of approximation of two or more previously distant technological areas as bidirectional convergence. It represents the effect of two technologies moving towards each other. Also, they mention unidirectional convergence (also called

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

fusion), which is the effect where only one technology moves towards the other, while the other remains in the same position.

The identification of large-scale convergence can be conducted through three main approaches, according to Curran, Bröring, and Leker (2010): monitoring scientific articles to capture scientific and technological convergence trends; patent analysis to map connections between technological areas and industrial sectors; and analysis of supplementary data, such as collaborative projects and market communications, for a comprehensive assessment of industrial, technological, or market convergences.

Thus, TC is one of the main constructs of this study, being understood as one of the driving forces of innovation. It has attracted increasing attention from researchers, universities, and industry, leading to interdisciplinary research and the development of new products and services, resulting from the combination or fusion of different technologies into new technologies or new applications (Kose; Sakata, 2019; Kwon; Sohn, 2024).

2.1 Patent Analysis (PA)

Patent documents represent a large source of technological knowledge, used according to Ernest (2003, p. 233) in "competitor monitoring, technology evaluation, Research and Development (R&D) portfolio management, identification and evaluation of potential sources for external generation of technological knowledge, especially through mergers and acquisitions, and HR management". Érdi et al. (2013) refer to patent data as a fruitful source of information on innovation and technological change.

Thus, PA plays a key role in TC research and innovation opportunity identification, providing qualified support for decision-making in various sectors (Afifuddin; Seo, 2024). It is one of the frequently used approaches for monitoring TC, as patents are objective and mature indicators of technological changes, serving as an analytical resource for studying the convergence between technologies (Kim; Lee, 2017).

For Moehrle et al. (2010), PA aims to ensure high-quality access to patent sources, involving data capture activities and analysis of patent content, as well as relationships between them, including text mining and citation analysis. The field of PA consists of processes for analyzing and

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

extracting knowledge from patent data, capable of serving multiple research objectives. These objectives include trend and competitive analysis, technology forecasting (TF), and technological strategic planning. It leverages a variety of techniques, including knowledge discovery in data (KDD), knowledge discovery in text (KDT), NLP, cluster analysis, network analysis, and graph theory (Charmanas et al., 2023).

2.2 Natural Language Processing (NLP) and Large Language Models (LLMs)

NLP is a subfield of Artificial Intelligence (AI) that deals with the ability of computers to understand and process human language, including speech and text (Al-Khalifa; AlOmar; AlOlyyan, 2024). As main tasks and themes in the area, Sawicki, Ganzha, and Paprzycki (2023) highlight: word embedding using Bidirectional Encoder Representations from Transformers (BERT) and other Transformer-based models, machine translation between English and other languages, sentiment analysis, question-answering models, Named Entity Recognition (NER), and text summarization.

NLP has witnessed remarkable advancements due to the emergence of LLMs, which have evolved from statistical models, through neural networks, and into pre-trained language models (PLMs) (D'Antonoli et al., 2024). The abilities of language models in syntactic and semantic comprehension allow for human-like interactions, representing the state-of-the-art in linguistic automation (Dugar; Asesh, 2023). Some challenges arise from the use of LLMs: the development of effective prompts; the explainability of LLMs to end-users; the requirement for robust computational resources for model training and execution; the observation of ethical and regulatory issues regarding data privacy; the management and control over content generation to avoid hallucination phenomena that occur in LLMs and lead to error and uncertainty (D'Antonoli et al., 2024; Leiser et al., 2024; Zhao et al., 2024).

As alternatives to qualify the results and performance of LLMs in specific contexts, fine-tuning and prompt engineering techniques are used. The former adjusts the internal parameters of the models, through additional training, to improve their performance in specific tasks that require domain knowledge. The latter optimizes interaction with the LLM through the formulation of effective

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

instructions, which are sequences constructed from natural language (Zhao et al., 2024; Xia et al., 2024; Zhao et al., 2024; D'Antonoli et al., 2024).

Another approach is Retrieval-Augmented Generation (RAG), which has the potential to provide information that may not be present in the LLM's training set or when domain information is continuously updated. This allows the model to ground and improve responses with factual data, effectively reducing the occurrence of inaccuracies or hallucinations (Miao et al., 2024; Radeva et al., 2024).

2.3 Knowledge Representation (KR) and Knowledge Graph (KG)

KR is a field of AI, relevant for the construction of artificially intelligent agents or systems (Paulius; Sun, 2019). KR in AI enables machines to store, organize, and utilize information effectively (Lee, 2024).

The representation approach establishes how information is extracted and how knowledge is defined and materialized. The reasoning process synthesizes the actions executed to achieve a desired outcome (Popescu; Dumitrache, 2023). KR in a machine-interpretable format supports reasoning about knowledge sources, shaping intelligent and user-centric information systems. These systems have the potential to dynamically adapt to the diverse and growing needs of users, based on the integration of AI and Information and Communication Technologies (ICTs). This integration lays the foundation for new technologies for interaction, processing, and visualization of information, increasing the customization and efficiency of these systems (Partarakis; Zabulis, 2024).

A practical form of emerging knowledge representation is the use of KGs. The AI community defines KG as a directed, labeled, and multi-relational graph that incorporates some type of semantics (Kejriwal, 2022). For Zhou et al. (2015), KG is a type of semantic network, where nodes represent concepts and edges represent semantic relationships between these concepts.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Research using KGs has been growing when associated with AI and big data resources, Liu et al. (2024) highlight that these have propelled the use of LLMs and KGs in industrial applications, as these technologies improve accuracy, consistency, and reasoning capability in specific domains. The integration of LLMs and KGs results in more robust and interpretable systems, reflecting a continuous movement toward more adaptable and efficient solutions (Huang et al., 2023; g4).

3 RELATED WORKS

A structured search was conducted in scientific databases to understand the models, frameworks, or architectures used for identifying and anticipating technological opportunities in the field of TC in patent analysis. The search considered English-language publications in the following databases: Scopus®, Web of Science®, ACM Digital Library®, Science Direct®, IEEE Xplore®, Google Scholar®, and Springerlink®. The standard string used the following combination: ((TITLE-ABS-KEY ("technology convergence") AND TITLE-ABS-KEY ("forecasting" OR "prediction" OR "foresight") AND TITLE-ABS-KEY ("patent") AND TITLE-ABS-KEY (“model” OR “framework” OR “architecture”)). When necessary, the strings were adapted with synonyms and variations of the search terms, respecting the specificities of each database.

A total of 956 journal and conference publications were retrieved between 2013 and 2024, which underwent three analysis cycles. In the 1st cycle, duplicate articles and non-pertinent documents were excluded, resulting in 704 publications. After the 2nd cycle, which verified full access to the documents, 670 articles remained. In the 3rd cycle, an analysis of titles and abstracts was applied, resulting in 86 eligible articles. An in-depth review was conducted on 21 articles, analyzing datasets, computational methods and techniques, and artifacts produced (models, frameworks, methods, methodologies, and implementations). The evaluated studies presented a series of methods and techniques employed for identifying convergence patterns between technological fields based on patent data.

Link prediction (LP) techniques were widely employed to identify TC patterns in different sectors. Kim and Cho (2022) applied these techniques in the defense sector, using support vector machine (SVM) and random forest (RF) algorithms, combined with social network analysis (SNA) to map technological interactions. Similarly, Choi, Afifuddin, and Seo (2022) investigated potential connections between unrelated technological fields; associations were represented by patent co-classification graphs. Seo and Afifuddin (2024) expanded the use of LP by combining topic modeling to capture the evolutionary nature of technological fields over time, technological topic networks, and supervised machine learning, revealing specific trends in the biomedicine field. In the framework proposed by Xi et al. (2024), the prediction of technological collaboration (TC) within International Patent Classification (IPC) co-occurrence networks is achieved through an approach that integrates LP, a RF model, and similarity indices. Furthermore, to interpret the obtained predictions and identify future technological patterns, the study applied the Clauset–Newman–Moore algorithm, which enabled the clustering of data and identification of technological communities. Wang and Lee (2023) applied a methodology guided by machine learning (ML) techniques integrated with network analysis to identify and analyze TC patterns in smart health-related patent data. In this methodology, network analysis identifies potential connections between technological areas, while co-occurrence analysis detects frequent patent classification patterns, both contributing to anticipating emerging technological innovations. A deep learning (DL) model centered on a transformer architecture adjusted for the LP task in dynamic knowledge networks was implemented by Zhao et al. (2024), demonstrating significant gains in predictive efficiency over traditional models, such as Recurrent Neural Network (RNN) and static ML models.

The use of semantic and dynamic modeling was evidenced in studies focused on anticipating TC. Zhang and Li (2021) applied semantic vectors based on the Bidirectional Encoder Representations from Transformers (BERT) model to identify TC in patent classifications, with results indicating significant improvements in Area Under the Curve (AUC) and Mean Average Precision (MAP) indicators, highlighting the method's effectiveness in predicting convergences based on textual classification. Liu et al. (2024) complement this approach with Graph Neural Networks (GNN) applied to KGs, combining the BERT model with Generative Adversarial Network (GAN) to identify subgraphs representative of emerging technologies in the intelligent manufacturing domain. Furthermore, Kim and Sohn (2020) demonstrated the impact of semantic analysis on telecommunications and automotive patent data, utilizing LP techniques, semantic vectors

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

(Doc2vec), and machine learning models (SVM, Neural Network (NN), Decision Tree (DT), RF) to identify and anticipate technological fusion patterns.

Predictive models for multi-technological convergences stood out as important contributions in different industrial domains. Kim and Lee (2017) introduced a multi-technological perspective for the IT and Biotechnology sectors, using dependency structure matrix (DSM) to model interactions and ANN to predict comprehensive future convergence patterns. Lee, Hong, and Kim (2021) complemented this approach by using a patent field-value matrix, tracking temporal changes in multi-technological convergence networks. In the environmental and chemical fields, Cho, Lee, and Sohn (2021) proposed a predictive structure that integrated association rule (AR) techniques, ML (RF, SVM, DT, and ANN), and LP, producing robust results for a prediction interval of up to four years, demonstrating the methodology's effectiveness in anticipating future trends and supporting strategic decisions in research and development (R&D&I). Similarly, the study by Lee, Han, and Sohn (2015) applied techniques such as AR and LP to identify intersections between technological fields and predict emerging trends, revealing emerging interconnections in medical and genetic technologies.

The application of TC frameworks in specific sectors was also widely explored. In the domain of defense technologies, Kim and Cho (2023) highlighted the importance of IT and aeronautical technologies, using SVM and RF models to predict more centralized interactions in convergence networks. However, in the field of green technology applied to construction, Wang and Hsu (2023) employed computational methods focused on network theory, constructing technological knowledge flow (TKF) networks to identify emerging technologies, and used the association rule mining (ARM) algorithm to discover multi-technology convergence patterns.

The use of ML techniques, DL, and graph-based networks were utilized to model, predict, and interpret TCs. Wang et al. (2024) employed objective indicators and entropy calculations to select high-value patents, implementing the Doc2Vec algorithm to process abstracts and identify textual similarities. TC relationships were modeled with Graph Convolutional Networks (GCN) and LP techniques, complemented by RF to combine technological solutions, aiming to generate strategic recommendations and identify innovation opportunities. Feng et al. (2020) used LP, network analysis, and text mining to discover TC topics in electric vehicles. In the model proposed by Feng et al. (2024), a two-step predictive approach uses Artificial Neural Network (ANN) to predict

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

eTECH
TECNOLOGIAS PARA
COMPETITIVIDADE INDUSTRIAL

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

the probability and technical impact of TC, with network indicators derived from LP feeding the process. The Shapley Additive exPlanations (SHAP) method was applied to make the model interpretable, revealing the precise influence of each factor on green innovation. Lim and Sohn (2024) applied advanced DL techniques, such as Graph Attributed Multiplex Heterogeneous Network Embedding (GATNE) and Dual Hypergraph Convolutional Network (Dual-HGCN), in multiplex networks to map convergences, providing an integrated perspective on promising areas of sustainable technological development. Afifuddin and Seo (2024) proposed an ML-based approach to predict TC, using technological topic networks applying Dynamic Topic Modeling (DTM) and LP measures to map interactions between topics, with multiple supervised learning algorithms. The methodology allows for anticipating technological trends, highlighting relevances in the biomedicine field and providing strategic support for R&D&I planning. A framework that allows for mapping potential technological convergences prospectively and robustly was proposed by Yang et al. (2024); the approach used three main modules: emerging node simulation, graph representation learning with the GraphSAGE algorithm to generate feature vectors, and evaluation criteria adjusted to the analyzed temporal dimension.

In the composition of the analyzed studies, data processing, and analysis, it was identified that traditional ML techniques integrated with LP techniques are still regularly implemented as effective means to identify and anticipate TC. In the view of Krestel et al. (2021), patents consist of text data and metadata, such as citation information and, sometimes, image data, corroborating the inferences that semantic relationships in the context of TC identification and anticipation are still underexplored given the possibilities of available digital processing and technologies. Thus, opportunities are identified for the advancement of new studies in the NLP area in the TC scenario, aligning also with the considerations of Zhang and Li (2021). The authors emphasize that modern semantic representation methods are still underutilized in studies seeking to estimate or measure the convergence between technologies, resources that could generate more precise results for studies in the area. Aligned with this perspective, Gao and Jiang (2024) are categorical in emphasizing that recent research lacks implementations that make use of semantic analysis for understanding technological convergence or fusion between technological domains.

Therefore, it was observed in the analysis of the studies that semantic relationships possess significant potential that has not yet been sufficiently explored, revealing a promising opportunity for the advancement of research in the field of NLP within the TC scenario (Lee; Han; Sohn, 2015; Kim; Lee, 2017; Kim; Sohn, 2020; Feng et al., 2020; Zhang; Li, 2021; Cho; Lee; Sohn, 2021; Lee; Hong, Kim, 2021; Kim; Cho, 2022; Choi; Afifuddin; Seo, 2022; Wang; Lee, 2023; Kim; Cho, 2023; Wang; Hsu, 2023).

Considering the above, we propose the development of a model that integrates computational methods and techniques with Knowledge Engineering, capable of identifying and anticipating convergent technological fields, providing managers with access to a tool that enhances and strengthens decision-making in R&D&I activities.

4 PROPOSED MODEL AND INSTANTIATION

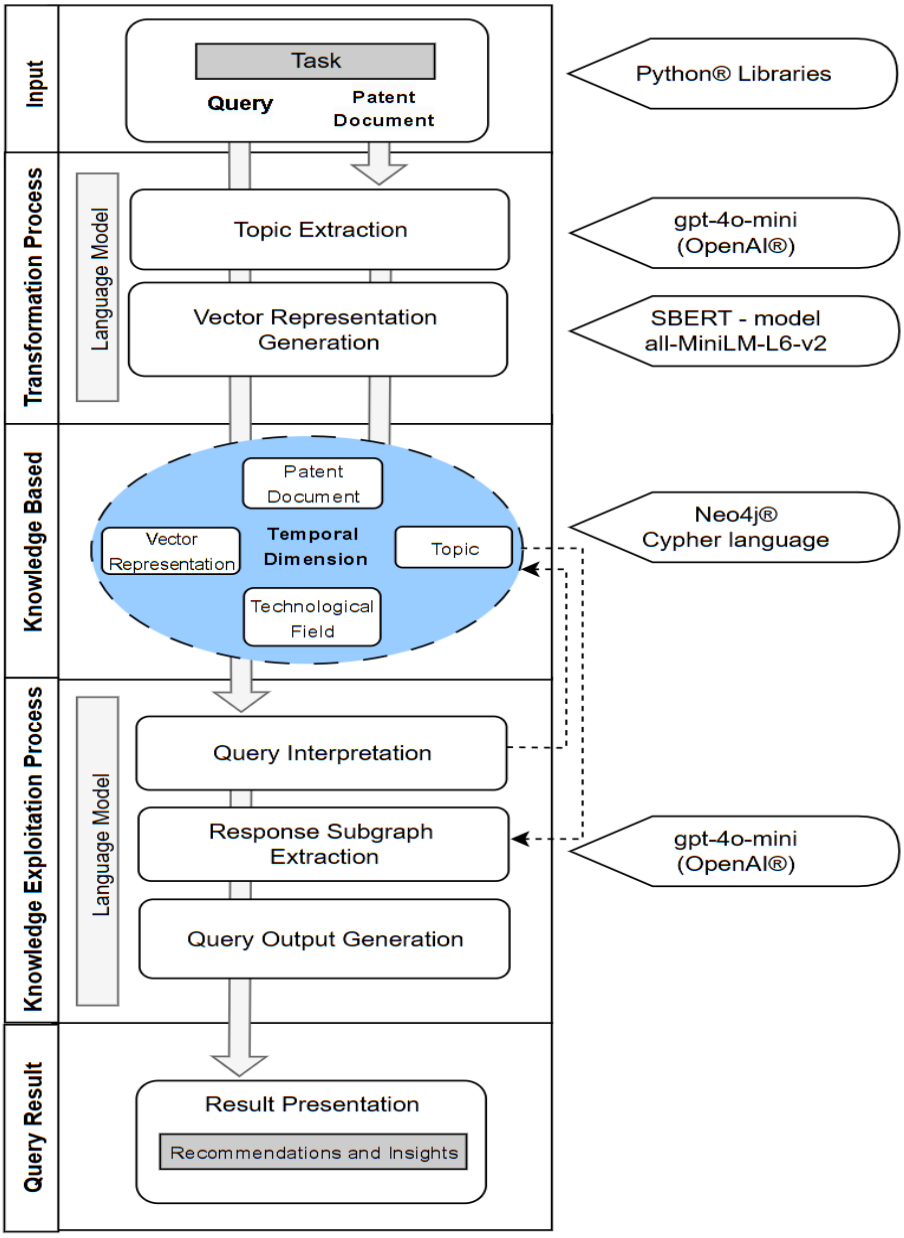
The logical structure of the model is represented in Figure 2, which follows a process-oriented perspective composed of the following components: Input, Transformation Process, Knowledge Base, Knowledge Exploitation Process, and Result Presentation.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Figure 2 - Logical representation of the proposed model



Source: The Authors (2025)

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

4.1 Input and Transformation Process

The database used is USPTO-2M, the patent dataset, which contains over 2 million records that were cleaned and organized in JSON (JavaScript Object Notation) format (USPTO, 2024). The patent documents are written in English; thus, the implemented computational techniques and resources considered the characteristics of the language and the analyzed documents.

The Python® language was used for data preprocessing. The dataset comprises multiple sub-datasets covering the period from 2006 to 2015. Each JSON file corresponds to a specific year, containing information about patents published in that period. A random sample of 1,000 patents was extracted from each dataset, aiming to preserve the statistical representativeness of the original dataset and ensure the reproducibility of the samples in the full dataset, totaling 10,000 patents.

The identification and extraction of the most relevant technical terms from patent titles and abstracts utilized the "gpt-4o-mini" language model from OpenAI®, instructed via prompt to extract the terms and act as a patent analysis expert, with knowledge of IPC (International Patent Classification) and CPC (Cooperative Patent Classification) classifications. An example of the prompt is presented in Figure 3:

Figure 3 - Prompt for term extraction

```
user_prompt = """
# Role and Objective
Act as a patent analysis expert with deep knowledge of the IPC (International Patent Classification)
and CPC (Cooperative Patent Classification) taxonomies. Using the provided text document, extract
the
most relevant technical terms, such as key concepts, techniques, and technologies, that can be
characterized
as entities within the context of patent analysis. Strictly follow these rules:

# Rules
    1)  Only singular forms of terms are included.
    2)  Highly similar terms (e.g., synonyms, plurals, or minor variations) must be aggregated under
        a single representative term.
    3)  Common words from the English language (e.g., general verbs, adjectives, prepositions, and
        non-domain-specific nouns) must be strictly excluded.
    4)  Simple words that do not represent a specific concept, technique, or technology must also be
        excluded, unless they are acronyms or recognized technical terms.
    5)  Prioritize terms that align with IPC and CPC classifications, ensuring that extracted terms
```

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

```
reflect key technological domains relevant to patent analysis.
6) Up to seven terms, mainly terms with at least two words.

# Output
The output must be a Python list containing only the extracted and processed terms without
additional text or formatting.
"""
```

Source: The Authors (2025)

From the files, the term (Term), the year (year), as well as the embeddings generated from the textual content of the patents (Title and Abstract), were incorporated into a dataframe (Figure 4). The vectorial representations were generated using transformer models via the sbert[1] library in Python and the “all-MiniLM-L6-v2” model, an ANN with wide application in generating embeddings from short texts.

Figure 4 - Structure of the dataset after embedding generation

	Subclass_labels	Abstract	Title	No	year	Terms	Embedding
6252	[H05K, E05D, F28F, E05F]	a translating hinge includes a translation ele...	translating hinge	US08270169	2012	["translating hinge", "translation element", "...	[-0.027831489220261574, 0.04053045064210892, -...
4684	[H02H]	a circuit for voltage transient protection is ...	apparatus and method for input voltage transie...	US07660090	2010	["input voltage transient protection", "low vo...	[-0.11418627947568893, 0.06106623262166977, 0....
1731	[C07C]	a process for preparing alkynecarboxylic acids...	process for preparing alkynecarboxylic acids b...	US07173149	2007	["alkynecarboxylic acid", "alkyne alcohol", "o...	[-2.736505302891601e-05, 0.015530739910900593,...
4742	[H02K]	a rotating electric machine which includes a s...	rotating electric machine and manufacturing me...	US07675206	2010	["rotating electric machine", "manufacturing m...	[-0.09851118922233582, 0.10133307427167892, -0...
4521	[H04N]	provided is a method of controlling a digital ...	method of controlling digital photographing ap...	US07760240	2010	["method of controlling", "digital photographi...	[-0.0763683095574379, 0.030493486672639847, -0...
...	...	...	...	...	...	...	...

Source: The Authors (2025)

The following section addresses data persistence, the tasks of storing, organizing, and retrieving the processed information in the proposed model's Knowledge Base.

4.2 Knowledge Base

The knowledge base uses the graph database, Neo4j®, which has the capacity to model and manipulate complex relationships between entities (nodes) and allows for a comprehensive

1SBERT can be found at https://sbert.net/.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

Bartholomeo Oliveira Barcelos
Alexandre Leopoldo Gonçalves
Lia Caetano Bastos

representation to facilitate the KG concept. The native interaction language for Neo4j® is Cypher, which sends instructions to the database to create nodes and relationships in the graph.

Regarding the basic structure of the KG, each node type (Patent, Subclass, Term, Year) represents an important entity in the context of PA and has attributes that store relevant information about that entity. These nodes are the building blocks of the KG, representing the entities that are interconnected through relationships. Each relationship type describes a significant connection between two nodes, such as the patent's relationship with the technological area (Subclass) it belongs to or with the technology topics (Term) it contains. These relationships enable the exploration of connections and patterns among patents.

For data persistence and organization, the model uses commands such as MATCH and MERGE, which allow locating or creating nodes while establishing connections like BELONGS_TO, CONTAINS_TERM, and PUBLISHED_IN. Additionally, co-occurrence connections are created between terms and between subclasses, allowing for the identification of patterns and inter-relationships, strengthening the KG structure and enabling advanced analyses. The terms (Terms), as well as embeddings are utilized via Retrieval-Augmented Generation (RAG) to retrieve relevant patents during searches.

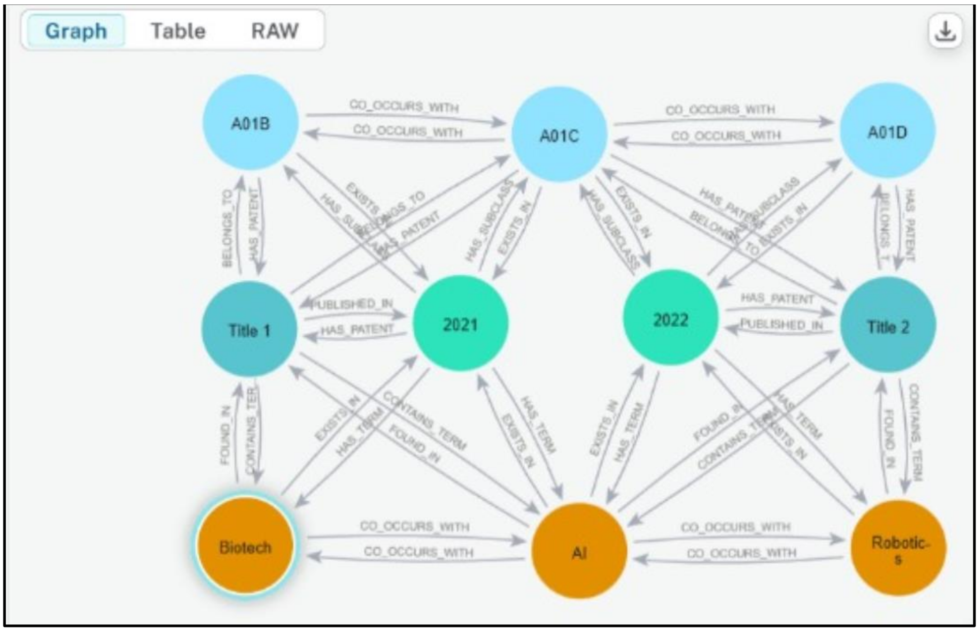
Figure 5 presents the knowledge graph used in the proposed model, which tests and verifies if the nodes and relationships were created correctly, following the structure defined in the dataframe. The dark green nodes represent patents, while the light green nodes represent years. The blue nodes represent subclasses, and the orange nodes represent terms.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Figure 5 - Example of the knowledge graph used in the proposed model



Source: The Authors (2025)

The following section comprises the Knowledge Exploitation Process component, which is activated by an external agent immediately after a natural language query (prompt) is sent to the model.

4.3 Knowledge Exploitation Process

The process is activated by an external agent who performs a natural language query to the LLM. This query is then sent to the Knowledge Base to interpret the demand and extract a response subgraph. Once the subgraph is obtained, it is sent to the language model, in the form of an input prompt, for the generation of the query results performed by the external agent.

Figure 6 presents an example of the search term (sentence) of interest to the external agent, used in the query to locate related patents. To enhance patent searching, based on the provided example, the sentence is enriched with additional contextual information. Instead of directly using the original search term, the system employs the LLM to generate a more comprehensive and informative text, which will serve as the basis for searching similar patents. Thus, a specific interest of an external agent, represented by a search term, for example, "Health and Wellness Technologies", is expanded by a prompt.

Figure 6 - Primary search term and expansion prompt

```
sentence = "Health and Wellness Technologies"
base_prompt = f"""
Given the term(s) {sentence}, generate a text to expand the context of this/these terms in order to improve semantic patent searching.
The response should include the following elements:

a) Related Terms: Identify and list terms that frequently co-occur with the given terms in the patente context.
b) Technical Explanation: Provide a clear and structured definition of each term, considering its usage in patents
and its role in the broader technological domain.
c) Potential Applications: Explain how these terms contribute to innovation, their relevance in different industries,
and potential areas of technological convergence.
d) The answer must be a maximum of 150 words in summary form, i.e. in a single paragraph.
Use a precise, technical, and analytical tone suitable for patent analysis, technology forecasting and technological convergence.
"""
```

Source: The Authors (2025)

As a result, the context of the initial sentence is expanded by the language model, which uses this text converted into an embedding to search for similar patents in the KG, and then displays the information of these patents in an organized and readable format. Figure 7 presents the final result of the generated text, from the sentence "Health and Wellness Technologies", which now includes themes such as: telemedicine, wearables, and AI applied to diagnosis.

Figure 7 - Expansion of the initial query sentence context

```
[ ] generated_text = run_prompt(base_prompt)
    print(break_text_by_words(generated_text, 10))

    Health and Wellness Technologies encompass a wide array of innovations
    including telemedicine, wearable health monitors, personalized health apps, and artificial
    intelligence for health diagnostics. Frequently co-occurring terms such as "biometrics,"
    "remote patient monitoring," "data analytics," and "health informatics" enhance the
    scope of innovation by enabling real-time health management and predictive
    analytics for better patient outcomes. These technologies facilitate a shift
    from reactive to proactive health management, significantly impacting the healthcare
    industry by improving patient engagement and system efficiencies. Additionally, terms
    like "interoperability" and "cloud computing" are critical, as they ensure
    seamless integration and access to health data across platforms, thus
    driving advancements in connectivity and user experience. In industries such
    as fitness, pharmaceuticals, and insurance, the convergence of these technologies
    fosters new business models and enhances service delivery, underscoring their
    relevance in addressing contemporary health challenges.

[ ] #sentence = "wireless and communication"
    #sentence = "paper"
    top_n = 20
    similar_patents = query_similar_patents(generated_text, top_n=top_n)
```

Source: The Authors (2025)

From the retrieved similar patents, the objective is to generate a high-level analysis capable of identifying technological trends and convergences. In Figure 8, the “insight_prompt” is a string containing instructions for the language model to act as an expert in Patent Analysis (PA), trends, and innovation, with expertise also in the fields of forecasting and Technology Convergence (TC).

Figure 8 - Prompt used to generate insights

```
[ ] insight_prompt = """
    You are an experienced patent, trend and innovation analyst used to conducting
    research and drafting reports presenting insights into scenarios involving technology
    forecasting and technological convergence. Based on the following patents, where each
    patent is represented by a title "Title:", an abstract "Abstract:" and a year "Year:"
    and separated by ';', draw up a general analysis of the content producing technological
    insights, as well as establishing possible trends in the main areas of the patents analyzed.
    Ensure that a trend analysis is carried out for each of the main areas identified.
    Finally, generate groupings of highly related technologies and explain in detail the
    connection between them in order to identify possible technological convergences.
    """
```

Source: The Authors (2025)

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

Bartholomeo Oliveira Barcelos
Alexandre Leopoldo Gonçalves
Lia Caetano Bastos

The core of the insight analysis step is executed by the “get_analysis” function, which interacts with the language model to generate patent analysis. This involves sending the prompt and patent information to the LLM, which returns the generated response in natural language (Figure 9).

Figure 9 - Generation of analyses and insights by the language model

```
def get_analysis(prompt, content, model=model_gpt, max_tokens=2000, temperature=0.4):
  completion = client_openai.chat.completions.create(
    model=model,
    max_tokens=max_tokens,
    max_completion_tokens=max_tokens,
    temperature=temperature,
    messages=[
      {"role": "user", "content": f"{prompt} \n Patents: -> {content}"}
    ]
  )
  return completion.choices[0].message.content
```

Source: The Authors (2025)

Upon activating the “get_analysis” function, the analysis prompt and the patent content are combined, allowing the LLM to produce an analysis with information on trends and TC, which will subsequently be displayed to the external agent.

4.4 Result Presentation

The final stage of the proposed model is the “Query Result” process, whose function is to generate recommendations and insights, with information and analyses in natural language. The outputs generated by the “Knowledge Exploitation Process” are actionable information, providing insights for decision-making.

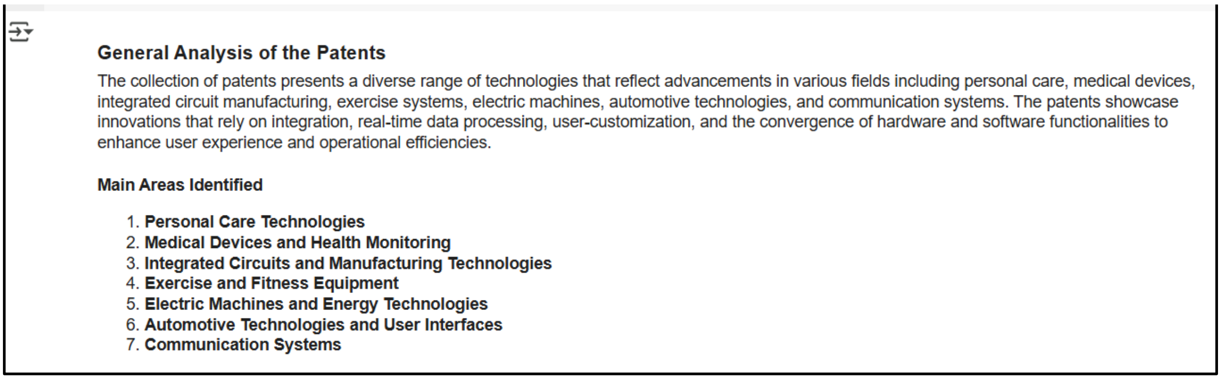
The results produced by the LLM based on the KG are analyzed in three distinct blocks, allowing for a detailed evaluation of its behavior and the quality of the generated insights. Figure 10 demonstrates the results of Block 1, which presents an overview of the analyzed patents, highlighting technological diversity and the presence of fields such as personal care, medical devices, integrated circuit manufacturing, exercise equipment, electrical machines, automotive technologies, and communication systems. It emphasizes integration, real-time data processing, personalization, and

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

hardware-software convergence. In this block, the language model demonstrates the ability to synthesize information from various sources, capturing the essence of the technologies involved.

Figure 10 - Block 1: General analysis and main areas identified



General Analysis of the Patents

The collection of patents presents a diverse range of technologies that reflect advancements in various fields including personal care, medical devices, integrated circuit manufacturing, exercise systems, electric machines, automotive technologies, and communication systems. The patents showcase innovations that rely on integration, real-time data processing, user-customization, and the convergence of hardware and software functionalities to enhance user experience and operational efficiencies.

Main Areas Identified

1. **Personal Care Technologies**
2. **Medical Devices and Health Monitoring**
3. **Integrated Circuits and Manufacturing Technologies**
4. **Exercise and Fitness Equipment**
5. **Electric Machines and Energy Technologies**
6. **Automotive Technologies and User Interfaces**
7. **Communication Systems**

Source: The Authors (2025)

Block 2 (Figure 11) details the trends observed in each of the areas identified in the previous block. For example, in "Personal Care Technologies", the trend is the increasing integration of smart technologies, with personal care devices evolving into intelligent systems capable of adjusting their operation based on user feedback. The result shows that the language model has the ability to go beyond simple topic identification and perform a deeper analysis, identifying the trends that drive innovation in each area.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Figure 11 - Block 2: Analysis of trends in the area

Trend Analysis for Each Area

1. **Personal Care Technologies**:
   - **Trend**: Increasing integration of smart technology into personal care devices.
   - **Analysis**: Devices are evolving from basic mechanical functions to smart systems capable of adjusting their operation based on real-time feedback from user interactions. Devices like the disclosed personal care technologies will likely become more customizable and connected.
2. **Medical Devices and Health Monitoring**:
   - **Trend**: Real-time patient monitoring and automated responses in medical devices.
   - **Analysis**: The shift towards using implantable devices and pattern recognition techniques for immediate diagnosis represents a trend of integrating AI-driven analytics in healthcare. They enable proactive medical interventions and data-driven patient management.
3. **Integrated Circuits and Manufacturing Technologies**:
   - **Trend**: Data-driven manufacturing processes leveraging AI for efficiency.
   - **Analysis**: The innovation in identifying the need for repairs based on historical data of the circuits suggests a movement towards smart manufacturing, where data collection and analytical tools streamline production processes and reduce waste.
4. **Exercise and Fitness Equipment**:
   - **Trend**: Increasing focus on user-centric design in fitness equipment.
   - **Analysis**: The emergence of adjustable fitness systems that can integrate with home environments points to personalized fitness solutions that accommodate varied user preferences and enhance convenience.
5. **Electric Machines and Energy Technologies**:
   - **Trend**: Development of more efficient electric machines with integrated sensors.
   - **Analysis**: The focus on detecting rotational positions and making real-time adaptations reflects ongoing efforts to enhance the performance and energy efficiency of electric machines through smarter design.

Source: The Authors (2025)

Figure 12 presents Block 3, which groups highly related technologies, such as "Medical Devices and Health Monitoring" and "Personal Care Technologies and Exercise Equipment", explaining the connections between them. The generated conclusion summarizes general trends, such as greater personalization, real-time monitoring, and intelligent integration, emphasizing adaptable and user-centric solutions. In this block, the language model demonstrates the ability to identify subtle relationships between different technological fields, revealing possible convergences and innovation opportunities.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

eTECH TECNOLOGIAS PARA COMPETITIVIDADE INDUSTRIAL

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

Figure 12 - Block 3: Technology clusters and conclusion

Groupings of Highly Related Technologies

Group 1: Medical Devices and Health Monitoring

- **Patents**: System for detecting cardiac ischemia; Method for moistening non-biological medical implant material.
- **Connection**: Both patents pertain to advancements in medical technology, specifically geared towards improving patient outcomes through innovations in device functionality and health monitoring.

Group 2: Personal Care Technologies and Exercise Equipment

- **Patents**: Personal care device; Seated row exercise system.
- **Connection**: These patents signify a convergence of health and wellness technologies, focusing on enhancing personal user health through smart and customizable solutions in both personal care and fitness applications.

Group 3: Integrated Circuit Manufacturing and Communication Systems

- **Patents**: Method for using data regarding manufacturing procedures in ICs; Over the air signaling service method and apparatus.
- **Connection**: Both reflect innovations in system efficiency and data communication, showcasing how improvements in manufacturing are directly linked to advanced communication infrastructure and signal processing technologies.

Conclusion

The patents analyzed reveal significant trends toward greater personalization, real-time monitoring, and intelligent integration across various technological domains. The converging technologies demonstrate an increasing emphasis on adaptability and user-centric solutions, indicating a broader movement towards smarter, connected devices across personal care, health monitoring, manufacturing, fitness, automotive applications, and communication systems. As these technologies continue to evolve, the potential for collaboration and convergence between these distinct areas becomes increasingly valuable, paving the way for holistic innovations.

Source: The Authors (2025)

Despite generally satisfactory performance, the model demonstrates some limitations when generating insights for the example in the "Health and Wellness Technologies" field. The inclusion of areas such as "Integrated Circuits and Manufacturing Technologies", "Electric Machines and Energy Technologies", and "Communication Systems" reveals a tendency to identify indirect, superficial, or even less relevant connections to the central theme. This behavior suggests the need to increase the number of analyzed patent samples, improve prompts, and provide more context, in order to direct the model towards areas with more direct and evident applications in Health and Wellness, ensuring the relevance and value of the generated insights.

5 FINAL CONSIDERATIONS

In a scenario of increasing digitalization and big data, driven by innovations such as AI, the Fourth Industrial Revolution (4IR) challenges organizations to explore convergent technologies to maintain competitiveness. This research proposes a model to identify and anticipate technological convergences based on Patent Analysis (PA), integrating Knowledge Graphs (KG), semantic technologies in Natural Language Processing (NLP), and Large Language Models (LLM).

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

The literature review, when analyzing models, frameworks, and architectures for technological convergence forecasting in patents, revealed a gap: the lack of integrated models or frameworks that consider recent techniques and methods, aimed at identifying technologies and technology groups, as well as natural language analysis with the goal of producing insights and recommendations focused on Technology Convergence (TC).

Preliminary results, based on USPTO data, suggest that the proposed model can support the identification of convergence patterns, with embeddings and KGs capturing the semantics and complexity of technological relationships.

Additional improvements, such as expanding the patent dataset sample, refining the prompt, and expanding the context, can enhance the recommendation generation capacity and the model's efficiency, boosting the accuracy of the generated insights and enabling the development of Technology Forecasting (TF).

The preliminary results of the instantiated model indicate the feasibility of identifying convergence patterns from patent data. Although the proposed model still requires refinements, it already outlines strategies for identifying and anticipating convergence between technologies and identifies characteristics for a knowledge representation structure.

REFERENCES

AFIFUDDIN, M.; SEO, W. Predictive modeling for technology convergence: A patent data-driven approach through technology topic networks. **Computers & Industrial Engineering**, v. 188, 2024. DOI: https://doi.org/10.1016/j.cie.2024.109909.

AL-KHALIFA, H.S.; ALOMAR, T.; ALOLYYAN, G. Natural Language Processing Patents Landscape Analysis. **Data**, v. 9, n. 52, 2024. DOI: https://doi.org/10.3390/data9040052.

ARDITO, L; NATALICCHIO, A.; PETRUZZELLI, A. M. Evidence on the Determinants of the Likelihood and Speed of Technological Convergence: A Knowledge Search and Recombination Perspective in Key Enabling Technologies. **IEEE Transactions on Engineering Management**, v. 70, n. 4, Apr. 2023. DOI: https://doi.org/10.1109/TEM.2021.3103878.

CHARMANAS, K. et al. Classifying the Main Technology Clusters and Assignees of Home Automation Networks Using Patent Classifications. **Computers**, n.12, v.10, out. 2023 DOI: https://doi.org/10.3390/computers12100211.

CHO, J.H.; LEE, J.; SOHN, S.Y. Predicting future technological convergence patterns based on machine learning using link prediction. **Scientometrics**, n.126, p.p 5413–5429, 2021. DOI: https://doi.org/10.1007/s11192-021-03999-8.

CHOI, S.; AFIFUDDIN, M.; SEO, W. A Supervised Learning-Based Approach to Anticipating Potential Technology Convergence. **IEEE Access**, v. 10, p. 19284-19300, 2022. DOI: https://doi.org/10.1109/ACCESS.2022.3151870.

CURRAN, C.; BRÖRING, S.; LEKER, J. Anticipating converging industries using publicly available data. **Technological Forecasting and Social Change**, n. 77, v. 3, p. 385-395, 2010. DOI: https://doi.org/10.1016/j.techfore.2009.10.002.

CURRAN, C.; LEKER, J. Patent indicators for monitoring convergence – examples from NFF and ICT. **Technological Forecasting and Social Change**, v. 78, n 2, 2011. DOI: https://doi.org/10.1016/j.techfore.2010.06.021.

DUGAR, M.; ASESH, A. Spatial Interpretation and LLMs. **IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT)**, San Salvador, El Salvador, 2023, p. 1-6, 2023. DOI: https://doi.org/10.1109/ICMLANT59547.2023.10372976.

D'ANTONOLI, T. A. et al. Large language models in radiology: fundamentals, applications, ethical considerations, risks, and future directions. **Diagnostic and Interventional Radiology**, 2024. DOI: https://doi.org/10.4274/dir.2023.232417.

ÉRDI, P. et al. Prediction of emerging technologies based on analysis of the US patent citation network. **Scientometrics**, n. 95, p. 225-242, 2013. DOI: https://doi.org/10.1007/s11192-012-0796-4.

ERNEST, H. Patent information for strategic technology management. **World Patent Information**, v. 25, n. 3, p. 233-242, 2003. DOI: https://doi.org/10.1016/S0172-2190(03)00077-2.

FENG, L. et al. A Review of Technological Forecasting from the Perspective of Complex Systems. **Entropy**, v. 24, n. 6, Jun. 2024. DOI: https://doi.org/10.3390/e24060787.

FENG, S. et al. The technology convergence of electric vehicles: Exploring promising and potential technology convergence relationships and topics. **Journal of Cleaner Production**, v. 260, 2020. DOI: https://doi.org/10.1016/j.jclepro.2020.120992.

GAO, Q; JIANG, M. Exploring technology fusion by combining latent Dirichlet allocation with Doc2vec: a case of digital medicine and machine learning. **Scientometrics**, n. 129, p. 4043–4070, 2024. DOI: https://doi-org.ez46.periodicos.capes.gov.br/10.1007/s11192-024- 05069-1.

HE, C.; SHI, F.; TAN, R. A synthetical analysis method of measuring technology convergence. **Expert Systems with Applications**, v. 209, dec. 2022. DOI:

https://doi.org/10.1016/j.eswa.2022.118262.

HUANG, Q. et al. Let's Chat to Find the APIs: Connecting Human, LLM and Knowledge Graph through AI Chain. **38th IEEE/ACM International Conference on Automated Software Engineering (ASE)**, Luxembourg, Luxembourg, p. 471-483, 2023. DOI: https://doi.org/10.1109/ASE56229.2023.00075.

JEONG, S; KIM, J.C; CHOI, J.Y. Technology convergence: What developmental stage are we in?. **Scientometrics**, n. 104, p. 841–871, 2015. DOI: https://doi.org/10.1007/s11192-015-1606-6.

KEJRIWAL, M. Knowledge Graphs: A Practical Review of the Research Landscape. **Information**, v. 13, n. 161, 2022. DOI: https://doi.org/10.3390/info13040161.

KIM, J.; LEE, S. Forecasting and identifying multi-technology convergence based on patent data: The case of IT and BT industries in 2020. *Scientometrics*, v. 111, p. 47-65, 2017. DOI: https://doi.org/10.1007/s11192-017-2275-4.

KIM, J.; LEE, S. Forecasting and identifying multi-technology convergence based on patent data: The case of IT and BT industries in 2020. *Scientometrics*, v. 111, p. 47-65, 2017. DOI: https://doi.org/10.1007/s11192-017-2275-4.

KIM, K.; CHO, N. A Study on Predictions of Technology Convergence in Defense Technologies. **ICIC Express Letters**, v. 14, n. 6, 2023. DOI: https://doi.org/10.24507/icicelb.14.06.613.

KIM, K. S.; CHO, N. W. Predicting the Patterns of Technology Convergence in Defense Technologies. **IEEE International Conference on Big Data and Smart Computing (BigComp)**, p. 72-75, 2022. DOI: https://doi.org/10.1109/BigComp54360.2022.00024.

KIM, M.; KIM, C. On A Patent Analysis Method for Technological Convergence. **Procedia - Social and Behavioral Sciences**, v. 40, p. 657-663, 2012. DOI: https://doi.org/10.1016/j.sbspro.2012.03.245.

KIM, S. T.; SOHN, S. Y. Machine-learning-based deep semantic analysis approach for forecasting new technology convergence. **Technological Forecasting and Social Change**, v. 157, 2020. DOI: https://doi.org/10.1016/j.techfore.2020.120095.

KOSE, T.; SAKATA, I. Identifying technology convergence in the field of robotics research. **Technological Forecasting and Social Change**, v. 146, p. 751-766, 2019. DOI: https://doi.org/10.1016/j.techfore.2018.09.005.

KRESTEL, R. et al. A survey on deep learning for patent analysis. **World Patent Information**. v. 65, 2021. DOI: https://doi.org/10.1016/j.wpi.2021.102035.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS

KWON, D.; SOHN, S. Y. Convergence technology opportunity discovery for firms based on technology portfolio using the stacked denoising autoencoder (SDAE). **IEEE Transactions on Engineering Management**, v. 71, p. 1804 -1818, 2024. DOI: https://doi.org/10.1109/TEM.2022.3208871.

LEE, C.; HONG, S; KIM, J. Anticipating multi-technology convergence: A machine learning approach using patent information. **Scientometrics**, v. 126, p. 1867-1896, 2021. DOI: https://doi.org/10.1007/s11192-020-03842-6.

LEE, D. A. Knowledge Representation and Reasoning in AI: Analyzing Different Approaches to Knowledge Representation and Reasoning in Artificial Intelligence Systems. **Journal of Artificial Intelligence Research**, Ahmedabad, India, v. 4, n. 1, p. 14–29, 2024. Available at: https://thesciencebrigade.com/JAIR/article/view/97. Accessed on: 31 may. 2024.

LEE, H. Converging technology to improve firm innovation competencies and business performance: Evidence from smart manufacturing technologies. **Technovation**, v. 123, 2023. DOI: https://doi.org/10.1016/j.technovation.2023.102724.

LEE, W. S.; HAN, E. J.; SOHN, S. Y. Predicting the pattern of technology convergence using big-data technology on large-scale triadic patents. **Technological Forecasting and Social Change**, v. 100, p. 317-329, 2015. DOI: https://doi.org/10.1016/j.techfore.2015.07.022.

LEISER, F. et al. Hallucination Identifier for Large Language Models. **Proceedings of the Conference on Human Factors in Computing Systems (CHI '24)**, Honolulu, HI, USA. May, 2024. DOI: . https://doi.org/10.1145/3613904.3642428.

LIM, D. H.; SOHN, S. Y. Identification of Sustainable Convergence Technology for Batteries via Multiplex Link Prediction. **IEEE Transactions on Engineering Management**, v. 71, p. 14365-14374, Aug. 2024. DOI: https://doi.org/10.1109/TEM.2024.3451154.

LIU, P. et al. Joint Knowledge Graph and Large Language Model for Fault Diagnosis and Its Application in Aviation Assembly. **IEEE Transactions on Industrial Informatics**, v. 20, n. 6, p. 8160-8169, jun. 2024. DOI: https://doi.org/10.1109/TII.2024.3366977.

MEYER, L. et al. LLM-assisted Knowledge Graph Engineering: Experiments with ChatGPT. In: **C. Zinke-Wehlmann und J. Friedrich (Hrsg.), First Working Conference on Artificial Intelligence Development for a Resilient and Sustainable Tomorrow, AIDRST 2023**, Informatik aktuell, S. p. 103–115, 2024. DOI: https://doi.org/10.1007/978-3-658-43705-3_8.

MIAO, J. et al. Integrating Retrieval-Augmented Generation with Large Language Models in Nephrology: Advancing Practical Applications. **Medicina**, v. 60, n. 3, 2024. DOI: https://doi.org/10.3390/medicina60030445.

MOEHRLE, M. G. et al. Patinformatics as a business process: A guideline through patent research tasks and tools. **World Patent Information**, v. 32, n. 4, p. 291-299, 2010. DOI: https://doi.org/10.1016/j.wpi.2009.11.003.

NGUYEN, N. U. P.; MOEHRLE, M. G. Combining the Analysis of Vertical and Horizontal Technology Convergence: Insights From the Case of Urban Innovation. **IEEE Transactions on Engineering Management**, v. 70, n. 4, p. 1402-1415, Apr. 2023, DOI: https://doi.org/10.1109/TEM.2021.3086320.

PARTARAKIS, N.; ZABULIS, X. A Review of Immersive Technologies, Knowledge Representation, and AI for Human-Centered Digital Experiences. **Electronics**, v. 13, n. 269, 2024. DOI: https://doi.org/10.3390/electronics13020269.

PAULIUS, D.; SUN, Y. A Survey of Knowledge Representation in Service Robotics. **Robotics and Autonomous Systems**, v. 118, p. 13-30, 2019. DOI: https://doi.org/10.1016/j.robot.2019.03.005.

POPESCU, D. C.; DUMITRACHE, I. Knowledge representation and reasoning using interconnected uncertain rules for describing workflows in complex systems. **Information Fusion**, v. 93, p. 412-428, 2023. DOI: https://doi.org/10.1016/j.inffus.2023.01.007.

RADEVA, I. et al. Web Application for Retrieval-Augmented Generation: Implementation and Testing. **Electronics**, v. 13, n. 7, 2024. DOI: https://doi.org/10.3390/electronics13071361.

SAWICKI, J.; GANZHA, M.; PAPRZYCKI, M. The State of the Art of Natural Language Processing - A Systematic Automated Review of NLP Literature Using NLP Techniques. **Data Intelligence**, v. 5, n. 3, p. 707–749, 2023. DOI: https://doi.org/10.1162/dint_a_00213.

SEO, W.; AFIFUDDIN, M. Developing a supervised learning model for anticipating potential technology convergence between technology topics. **Technological Forecasting and Social Change**, v. 203, 2024. DOI: https://doi.org/10.1016/j.techfore.2024.123352.

SONG, B.; LUAN, C.; LIANG, D. Identification of emerging technology topics (ETTs) using BERT-based model and sematic analysis: a perspective of multiple-field characteristics of patented inventions (MFCOPIs). **Scientometrics**, v. 128, p. 5883–5904, 2023. DOI: https://doi.org/10.1007/s11192-023-04819-x.

UNITED STATE PATENT TRADEMARK OFFICE (USPTO) . **USPTO-2M [database]**. Available at: https://github.com/JasonHoou/USPTO-2M. Accessed: March 18, 2024.

WAMBSGANSS, A. et al. Technology Strategies in Converging Technology Systems: Evidence from Printed Electronics. **Journal of Product Innovation Management**, v. 40, n. 5, p. 705–732, 2023. DOI: https://doi.org/10.1111/jpim.12693.

WANG, J.; HSU, T. Y. Early discovery of emerging multi-technology convergence for analyzing technology opportunities from patent data: the case of smart health. **Scientometrics**, n. 128, p. 4167–4196, 2023. DOI: https://doi.org/10.1007/s11192-023-04760-z.

WANG, J.; LEE, J. Predicting and analyzing technology convergence for exploring technological opportunities in the smart health industry. **Computers & Industrial Engineering**, v. 182, 2023. DOI: https://doi.org/10.1016/j.cie.2023.109352.

WANG, Z. et al*.* From technology opportunities to solutions generation via patent analysis: Application of machine learning-based link prediction. **Advanced Engineering Informatics**, v. 62, 2024. DOI: https://doi.org/10.1016/j.aei.2024.102944.

WORLD INTELLECTUAL PROPERTY ORGANIZATION (WIPO). **Global Innovation Index 2023: Innovation in the face of uncertainty**. Geneva: WIPO, 2023a. DOI: https://doi.org/10.34667/tind.48220.

XI, X. et al. Exploring the potentials of artificial intelligence towards carbon neutrality: Technological convergence forecasting through link prediction and community detection. **Computers & Industrial Engineering**, v. 190, 2024, DOI: https://doi.org/10.1016/j.cie.2024.110015.

XIA, L. et al. Leveraging error-assisted fine-tuning large language models for manufacturing excellence. **Robotics and Computer-Integrated Manufacturing**, v. 88. 2024, DOI: https://doi.org/10.1016/j.rcim.2024.102728.

YANG, G. et al*.* A framework armed with node dynamics for predicting technology convergence. **Journal of Informetrics**, v. 18, n. 4, 2024. DOI: https://doi.org/10.1016/j.joi.2024.101583.

ZHANG, J; LI, Y. Technology convergence prediction based on semantic representation of patent classification's text. **Proceedings of the 4th International Conference on Computer Science and Software Engineering (CSSE '21)**, p. 285-289, 2021. DOI: https://doi.org/10.1145/3494885.3494937.

ZHAO, H. et al. Explainability for Large Language Models: A Survey. **ACM Transactions on Intelligent Systems and Technology**, v. 15, n. 2, p. 1–38, 2024 DOI: https://doi.org/10.1145/3639372.

ZHAO, Y. et al. Improving the Robustness of Large Language Models via Consistency Alignment. In: **LREC-COLING 2024**, Shandong University, Qingdao, China; Baidu Inc., Beijing, China; Leiden University, Leiden, The Netherlands, p. 8931-8941, may-2024.

ZHOU, P. et al. Automatically Constructing of Association Semantic Link Enabled Subject Cognitive Mapping. **International Conference on Identification, Information, and Knowledge in the Internet of Things**, Beijing, China, p. 50-53, 2015. DOI: https://doi.org/10.1109/IIKI.2015.17.

ISSN 1983-1838

(DOI): 10.18624/ etech.v18i1.1444

TECHNOLOGICAL CONVERGENCE IDENTIFICATION MODEL (TCIM) FOR R&D&I ACTIVITIES -

BARTHOLOMEO OLIVEIRA BARCELOS
ALEXANDRE LEOPOLDO GONÇALVES
LIA CAETANO BASTOS