# ESTRUTURA SEMÂNTICA, CONVERGÊNCIA TECNOLÓGICA E DESEMPENHO PATENTÁRIO: EVIDÊNCIAS DE INOVAÇÕES EM EMBALAGENS PARA EXPORTAÇÃO DE FRUTAS TROPICAIS

*SEMANTIC STRUCTURE, TECHNOLOGICAL CONVERGENCE, AND PATENT PERFORMANCE: EVIDENCE FROM PACKAGING INNOVATIONS FOR TROPICAL FRUIT EXPORTS STRUCTURED ABSTRACT*

**EDILANE DO AMARAL HELENO**
UFPB - UNIVERSIDADE FEDERAL DA PARAÍBA

**PRISCILA REZENDE DA COSTA**
ESPM

**Comunicação:**
O XIV SINGEP foi realizado em conjunto com a 14th Conferência Internacional do CIK (CYRUS Institute of Knowledge) e com a SPARK (Society for Promoting Academic Research and Knowledge), em formato híbrido, com sede presencial na UNINOVE - Universidade Nove de Julho, no Brasil.

**Agradecimento à orgão de fomento:**
As autoras agradecem à CAPES pelo apoio financeiro e à Escola Superior de Propaganda e Marketing (ESPM) pelo suporte institucional à realização desta pesquisa.

# ESTRUTURA SEMÂNTICA, CONVERGÊNCIA TECNOLÓGICA E DESEMPENHO PATENTÁRIO: EVIDÊNCIAS DE INOVAÇÕES EM EMBALAGENS PARA EXPORTAÇÃO DE FRUTAS TROPICAIS

**Objetivo do estudo**
Analisar a relação entre estrutura do conhecimento tecnológico e desempenho inovativo em patentes relacionadas às tecnologias de embalagens aplicadas à exportação de frutas tropicais, considerando simultaneamente dimensões semânticas, classificatórias e relacionais, bem como sua associação com impacto e apropriação da inovação

**Relevância/originalidade**
O estudo integra análise semântica, agrupamentos tecnológicos, indicadores de diversidade, convergência e coesão semântica para examinar patentes relacionadas às tecnologias de embalagens aplicadas à exportação de frutas tropicais, abordando de forma integrada dimensões pouco exploradas conjuntamente na literatura.

**Metodologia/abordagem**
A pesquisa combinou análise semântica de documentos patentários, técnicas de agrupamento não supervisionado e modelagem econométrica para caracterizar diferentes configurações do conhecimento tecnológico e examinar sua associação com indicadores de desempenho inovativo.

**Principais resultados**
Os resultados evidenciaram elevada heterogeneidade tecnológica, organizada em múltiplas trajetórias de conhecimento. Diversidade tecnológica e convergência não estiveram automaticamente associadas ao maior impacto tecnológico, enquanto características dos clusters explicaram melhor o impacto das invenções do que sua apropriação.

**Contribuições teóricas/metodológicas**
Os achados indicam que diversidade tecnológica, convergência e coesão semântica representam dimensões distintas da estrutura do conhecimento. Metodologicamente, demonstram a utilidade da integração entre análise semântica, agrupamento não supervisionado e modelagem econométrica para estudos de inovação baseada em patentes.

**Contribuições sociais/para a gestão**
Os resultados contribuem para a compreensão das relações entre estrutura do conhecimento tecnológico, impacto e apropriação da inovação, oferecendo subsídios para decisões relacionadas à gestão da inovação e ao direcionamento de estratégias tecnológicas em diferentes contextos organizacionais

**Palavras-chave:** Análise de patentes, Estrutura do conhecimento tecnológico, Convergência tecnológica, Análise semântica, embalagens de frutas tropicais

# *SEMANTIC STRUCTURE, TECHNOLOGICAL CONVERGENCE, AND PATENT PERFORMANCE: EVIDENCE FROM PACKAGING INNOVATIONS FOR TROPICAL FRUIT EXPORTS STRUCTURED ABSTRACT*

**Study purpose**
To analyze the relationship between the structure of technological knowledge and innovative performance in patents related to packaging technologies applied to tropical fruit exports, considering semantic, classificatory, and relational dimensions, as well as their association with technological impact and innovation appropriation.

**Relevance / originality**
The study integrates semantic analysis, technological clustering, and indicators of diversity, convergence, and semantic cohesion to examine patents related to packaging technologies applied to tropical fruit exports, addressing dimensions that have rarely been investigated jointly in the literature.

**Methodology / approach**
The research combined semantic analysis of patent documents, unsupervised clustering techniques, and econometric modeling to characterize different configurations of technological knowledge and examine their association with indicators of innovative performance.

**Main results**
The results revealed high technological heterogeneity organized into multiple knowledge trajectories. Technological diversity and convergence were not automatically associated with higher technological impact, while cluster characteristics explained the impact of inventions more effectively than their innovation appropriation.

**Theoretical / methodological contributions**
The findings indicate that technological diversity, convergence, and semantic cohesion represent distinct dimensions of knowledge structure. Methodologically, the study demonstrates the usefulness of integrating semantic analysis, unsupervised clustering, and econometric modeling in patent-based innovation research.

**Social / management contributions**
he results contribute to a better understanding of the relationships among technological knowledge structure, technological impact, and innovation appropriation, providing insights to support innovation management decisions and the development of technological strategies across different organizational contexts.

**Keywords:** Patent analysis, Technological knowledge structure;, Technological convergence, Semantic analysis, Tropical fruit packaging

# ESTRUTURA SEMÂNTICA, CONVERGÊNCIA TECNOLÓGICA E DESEMPENHO PATENTÁRIO: EVIDÊNCIAS DE INOVAÇÕES EM EMBALAGENS PARA EXPORTAÇÃO DE FRUTAS TROPICAIS

## 1 Introdução

A crescente inserção das frutas tropicais nos mercados internacionais tem intensificado a demanda por soluções tecnológicas capazes de garantir qualidade, segurança e eficiência logística ao longo de cadeias de suprimento globalmente distribuídas. Nesse contexto, as tecnologias de embalagem assumem papel estratégico, não apenas como mecanismos de contenção, mas como sistemas ativos de preservação, controle ambiental e proteção pós-colheita, fundamentais para reduzir perdas e assegurar a competitividade em mercados externos (Kader, 2005; Yam et al., 2005; Realini & Marcos, 2014; Tilahun et al., 2019).

A evolução desse campo tem sido marcada pela incorporação de funcionalidades avançadas, como embalagens ativas e inteligentes, sistemas de atmosfera modificada e controlada, materiais poliméricos funcionais e dispositivos de monitoramento da qualidade. Essas transformações evidenciam a crescente integração entre diferentes áreas do conhecimento, indicando que as soluções de embalagem devem ser compreendidas como sistemas tecnológicos complexos, nos quais se articulam conhecimentos provenientes da engenharia de materiais, química aplicada, ciência dos alimentos e tecnologias digitais (Brody et al., 2001; Kerry & Butler, 2008; Marsh & Bugusu, 2007; Kumar et al., 2025).

Do ponto de vista da dinâmica inovativa, esse conjunto de tecnologias apresenta características típicas de sistemas baseados em recombinação tecnológica, nos quais novas soluções emergem da combinação incremental de conhecimentos existentes, em vez de rupturas radicais (Fleming, 2001; Nur et al., 2025). Esse processo resulta em elevada heterogeneidade conceitual e técnica, refletida tanto na diversidade de terminologias quanto na distribuição das tecnologias em múltiplas classificações patentárias (Hall et al., 2001; Chonhenchob et al., 2012; Panou et al., 2025). Como consequência, compreender a evolução das tecnologias relacionadas às embalagens aplicadas à exportação de frutas tropicais requer abordagens capazes de captar não apenas a estrutura classificatória das invenções, mas também sua organização conceitual e funcional.

A literatura de inovação baseada em patentes utiliza amplamente indicadores como citações, tamanho da família patentária e cobertura geográfica para mensurar diferentes dimensões do desempenho inovativo (Archibugi & Planta, 1996; Hall et al., 2001). Entretanto, permanece limitada a compreensão sobre como diferentes configurações da estrutura do conhecimento tecnológico — expressas por meio de diversidade, convergência e proximidade conceitual — se associam simultaneamente aos padrões de impacto tecnológico e apropriação observados em documentos patentários.

Nesse contexto, avanços recentes em inteligência artificial e processamento de linguagem natural têm ampliado as possibilidades de análise de patentes ao permitir a identificação de padrões semânticos, agrupamentos tecnológicos e relações latentes entre diferentes domínios do conhecimento (Nimitkeatkai et al., 2022; Lee et al., 2023; Lin & Chou, 2025). Apesar desses avanços, ainda são escassos os estudos que integram análise semântica, indicadores de diversidade e convergência tecnológica e métricas de desempenho patentário em uma perspectiva unificada, particularmente em tecnologias relacionadas às embalagens aplicadas à exportação de frutas tropicais.

Diante dessa lacuna, este estudo busca responder à seguinte questão de pesquisa: como diferentes configurações da estrutura do conhecimento tecnológico presentes em patentes relacionadas às tecnologias de embalagens aplicadas à exportação de frutas tropicais estão associadas aos seus padrões de impacto tecnológico e estratégias de apropriação?

O objetivo do estudo consiste em analisar a relação entre estrutura do conhecimento tecnológico e desempenho inovativo em patentes relacionadas às tecnologias de embalagens aplicadas à exportação de frutas tropicais, considerando simultaneamente dimensões semânticas, classificatórias e relacionais. Especificamente, busca-se: (i) caracterizar a estrutura semântica do corpus analisado; (ii) identificar agrupamentos tecnológicos e diferentes configurações do conhecimento; (iii) examinar indicadores de diversidade tecnológica, convergência e coesão semântica; e (iv) analisar a associação dessas características com indicadores de impacto tecnológico e apropriação da inovação.

Assim, o estudo contribui para a literatura ao integrar análise semântica, classificação tecnológica, agrupamento não supervisionado e modelagem econométrica em uma estrutura analítica voltada à compreensão das relações entre conhecimento tecnológico e desempenho inovativo. Ao explorar simultaneamente dimensões conceituais, classificatórias e relacionais das patentes, a pesquisa amplia o uso de técnicas de inteligência artificial e mineração textual na investigação da dinâmica tecnológica e da inovação.

## 2 Referencial teórico

As tecnologias de embalagem desempenham papel estratégico na cadeia de exportação de frutas tropicais ao contribuírem para a preservação da qualidade, redução de perdas e ampliação da vida útil dos produtos ao longo de cadeias logísticas cada vez mais complexas e internacionalizadas (Kader, 2005; Realini & Marcos, 2014; Tilahun et al., 2019). A evolução desse campo tem sido marcada pela incorporação de funcionalidades associadas às embalagens ativas e inteligentes, aos sistemas de atmosfera modificada e controlada, aos materiais poliméricos funcionais e aos dispositivos de monitoramento da qualidade, evidenciando uma crescente integração entre diferentes áreas do conhecimento (Brody et al., 2001; Yam et al., 2005; Kerry & Butler, 2008; Caleb et al., 2013; Mahajan et al., 2014). Nesse contexto, as soluções tecnológicas deixam de exercer apenas funções de acondicionamento e passam a operar como sistemas tecnológicos complexos que articulam conhecimentos provenientes da engenharia de materiais, química aplicada, ciência dos alimentos, biotecnologia e tecnologias digitais (Marsh & Bugusu, 2007; Kumar et al., 2025).

A literatura sobre inovação tecnológica sugere que a evolução desses sistemas ocorre predominantemente por processos de recombinação de conhecimentos existentes, nos quais novas soluções emergem da integração incremental de componentes tecnológicos previamente desenvolvidos (Fleming, 2001). Em ambientes caracterizados por elevada interdisciplinaridade, esse processo tende a produzir múltiplas trajetórias tecnológicas, marcadas pela coexistência de diferentes configurações de conhecimento e distintos níveis de especialização e convergência. Nesse sentido, a convergência tecnológica pode ser compreendida como o resultado da articulação progressiva entre diferentes domínios científicos e tecnológicos, gerando soluções multifuncionais e ampliando a complexidade dos sistemas de inovação (Choi et al., 2020; Kumar et al., 2025).

A crescente disponibilidade de bases patentárias e o avanço das técnicas de mineração textual têm ampliado as possibilidades de investigação desses processos. Estudos recentes demonstram que abordagens baseadas em inteligência artificial, processamento de linguagem natural e análise semântica permitem identificar estruturas latentes do conhecimento, padrões de proximidade tecnológica e agrupamentos conceituais que não são plenamente capturados pelas classificações tecnológicas tradicionais (Hsu et al., 2020; Nimitkeatkai et al., 2022; Lee et al., 2023; Lin & Chou, 2025). Dessa forma, a análise semântica complementa os sistemas classificatórios ao possibilitar a investigação simultânea da diversidade tecnológica, da convergência do conhecimento e das relações entre diferentes trajetórias de inovação.

As patentes constituem uma das principais fontes de informação para o estudo da dinâmica tecnológica e da atividade inovativa, sendo amplamente utilizadas como indicadores

de produção de conhecimento, mudança tecnológica e desempenho inovativo (Pavitt, 1985; Griliches, 1990; Hall et al., 2001). Além das classificações tecnológicas, os documentos patentários permitem examinar diferentes dimensões associadas ao impacto e à apropriação das inovações. Indicadores como citações recebidas, tamanho da família patentária e abrangência geográfica da proteção são frequentemente empregados para mensurar a relevância tecnológica e os mecanismos de difusão e proteção do conhecimento (Archibugi & Planta, 1996; Hall et al., 2001; Arundel, 2001).

Recentemente, a integração entre análise patentária e inteligência artificial tem contribuído para ampliar a compreensão sobre a organização do conhecimento tecnológico e sua relação com o desempenho inovativo (Mejia & Kajikawa, 2025). Apesar desses avanços, permanecem limitadas as evidências empíricas sobre como diferentes configurações da estrutura do conhecimento tecnológico, expressas por dimensões semânticas, classificatórias e relacionais, estão associadas aos padrões de impacto tecnológico e apropriação observados em contextos caracterizados por elevada convergência tecnológica, como aqueles relacionados às tecnologias de embalagens aplicadas à exportação de frutas tropicais. É nessa lacuna que o presente estudo se insere.

## 3 Metodologia

### 3.1 Construção do corpus e preparação dos dados

Os dados foram coletados na plataforma Lens.org por meio de uma estratégia de busca estruturada para recuperar documentos patentários relacionados a tecnologias de embalagem, conservação, acondicionamento, transporte e monitoramento aplicadas a frutas e produtos tropicais destinados a cadeias logísticas de exportação. A unidade de análise adotada foi a família de patentes, com o objetivo de reduzir duplicidades associadas a múltiplos depósitos da mesma invenção em diferentes jurisdições (Hall et al., 2001; Tong et al., 2026). A busca foi realizada utilizando a seguinte string:

<(packag* OR "packaging system" OR "packaging solution" OR container* OR crate* OR box*) AND (fruit* OR "tropical fruit*" OR "tropical produce") AND (export* OR shipping OR transportation OR logistics OR distribution) AND (technolog* OR innovation OR smart packaging OR active packaging OR intelligent packaging OR "modified atmosphere" OR "controlled atmosphere" OR MAP OR CAP OR "shelf life" OR freshness) AND (preserv* OR protection OR postharvest OR "quality maintenance" OR "damage reduction")>

Após a recuperação, os registros foram submetidos à remoção de duplicidades residuais, verificação de campos essenciais e padronização por família de patentes. Foram extraídos os campos de título e resumo, além de metadados associados aos documentos, incluindo códigos IPC/CPC, citações, tamanho da família patentária e alcance geográfico da proteção. O pré-processamento textual incluiu conversão para letras minúsculas, remoção de caracteres não informativos, eliminação de stopwords, lematização e normalização de termos recorrentes, resultando em um corpus consolidado para análise semântica e tecnológica (Cunha et al., 2025; Lin & Chou, 2025).

### 3.2 Análise semântica e identificação dos clusters tecnológicos

Os documentos foram representados por vetores TF-IDF construídos a partir da concatenação dos campos de título e resumo. Em seguida, foi aplicada a técnica de decomposição em valores singulares (SVD), gerando uma representação semântica latente do corpus em 200 dimensões, utilizada para reduzir a esparsidade textual e preservar padrões conceituais relevantes (Cunha et al., 2025; Hsu et al., 2020).

A estrutura semântica do corpus foi caracterizada por meio da similaridade do cosseno, considerando medidas de proximidade local, global e em relação ao centróide semântico. Essas

métricas foram utilizadas para avaliar a coesão e a heterogeneidade conceitual da base analisada.

A identificação dos agrupamentos tecnológicos foi realizada por meio do algoritmo k-means aplicado à representação semântica obtida após a redução dimensional. A definição do número de clusters considerou critérios de estabilidade, distribuição dos documentos e interpretabilidade dos resultados, resultando na identificação de dez clusters tecnológicos. A caracterização dos agrupamentos foi realizada com base nos termos mais representativos e nos códigos CPC predominantes em cada conjunto (Cunha et al., 2025; Lin & Chou, 2025).

### 3.3 Diversidade tecnológica, convergência e coesão semântica

Os clusters identificados foram analisados por meio de indicadores de diversidade tecnológica, convergência entre domínios do conhecimento e coesão semântica. A diversidade tecnológica foi avaliada a partir do número de classes CPC distintas, da participação relativa da classe dominante e dos índices de Shannon e Simpson. A convergência tecnológica foi examinada por meio da combinação de diferentes classes e seções CPC em uma mesma patente, utilizando indicadores de sobreposição tecnológica, proporção de documentos multisseção, similaridade de Jaccard e densidade das redes de coocorrência entre classes tecnológicas (Fleming, 2001; Choi et al., 2020; Kim et al., 2019).

Complementarmente, a coesão semântica foi mensurada a partir da similaridade cosseno entre documentos pertencentes a cada cluster, considerando a média e o percentil 75 das similaridades internas. Em conjunto, esses indicadores permitiram caracterizar diferentes configurações da estrutura do conhecimento tecnológico presentes no corpus e subsidiaram as análises comparativas realizadas na seção de resultados (Cunha et al., 2025; Lin & Chou, 2025).

## 4 Análise dos resultados

### 4.1 Estrutura semântica do corpus de patentes

A análise semântica dos documentos patentários permitiu caracterizar a organização conceitual do corpus recuperado pela estratégia de busca adotada, evidenciando padrões que não são diretamente observáveis por meio das classificações tecnológicas formais. A partir do processamento textual utilizando representação TF-IDF e redução de dimensionalidade por decomposição em valores singulares (Archibugi & Planta, 1996; Hall et al., 2001; Cunha et al., 2025), observou-se que o corpus analisado apresentou ampla cobertura documental e elevada diversidade lexical.

O vocabulário composto por 72.551 termos distintos indica a presença de múltiplos domínios tecnológicos, característica recorrente em bases patentárias abrangentes, nas quais diferentes aplicações e contextos tecnológicos coexistem em um mesmo espaço de análise (Kim et al., 2019; Lee et al., 2023; Lin & Chou, 2025). Os principais indicadores semânticos e estruturais do corpus são apresentados na Tabela 1.

Tabela 1: Características semântica e estrutural do corpus patentário

| Dimensão analítica | Indicador | Valor | Interpretação |
|---|---|---|---|
| Cobertural documental | Número de documentos (n_docs) | 33.004 | Total de famílias e patentes analisadas |
| Cobertural documental | Vocabulário distinto (vocab_size) | 72.551 | Quantidade de termos únicos identificados |
| Estrutura textual | Densidade TF-IDF (tfidf_density) | 0.00076611607434280 26 | Indica elevada esparsidade textual |
| Estrutura semântica | Dimensões latentes (svd_dims) | 200 | Números de componentes semânticos retidos |
| Estrutura semântica | Variância explicativa pelo SVD | 24,56% | Informações preservada após redução dimensional |

| | | | |
|---|---|---|---|
| | (svd_explanaid_variance_sum | | |
| Similaridade local | Números vizinhos (knn_k) | 20 | Parâmetro utilizado na análise de vizinhança |
| Similaridade local | Similaridade Média k-NN (knn_cosine_mean_avg) | 0.4895165862193725 | Grau médio de proximidade entre documentos semanticamente próximos |
| Similaridade local | Mediana da similaridade aleatória (knn_cosine_mean_p50) | 0.4707521420766597 | Similaridade típica entre vizinhos próximos |
| Similaridade local | Similaridade média aleatoria (pairwise_cosine_mean) | 0.04906688519220837 | Grau médio de proximidade entre documentos do corpus |
| Similaridade global | Mediana da similaridade aleatória (pairwise_cosine_p50) | 0.0189298482122693 | Similaridade típica entre documentos selecionados aleatoriamente |
| Similaridade global | Percentual 95 da similaridade aleatória (pairwise_cosine_p95) | 0.21585144022747924 | Limite superior da relações semânticas observadas no corpus |
| Similaridade global | Similaridade média ao centróide (cos_to_centroid_mean) | 0.2236684862275743 | Proximidade média dos documentos ao núcleo semântico |
| Núcleo conceitual | Percentil 95 da similaridade ao centróide (cos_to_centroid_p95) | 0.38373567470900083 | Proximidade dos documentos mais alinhados ao núcleo conceitual |

Fonte: Dados da pesquisa, 2026.

A matriz termo-documento apresentou elevada esparsidade (TF-IDF = 0,000766), indicando que cada documento mobiliza apenas uma pequena parcela do vocabulário total. Esse padrão é comum em bases patentárias de grande escala e justifica a utilização de técnicas de redução dimensional para identificar relações semânticas latentes (Cunha et al., 2025; Hsu et al., 2020). A representação em 200 dimensões preservou 24,56% da variância do corpus, valor compatível com domínios tecnológicos caracterizados por elevada heterogeneidade e pela coexistência de múltiplas trajetórias de desenvolvimento (Hsu et al., 2020; Jang et al., 2023; Lin & Chou, 2025).

As métricas de similaridade revelaram um padrão de elevada proximidade conceitual em nível local e baixa similaridade em nível global. A similaridade média entre os 20 vizinhos mais próximos (0,4895) sugere a existência de agrupamentos conceituais relativamente coesos, compatíveis com processos de inovação incremental baseados na recombinação de conhecimentos previamente estabelecidos (Fleming, 2001). Em contrapartida, a baixa similaridade média entre documentos selecionados aleatoriamente (0,0491) evidencia a presença de múltiplas trajetórias tecnológicas coexistindo no corpus analisado.

A análise da proximidade ao centróide semântico reforçou essa interpretação. Embora os documentos compartilhem determinados fundamentos técnicos, os valores observados indicam ampla variação em suas aplicações e configurações tecnológicas. Esse comportamento é característico de domínios orientados à aplicação, nos quais conhecimentos comuns são adaptados a diferentes contextos produtivos, logísticos e tecnológicos (Hsu et al., 2020; Song et al., 2023; Tong et al., 2025).

De forma integrada, os resultados indicam que o corpus combina elevada diversidade lexical, forte esparsidade textual e heterogeneidade tecnológica, ao mesmo tempo em que apresenta agrupamentos conceituais relativamente coesos. Esse padrão é consistente com abordagens que associam a dinâmica inovativa à recombinação incremental do conhecimento e à coexistência de múltiplas trajetórias tecnológicas dentro de um mesmo domínio de análise (Fleming, 2001; Bortolini et al., 2018; de Mello Júnior et al., 2025).

### 4.2 Identificação de clusters tecnológicos por aprendizado não supervisionado

A aplicação de técnicas de agrupamento não supervisionado permitiu identificar dez clusters tecnológicos distintos, conforme apresentado na Tabela 2. Esses agrupamentos representam padrões de proximidade semântica observados entre os documentos recuperados pela estratégia de busca adotada, permitindo identificar diferentes configurações tecnológicas presentes no corpus. Ao serem construídos com base na similaridade conceitual entre documentos, os clusters possibilitaram examinar relações tecnológicas que não são integralmente capturadas pelos sistemas classificatórios tradicionais, como IPC e CPC (Cunha et al., 2025; Lin & Chou, 2025).

Tabela 2: Características semântica e estrutural do corpus patentário

| Dimensão analítica | Indicador | Valor | Interpretação |
|---|---|---|---|
| Cobertural documental | Número de documentos (n_docs) | 33004 | Total de famílias e patentes analisadas |
| Cobertural documental | Vocabulário distinto (vocab_size) | 72551 | Quantidade de termos únicos identificados |
| Estrutura textual | Densidade TF-IDF (tfidf_density) | 0.00076611607434280 26 | Indica elevada esparsidade textual |
| Estrutura semântica | Dimensões latentes (svd_dims) | 200 | Números de componentes semânticos retidos |
| Estrutura semântica | Variância explicativa pelo SVD (svd_explanaid_variance_sum | 24,56% | Informações preservada após redução dimensional |
| Similaridade local | Números vizinhos (knn_k) | 20 | Parâmetro utilizado na análise de vizinhança |
| Similaridade local | Similaridade Média k-NN (knn_cosine_mean_avg) | 0.4895165862193725 | Grau médio de proximidade entre documentos semanticamente próximos |
| Similaridade local | Mediana da similaridade aleatória (knn_cosine_mean_p50) | 0.4707521420766597 | Similaridade típica entre vizinhos próximos |
| Similaridade local | Similaridade média aleatoria (pairwise_cosine_mean) | 0.04906688519220837 | Grau médio de proximidade entre documentos do corpus |
| Similaridade global | Mediana da similaridade aleatória (pairwise_cosine_p50) | 0.0189298482122693 | Similaridade típica entre documentos |

| | | | |
|---|---|---|---|
| | | | selecionados aleatoriamente |
| Similaridade global | Percentual 95 da similaridade aleatória (pairwise_cosine_p95) | 0.21585144022747924 | Limite superior da relações semânticas observadas no corpus |
| Similaridade global | Similaridade média ao centróide (cos_to_centroid_mean) | 0.2236684862275743 | Proximidade média dos documentos ao núcleo semântico |
| Núcleo conceitual | Percentil 95 da similaridade ao centróide (cos_to_centroid_p95) | 0.38373567470900083 | Proximidade dos documentos mais alinhados ao núcleo conceitual |

Fonte: Dados da pesquisa, 2026.

Os resultados evidenciaram que os clusters foram caracterizados por diferentes combinações de classes tecnológicas, refletindo múltiplas trajetórias de desenvolvimento presentes no corpus. Os clusters 1, 2 e 3 concentraram maior volume documental e apresentaram maior diversidade classificatória, sugerindo configurações tecnológicas mais amplas e heterogêneas em comparação aos demais agrupamentos.

Em termos tecnológicos, observou-se a predominância de classes associadas à química aplicada, biotecnologia, ciência dos alimentos e materiais funcionais, especialmente A61K, A61P, A23L, A01N e C12N. A recorrência dessas classificações indica a relevância de conhecimentos relacionados à conservação, proteção, monitoramento e manutenção da qualidade de produtos perecíveis, resultado compatível com a literatura sobre sistemas tecnológicos voltados ao acondicionamento e preservação de alimentos (Brody et al., 2001; Kerry & Butler, 2008; Mahajan et al., 2014).

Também foram identificados agrupamentos associados a materiais avançados, sistemas digitais, tecnologias computacionais e melhoramento vegetal, evidenciando a incorporação de conhecimentos provenientes de diferentes áreas tecnológicas. Essa configuração sugere processos de convergência tecnológica nos quais competências oriundas de distintos domínios são articuladas em soluções multifuncionais (Choi et al., 2020; Kim et al., 2019; Kumar et al., 2025).

Os clusters de menor dimensão apresentaram maior concentração em conjuntos específicos de classes tecnológicas, indicando trajetórias mais especializadas. Em contraste, os agrupamentos mais amplos incorporaram maior diversidade de conhecimentos, sugerindo diferentes níveis de recombinação tecnológica e especialização funcional.

De forma integrada, os resultados indicam que o corpus analisado reúne múltiplas trajetórias tecnológicas caracterizadas por distintos níveis de diversidade, especialização e integração do conhecimento. Esse padrão reforça a interpretação de que a inovação ocorre por meio da articulação progressiva de diferentes competências tecnológicas, conforme proposto por abordagens que compreendem a inovação como um processo cumulativo e baseado na recombinação do conhecimento (Fleming, 2001; Hsu et al., 2020).

### 4.3 Diversidade, convergência e impacto dos clusters tecnológicos

Os indicadores de diversidade tecnológica, convergência, estrutura de rede, coesão semântica e impacto permitiram examinar diferentes configurações do conhecimento presentes nos clusters identificados. Esses indicadores oferecem uma visão integrada da forma como os conhecimentos tecnológicos são organizados e articulados dentro do corpus analisado.

Tabela 3. Indicadores de diversidade, convergência e impacto dos clusters tecnológicos

| Indicadores | Dimensão e diversidade tecnológica | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Clus. 1** | **Clus. 2** | **Clus. 3** | **Clus. 4** | **Clus. 5** | **Clus. 6** | **Clus. 7** | **Clus. 8** | **Clus. 9** | **Clus. 10** |
| (n_docs) Número de patentes pertencentes ao cluster | 3649 | 1910 | 1542 | 1510 | 1499 | 1272 | 1155 | 1046 | 1038 | 1020 |
| (cpc4_unique) Quantidade de códigos CPC4 distintos presentes no cluster | 256 | 256 | 149 | 125 | 131 | 152 | 67 | 222 | 130 | 59 |
| (cpc4_top_share) Participação da classe CPC dominante no Cluster | 0,093 | 0,064 | 0,132 | 0,180 | 0,162 | 0,171 | 0,263 | 0,129 | 0,173 | 0,293 |
| (div_shannon) Índice de Shannon | 4,550 | 4,558 | 3,686 | 3,541 | 3,348 | 3,709 | 2,646 | 3,905 | 3,473 | 2,516 |
| (div_shannon_norm) Versão normalizada do índice de shannon(0-1) | 0,820 | 0,822 | 0,737 | 0,733 | 0,687 | 0,738 | 0,629 | 0,723 | 0,713 | 0,617 |
| (div_simpson) Índice de Simpson | 0,978 | 0,981 | 0,949 | 0,937 | 0,932 | 0,943 | 0,862 | 0,953 | 0,930 | 0,851 |
| **Indicadores** | **Indicadores de convergência tecnológica** | | | | | | | | | |
| | **Clus. 1** | **Clus. 2** | **Clus. 3** | **Clus. 4** | **Clus. 5** | **Clus. 6** | **Clus. 7** | **Clus. 8** | **Clus. 9** | **Clus. 10** |
| (conv_jaccard_medio) Grau média de sobreposição entre classes CPC presentes nas patentes do cluster | 0,065 | 0,030 | 0,121 | 0,167 | 0,166 | 0,127 | 0,353 | 0,109 | 0,137 | 0,329 |
| (conv_densidade_rede_cpc4) Densidade das conexões entre classes CPC na rede tecnológica do cluster | 0,087 | 0,075 | 0,115 | 0,135 | 0,107 | 0,097 | 0,179 | 0,071 | 0,106 | 0,180 |
| (conv_mix_secao_share_docs) Proporção de patentes que combinam mais de uma seção tecnologica CPC | 0,620 | 0,608 | 0,657 | 0,540 | 0,718 | 0,414 | 0,572 | 0,565 | 0,480 | 0,508 |
| (conv_secao_media_por_doc) Número médio de seções tecnológicas preentes em cada patente | 1,961 | 1,918 | 1,894 | 1,746 | 1,949 | 1,538 | 1,764 | 1,770 | 1,638 | 1,672 |
| **Indicadores** | **Estrutura de rede tecnológica** | | | | | | | | | |
| | **Clus. 1** | **Clus. 2** | **Clus. 3** | **Clus. 4** | **Clus. 5** | **Clus. 6** | **Clus. 7** | **Clus. 8** | **Clus. 9** | **Clus. 10** |
| (rede_n_nos_cpc4) Número de classes CPC4 distintas representadas como nós na rede tecnológica | 256 | 256 | 149 | 125 | 131 | 152 | 67 | 222 | 130 | 59 |
| (rede_n_arestas_coc) Números de conexões (coocorrências) entre | 2830 | 2443 | 1266 | 1049 | 910 | 1114 | 395 | 1741 | 891 | 308 |

| classes CPC4 na rede tecnológica | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Indicadores** | **Coesão semântica e impacto tecnológico** | | | | | | | | | |
| | **Clus. 1** | **Clus. 2** | **Clus. 3** | **Clus. 4** | **Clus. 5** | **Clus. 6** | **Clus. 7** | **Clus. 8** | **Clus. 9** | **Clus. 10** |
| (idade_media_anos) Idade média das patentes pertencentes ao cluster | 7,35 | 7,80 | 8,31 | 7,33 | 6,73 | 7,25 | 6,43 | 6,47 | 5,75 | 7,62 |
| (cohesao_media_cos) Similaridade semântica média entre documentos do mesmo cluster | 0,16 | 0,36 | 0,44 | 0,49 | 0,34 | 0,51 | 0,47 | 0,60 | 0,50 | 0,40 |
| (cohesao_p75_cos) Percentil 75 da similaridade semântica interna do cluster | 0,20 | 0,42 | 0,51 | 0,56 | 0,40 | 0,59 | 0,54 | 0,70 | 0,58 | 0,48 |
| (citacoes_forward_proxy_score) Indicador padronizado de impacto tecnológico baseado em citações recebidas | -1,15 | -0,25 | 0,32 | -0,15 | -0,97 | -0,12 | -0,75 | -0,34 | -1,10 | -0,23 |
| (citacoes_forward_proxy_0_100) Indicador de impacto tecnológico em escala de 0 a 100 | 0,00 | 32,12 | 52,44 | 35,88 | 6,17 | 36,77 | 14,18 | 28,88 | 1,82 | 32,82 |

Fonte: Dados da pesquisa, 2026.

A distribuição dos documentos revelou diferenças expressivas entre os agrupamentos tecnológicos. Os clusters 1 e 2 concentraram maior volume documental e apresentaram os maiores níveis de diversidade classificatória, refletidos pela quantidade de códigos CPC4 distintos e pelos elevados valores dos índices de Shannon e Simpson. Em contraste, os clusters 7 e 10 apresentaram estruturas mais especializadas, caracterizadas por menor variedade tecnológica e maior concentração em conjuntos específicos de classes tecnológicas. Esse padrão é compatível com abordagens que associam a recombinação do conhecimento à geração de diferentes trajetórias tecnológicas (Fleming, 2001; Cunha et al., 2025; Song et al., 2023).

As métricas de convergência tecnológica revelaram diferenças importantes na forma como os conhecimentos foram combinados nos distintos agrupamentos. Os clusters 7 e 10 apresentaram maiores níveis de sobreposição tecnológica e densidade de conexões entre classes CPC, indicando combinações mais recorrentes entre domínios específicos do conhecimento. Por outro lado, os clusters mais diversificados exibiram redes mais amplas, porém relativamente menos densas, sugerindo processos de recombinação tecnológica distribuídos por um espectro mais amplo de conhecimentos (Hall et al., 2001; Kim et al., 2019; Lin & Chou, 2025).

A análise da combinação entre seções tecnológicas também evidenciou diferenças relevantes. O cluster 5 apresentou a maior proporção de documentos que combinam múltiplas seções CPC, indicando níveis mais elevados de interdisciplinaridade tecnológica, enquanto o cluster 6 apresentou uma configuração mais concentrada. Esses resultados sugerem que os agrupamentos tecnológicos diferem não apenas em termos de diversidade, mas também na intensidade com que articulam conhecimentos provenientes de distintos domínios tecnológicos (Brody et al., 2001; Marsh & Bugusu, 2007; Bortolini et al., 2018).

Em relação à coesão semântica, observou-se variação significativa entre os clusters. Os agrupamentos 8, 6 e 9 apresentaram os maiores níveis de similaridade interna, indicando

estruturas conceituais mais homogêneas. Em contraste, o cluster 1 apresentou menor coesão semântica, resultado compatível com sua elevada diversidade tecnológica e abrangência temática. Esse comportamento reforça a coexistência de trajetórias tecnológicas com diferentes graus de especialização e integração conceitual (Hsu et al., 2020; Jang et al., 2023; Tong et al., 2025).

Os indicadores de impacto tecnológico evidenciaram que a diversidade e a convergência não estiveram automaticamente associadas aos maiores níveis de desempenho inovativo. Embora o cluster 3 tenha apresentado os maiores valores de impacto tecnológico, medidos por citações forward, outros agrupamentos com elevada diversidade ou convergência exibiram desempenhos mais modestos. Esse resultado sugere que o impacto das invenções depende de configurações específicas do conhecimento tecnológico, e não apenas da amplitude ou intensidade das combinações realizadas (Hall et al., 2001; Tong et al., 2026).

De forma integrada, os resultados indicam que os clusters tecnológicos diferem significativamente quanto à diversidade, convergência, coesão semântica e impacto, evidenciando a existência de múltiplas configurações da estrutura do conhecimento tecnológico. Esses achados reforçam a natureza multidimensional da dinâmica inovativa e fornecem evidências de que diferentes trajetórias tecnológicas podem gerar resultados distintos em termos de desempenho inovativo.

### 4.4 Modelos de associação entre estrutura tecnológica e desempenho inovativo

A análise econométrica buscou avaliar em que medida características relacionadas à diversidade tecnológica, convergência, estrutura de rede e coesão semântica contribuíram para explicar diferenças observadas nos indicadores de desempenho das patentes. Para isso, foram estimados modelos OLS, Ridge e Lasso, permitindo comparar a capacidade explicativa dessas variáveis em relação ao impacto tecnológico e aos indicadores de apropriação da inovação.

Tabela 4. Desempenho dos modelos na explicação dos indicadores de impacto e apropriação tecnológica

| Dimensão analisada | Indicador | Número de variáveis do modelo | OLS ($R^2$ em amostra) | OLS ($R^2$ validação cruzada) | Ridge (α) | Ridge ($R^2$ Validação cruzada) | Lasso (α) | Lasso ($R^2$ Validação cruzada) |
|---|---|---|---|---|---|---|---|---|
| **Impacto tecnológico** | Citações forward (0-100) | 40 | 0,894 | 0,498 | 4,924 | 0,596 | 0,527 | 0,586 |
| **Apropriação tecnológico** | Tamanho da família de patentes (média de jurisdições) | 40 | 0,398 | -0,029 | 0,001 | -0,029 | 0,000 | -0,029 |
| **Apropriação tecnológica** | Alcance geográfico da proteção (jurisdições únicas) | 40 | 0,307 | -0,399 | 1,701 | -0,211 | 0,123 | -0,239 |

Fonte: Dados da pesquisa, 2026.

Os resultados evidenciam diferenças relevantes entre os indicadores analisados. Para o impacto tecnológico, representado pelas citações forward, os modelos apresentaram capacidade explicativa moderada, especialmente após a aplicação das técnicas de regularização. Embora o modelo OLS tenha alcançado elevado ajuste em amostra ($R^2 = 0,894$), a redução observada na validação cruzada ($R^2 = 0,498$) sugere sensibilidade à complexidade do conjunto de variáveis utilizado. Os modelos Ridge ($R^2 = 0,596$) e Lasso ($R^2 = 0,586$) apresentaram melhor

desempenho preditivo, indicando que parte das relações observadas entre as características tecnológicas dos clusters e o impacto das invenções permaneceu estável fora da amostra (Hall et al., 2001; Hsu et al., 2020; Tong et al., 2025).

Esses resultados sugerem que diferentes configurações da estrutura do conhecimento tecnológico estão associadas ao impacto das invenções, ainda que essa relação não seja explicada por uma única dimensão isolada. Em conjunto, os achados reforçam a interpretação apresentada na seção anterior de que diversidade, convergência e coesão semântica interagem de forma complexa na geração de desempenho inovativo.

Em contraste, os resultados associados aos indicadores de apropriação tecnológica apresentaram baixo poder explicativo. Tanto para o tamanho da família patentária quanto para o alcance geográfico da proteção, os modelos exibiram desempenho limitado em validação cruzada, inclusive com valores negativos de $R^2$. Esses resultados indicam que as características tecnológicas dos clusters, embora relevantes para compreender o impacto das invenções, não explicam adequadamente as estratégias de proteção e difusão internacional observadas no corpus.

Esse padrão sugere que decisões relacionadas à extensão territorial da proteção patentária dependem de fatores que transcendem a estrutura tecnológica das invenções, incluindo aspectos organizacionais, econômicos e estratégicos associados ao processo de apropriação do conhecimento (Archibugi & Planta, 1996; Hall et al., 2001). Dessa forma, os resultados apontam para uma distinção entre os mecanismos associados à geração de impacto tecnológico e aqueles relacionados à proteção e difusão das inovações.

De forma integrada, a análise econométrica indica que a estrutura tecnológica apresenta maior capacidade de explicar diferenças associadas ao impacto das invenções do que aos indicadores de apropriação tecnológica. Esse resultado reforça a ideia de que impacto tecnológico e estratégias de apropriação representam dimensões distintas do desempenho inovativo, evidenciando a natureza multidimensional da dinâmica tecnológica observada no corpus analisado (Fleming, 2001; Song et al., 2023; Lin & Chou, 2025).

## 5. Discussão

A análise integrada dos resultados permite examinar como diferentes configurações da estrutura do conhecimento tecnológico se manifestam em bases patentárias caracterizadas por elevada diversidade temática e múltiplas trajetórias tecnológicas. As evidências obtidas a partir das dimensões semântica, classificatória e econométrica demonstraram que os documentos recuperados pela estratégia de busca se organizam em diferentes configurações tecnológicas, caracterizadas por distintos níveis de diversidade, convergência, coesão semântica e desempenho patentário.

Os resultados da análise semântica revelaram a coexistência de elevada proximidade conceitual em nível local e significativa heterogeneidade em nível global. Esse padrão é consistente com a literatura sobre recombinação tecnológica, segundo a qual a inovação ocorre por meio da reorganização progressiva de conhecimentos existentes, e não pela substituição completa de trajetórias tecnológicas anteriores (Fleming, 2001). A combinação entre elevada coesão local e baixa similaridade global indica a presença de agrupamentos conceituais relativamente especializados inseridos em um conjunto mais amplo e heterogêneo de documentos, reforçando a interpretação de que diferentes trajetórias tecnológicas compartilham determinados fundamentos conceituais, mas se distinguem pelas formas de aplicação e combinação do conhecimento (Fleming, 2001; Hsu et al., 2020; Lin & Chou, 2025).

A identificação dos clusters tecnológicos evidenciou diferentes formas de articulação entre conhecimentos provenientes da química aplicada, biotecnologia, materiais funcionais, tecnologias digitais e sistemas agrícolas. Os agrupamentos identificados revelaram padrões diferenciados de especialização e integração tecnológica, indicando que os conhecimentos

mobilizados no corpus não se distribuem de forma homogênea, mas se organizam em trajetórias com diferentes graus de articulação entre domínios tecnológicos. Essa configuração é compatível com abordagens que compreendem a convergência tecnológica como resultado da combinação progressiva de competências oriundas de diferentes áreas científicas e tecnológicas (Choi et al., 2020; Kim et al., 2019; Kumar et al., 2025).

Entretanto, os achados demonstraram que diversidade tecnológica e convergência tecnológica representam dimensões relacionadas, mas não equivalentes. Os clusters mais diversos não foram necessariamente aqueles que apresentaram maior densidade relacional ou maior sobreposição entre classes tecnológicas, assim como agrupamentos mais especializados não exibiram, obrigatoriamente, menor capacidade de articulação entre conhecimentos distintos. Os achados indicam que diversidade e convergência representam dimensões analíticas distintas da estrutura tecnológica e, portanto, fornecem informações complementares sobre a organização do conhecimento.

A análise dos indicadores de impacto tecnológico reforçou essa interpretação. Os resultados mostraram que níveis mais elevados de diversidade tecnológica ou convergência classificatória não estiveram automaticamente associados aos maiores níveis de impacto medidos por citações forward. Alguns clusters relativamente especializados apresentaram desempenho comparável ou superior ao de agrupamentos mais amplos. Esse comportamento reforça a ideia de que o impacto tecnológico decorre menos da amplitude das combinações realizadas e mais da capacidade dessas combinações gerarem soluções tecnológicas reconhecidas e reutilizadas por desenvolvimentos posteriores, aspecto compatível com a literatura que discute a natureza multifatorial do impacto tecnológico e a importância da qualidade das combinações tecnológicas na geração de valor inovativo (Hall et al., 2001; Song et al., 2023).

Os resultados econométricos forneceram evidências adicionais ao indicar que as características tecnológicas dos clusters apresentaram capacidade moderada de explicar diferenças associadas ao impacto das invenções, mas baixo poder explicativo para os indicadores de apropriação tecnológica. Esse padrão sugere que impacto tecnológico e estratégias de proteção representam dimensões distintas do desempenho inovativo, resultado compatível com a literatura que diferencia circulação do conhecimento, reconhecimento tecnológico e apropriação econômica da inovação (Archibugi & Planta, 1996; Hall et al., 2001).

Nesse sentido, os resultados indicam que a estrutura tecnológica parece estar mais associada à geração e circulação do conhecimento do que às decisões relacionadas à extensão internacional da proteção patentária. Tais decisões tendem a depender de fatores adicionais, como condições institucionais, estratégias organizacionais, custos de patenteamento e características dos mercados de atuação, elementos que não foram capturados pelos indicadores tecnológicos utilizados neste estudo.

As evidências apresentadas permitem destacar três implicações teóricas diretamente associadas aos resultados obtidos. Primeiro, evidenciam a utilidade da integração entre análise semântica e análise classificatória para a compreensão da estrutura do conhecimento tecnológico. Segundo, demonstram que diversidade tecnológica, convergência e coesão semântica representam dimensões complementares, porém distintas, da organização tecnológica. Terceiro, indicam que diferentes dimensões do desempenho inovativo respondem de forma diferenciada às características estruturais do conhecimento, reforçando a natureza multidimensional e não linear dos processos de inovação.

Além disso, os resultados indicam que abordagens baseadas exclusivamente em classificações tecnológicas podem não capturar integralmente as relações conceituais presentes nos documentos patentários. A incorporação de técnicas de análise semântica permitiu identificar padrões de proximidade, especialização e organização do conhecimento que complementam as informações derivadas dos sistemas tradicionais de classificação, oferecendo

alternativas metodológicas para investigar processos de convergência tecnológica, especialização e recombinação do conhecimento em bases patentárias caracterizadas por elevada heterogeneidade tecnológica.

## 6. Conclusão

Este estudo analisou a relação entre estrutura do conhecimento tecnológico e desempenho inovativo em um corpus de documentos patentários recuperados por meio de uma estratégia de busca relacionada a tecnologias de embalagens, conservação, monitoramento e acondicionamento de produtos perecíveis. Para isso, foram integradas técnicas de análise semântica, classificação tecnológica, agrupamento não supervisionado e modelagem econométrica, permitindo examinar diferentes dimensões da organização do conhecimento tecnológico e sua associação com indicadores de desempenho patentário.

Os resultados evidenciaram que o corpus analisado apresenta elevada heterogeneidade tecnológica, organizada em múltiplas trajetórias de conhecimento. A análise semântica revelou a coexistência de agrupamentos conceituais relativamente coesos em um conjunto mais amplo e diversificado de documentos, enquanto a identificação dos clusters tecnológicos evidenciou diferentes formas de articulação entre conhecimentos provenientes da química aplicada, biotecnologia, materiais funcionais, tecnologias digitais e sistemas agrícolas. Em conjunto, os resultados indicaram que diversidade tecnológica, convergência tecnológica e coesão semântica representam dimensões complementares, porém distintas, da estrutura do conhecimento. Observou-se ainda que níveis mais elevados de diversidade ou convergência não estiveram automaticamente associados aos maiores níveis de impacto tecnológico, reforçando a natureza não linear das relações entre estrutura tecnológica e desempenho inovativo.

A análise econométrica mostrou que as características tecnológicas dos clusters apresentaram maior capacidade de explicar diferenças associadas ao impacto tecnológico das invenções, medido por citações forward, do que aos indicadores de apropriação tecnológica. Em contraste, o tamanho da família patentária e o alcance geográfico da proteção apresentaram baixa capacidade de explicação a partir das variáveis tecnológicas analisadas. Esse resultado reforça a distinção entre geração de impacto tecnológico e estratégias de apropriação da inovação, em linha com as discussões propostas por Archibugi e Planta (1996) e Hall et al. (2001).

Os achados reforçam a utilidade da integração entre análise semântica, classificação tecnológica e modelagem econométrica para investigar a organização do conhecimento em bases patentárias. Além disso, indicam que diversidade tecnológica, convergência e coesão semântica capturam diferentes dimensões da estrutura tecnológica e apresentam associações distintas com os indicadores de desempenho inovativo, aspecto consistente com as contribuições de Fleming (2001), Hsu et al. (2020) e Lin e Chou (2025).

Como limitações, destaca-se a ampla abrangência temática do corpus recuperado, a utilização exclusiva de documentos patentários e a ausência de variáveis institucionais, organizacionais e mercadológicas nos modelos analisados. Pesquisas futuras podem incorporar procedimentos adicionais de refinamento temático, integrar variáveis complementares associadas à apropriação tecnológica e realizar análises comparativas entre diferentes setores tecnológicos, contribuindo para aprofundar a compreensão das relações entre estrutura do conhecimento, convergência tecnológica e desempenho inovativo.

## Referências

Archibugi, D., Planta, M., 1996. Measuring technological change through patents and innovation surveys. *Technovation* 16 (9), 451–519. https://doi.org/10.1016/0166-4972(96)00031-4

Arundel, A. 2001. The relative effectiveness of patents and secrecy for appropriation. Research Policy, 30(4), 611–624. https://doi.org/10.1016/s0048-7333(00)00100-1

Bortolini, M., Galizia, F.G., Mora, C., Botti, L., Rosano, M., 2018. Bi-objective design of fresh food supply chain networks with reusable and disposable packaging containers. *J. Clean. Prod.* 184, 375–388. https://doi.org/10.1016/j.jclepro.2018.02.231

Brody, A.L., Strupinsky, E.P., Kline, L.R., 2001. *Active Packaging for Food Applications*. CRC Press. https://doi.org/10.1201/9780367801311

Caleb, O.J., Mahajan, P.V., Al-Said, F.A.J., Opara, U.L., 2013. Modified atmosphere packaging technology of fresh and fresh-cut produce and the microbial consequences—a review. *Food Bioprocess Technol.* 6 (2), 303–329. https://doi.org/10.1007/s11947-012-0932-4

Choi, D.S., Son, J.Y., Kim, J.S., Kim, Y.H., Park, C.W., Jung, H.M., Hwang, S.W., 2020. Optimum packaging design of packaging tray and cushion pad of Korean pears for exporting using FEA simulation. *J. Korean Soc. Ind. Converg.* 23 (5), 843–852.

Chonhenchob, V., Chinsirikul, W., Singh, S.P., 2012. Current and innovative packaging technologies for tropical and subtropical fruits. In: *Tropical and Subtropical Fruits: Postharvest Physiology, Processing and Packaging*, pp. 115–134. https://doi.org/10.1002/9781118324097.ch7

Cunha, K.C.T., Martens, C.D.P., Marcolin, C.B., 2025. Mining textual fields from patent documents: systematic review. *Manag. Rev. Q.* 1–37. https://doi.org/10.1007/s11301-025-00555-z

de Mello Júnior, N.R.C., Barbosa, L.M.A., Batista, E.V., de Souza Santos, M., Aurélio, R.L., Ospino, A.D.G., dos Santos Cavalcante, P.R., 2025. Modified atmosphere packaging in the export chain of tropical fruits: efficiency, limitations, and logistical implications. *Obs. Econ. Latinoam.* 23 (10), e11821. https://doi.org/10.55905/oelv23n10-070

Fleming, L., 2001. Recombinant uncertainty in technological search. *Manage. Sci.* 47 (1), 117–132. http://www.jstor.org/stable/2661563

Griliches, Z. 1990. Patent Statistics as Economic Indicators: A Survey. National Bureau of Economic Research. https://doi.org/10.3386/w3301

Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. The NBER patent citation data file: lessons, insights and methodological tools. https://doi.org/10.3386/w8498.

Hsu, P.H., Lee, D., Tambe, P., Hsu, D.H., 2020. Deep learning, text, and patent valuation. *Text and Patent Valuation*. https://dx.doi.org/10.2139/ssrn.3758388

Jang, H., Kim, S., Yoon, B., 2023. An explainable AI (XAI) model for text-based patent novelty analysis. *Expert Syst. Appl.* 231, 120839. https://doi.org/10.1016/j.eswa.2023.120839

Kader, A.A., 2005. Increasing food availability by reducing postharvest losses of fresh produce. In: *V International Postharvest Symposium* 682, pp. 2169–2176. https://doi.org/10.17660/ActaHortic.2005.682.296

Kerry, J., Butler, P. (Eds.), 2008. *Smart Packaging Technologies for Fast Moving Consumer Goods*. John Wiley & Sons. https://doi.org/10.1002/9780470753699

Kim, K.H., Han, Y.J., Lee, S., Cho, S.W., Lee, C., 2019. Text mining for patent analysis to forecast emerging technologies in wireless power transfer. *Sustainability* 11 (22), 6240. https://doi.org/10.3390/su11226240

Kumar, A., Kumar, N., Panghal, A., 2025. Sustainable active packaging: cellulose-based vs LDPE films for extending jujube shelf life. *Nutr. Food Sci.* 55 (6), 1075–1090. https://doi.org/10.1108/NFS-04-2025-0148

Lee, J.Y., Choi, J.W., Choi, J.H., Lee, B.H., 2023. Text-mining analysis using national R&D project data of South Korea to investigate innovation in graphene environment technology. *Int. J. Innov. Stud.* 7 (1), 87–99. https://doi.org/10.1016/j.ijis.2022.09.005

Lin, T.Y., Chou, L.C., 2025. A systematic review of artificial intelligence applications and methodological advances in patent analysis. *World Pat. Inf.* 82, 102383. https://doi.org/10.1016/j.wpi.2025.102383

Mahajan, P.V., Caleb, O.J., Singh, Z., Watkins, C.B., Geyer, M., 2014. Postharvest treatments of fresh produce. *Philos. Trans. R. Soc. A* 372 (2017). https://doi.org/10.1098/rsta.2013.0309

Marsh, K., Bugusu, B., 2007. Food packaging—roles, materials, and environmental issues. *J. Food Sci.* 72 (3), R39–R55. https://doi.org/10.1111/j.1750-3841.2007.00301.x

Mejia, C., & Kajikawa, Y. 2025. Patent research in academic literature: Landscape and trends with a focus on patent analytics. Frontiers in Research Metrics and Analytics, 9. https://doi.org/10.3389/frma.2024.1484685

Nimitkeatkai, H., Techavuthiporn, C., Boonyaritthongchai, P., Supapvanich, S., 2022. Commercial active packaging maintaining physicochemical qualities of carambola fruit during cold storage. *Food Packag. Shelf Life* 32, 100834. https://doi.org/10.1016/j.fpsl.2022.100834

Nur, P.S., Azhar, A.H., Pak-Dek, M.S., Ramli, N.S., Rukayadi, Y., Jaafar, A.H., Mediani, A., 2025. Potential application and mechanism of hexanal in extending shelf life and safety of tropical fruits: A systematic review. *Int. Food Res. J.* 32 (2). https://doi.org/10.47836/ifrj.32.2.03

Panou, A., Lazaridis, D.G., Karabagias, I.K., 2025. Application of smart packaging on the preservation of different types of perishable fruits. *Foods* 14 (11), 1878. https://doi.org/10.3390/foods14111878

Pavitt, K. 1985. Patent statistics as indicators of innovative activities: Possibilities and problems. Scientometrics, 7(1-2), 77–99. https://doi.org/10.1007/bf02020142

Realini, C.E., Marcos, Begônia., 2014. Active and intelligent packaging systems for a modern society. *Meat Sci.* 98 (3), 404–419. https://doi.org/10.1016/j.meatsci.2014.06.031

Song, B., Luan, C., Liang, D., 2023. Identification of emerging technology topics (ETTs) using BERT-based model and semantic analysis: A perspective of multiple-field characteristics of patented inventions (MFCOPIs). *Scientometrics* 128 (11), 5883–5904. https://doi.org/10.1007/s11192-023-04819-x

Tilahun, S., Choi, H.R., Kim, S.H., Park, D.S., Lee, Y.M., Solomon, T., Jeong, C.S., 2019. Fruit size and placement in packaging affect firmness-related quality attributes of muskmelon (*Cucumis melo* L.) fruit. *Hortic. Sci. Technol.* 37 (2), 246–255. https://doi.org/10.12972/kjhst.20190024

Tong, X., Lu, Y., Motohashi, K., 2026. Prediction of patent grant using interpretable citation-guided graph neural networks. *Scientometrics* 1–32. https://doi.org/10.1007/s11192-025-05515-8

Tong, X., Zeng, L., Lu, Y., 2025. What matters in patent claims: structural and semantic-based feature extraction and embedding optimization. *Scientometrics* 1–25. https://doi.org/10.1007/s11192-025-05427-7

Yam, K.L., Takhistov, P.T., Miltz, J., 2005. Intelligent packaging: concepts and applications. *J. Food Sci.* 70 (1), R1–R10. https://doi.org/10.1111/j.1365-2621.2005.tb09052.x