sustainability

MDPI

Article

Exploring the Co-Evolutionary Dynamics Between Scientific Findings and Industrial Innovations in Tunnel Construction Safety: A 22-Year Empirical Analysis

Guanfeng Yan [1,2,*] and Lingrong Yan [3]

[1] School of Engineering, Sichuan Normal University, Chengdu 610101, China
[2] School of Civil Engineering, Southwest Jiaotong University, Chengdu 610031, China
[3] College of Arts, Humanities and Social Sciences, The University of Edinburgh, Edinburgh EH8 9YL, UK
* Correspondence: yanguanfeng@swjtu.edu.cn

Abstract

The interaction between safety-related scientific findings and industrial innovation is of great benefit in improving safety management during the tunnel construction. However, few studies have aimed to investigate the co-evolution between scientific knowledge and technological development in underground project safety, where substantial lexical heterogeneity exists in the cross-domain corpora. To address this research gap, this study proposes an integrated framework combining a novel natural language processing method with temporal econometrics. A cross-domain corpus comprising 6404 Web of Science (WoS) publications and 20,932 high-quality patent documents over a 22-year period (2004–2026) is collected in the field of underground project safety. Then, the dense semantic encoder bge-large-en-v1.5 is adopted to map the textual dataset onto a standardized high-dimensional vector space. The trajectories of quarterly safety-related science-technology semantic overlay are obtained, and then the Vector Autoregressive (VAR) method is applied to the time series stabilized via log-differencing to measure the temporal interaction patterns among the scientific activity, technological innovation, and semantic alignment. Empirical results show that the semantic overlay metric stabilized at a plateau ranging from 0.91 to 0.93. Moreover, a temporal permutation benchmark indicates that part of this similarity results from the shared engineering domain between the two datasets. The impulse response analysis based on a VAR model reveals the temporal fluctuations and dynamic adjustment among scientific activity, technological innovation, and semantic alignment rather than a unidirectional transmission pathway. In addition, the keyword analysis based on normalized document frequency exhibits a divergence in priorities between scientific research and patent-based innovation, while the shared keywords indicate both domains pay great attention to primary risks, such as water inrush and rock instability. These empirical data-driven discoveries provide insights into the science-technology co-evolution in tunnel construction safety. The integrated framework offers methodological support for monitoring knowledge alignment and developing proactive safety management strategies for the sustainable development of underground infrastructure.

Keywords: sustainable underground infrastructure; tunnel construction safety; the vector autoregressive (VAR) method; science-technology co-evolution; semantic alignment

Check for updates

Academic Editor: Payam Rahnamayiezekavat

Received: 20 September 2026
Revised: 21 October 2026
Accepted: 23 October 2026
Published: 24 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Rapid urbanization across the world has intensified the demand for more comprehensive infrastructure networks [1]. The growing demand drives new projects, particularly ultra-deep underground projects, into complex and challenging geological domains, where traditional safety management frameworks fail to adequately accommodate these geological conditions [2]. Challenging geotechnical environments, such as densely faulted ground [3], highly non-uniform geostresses [4], and karst water networks with high pressure [5], make tunneling projects vulnerable to various hazards. These range from sudden failures like severe rockbursts and progressive deformation of surrounding rocks to catastrophic water and mud inrushes. To address these risks, the global ecosystem of tunnel engineering has followed two parallel trajectories: the scientific and industrial frontiers.

Academic scholars from universities and research institutions have adopted a range of analytical and numerical methods to establish failure criteria and predict potential failure risks. Researchers have developed constitutive models to assess safety by predicting the deformation responses of surrounding rock [6]. Scholars have also employed numerical simulations to explore the limit support pressure of the tunnel face under various layered ground conditions, with the aim of optimizing support design [7]. More recently, researchers have proposed a framework based on machine learning algorithms to predict risk for tunnel crossing water-rich faults to improve risk forecasting capacity [8]. On the industrial side, construction equipment suppliers and contractors prioritize practicality and aim to develop site-deployable equipment and operating protocols [9]. For example, an adaptive cutterhead has been designed to actively adjust its cutting diameter through hydraulic expansion to respond to varying geological conditions [10]. In addition, a real-time warning system based on geological conditions and tunneling parameters has been designed to alert operators to potential jamming incidents [11]. Collectively, the work from both frontiers has contributed to tunnel construction safety, each addressing important challenges from revealing rock strain-stress behavior to delivering sites with construction hardware and protocols, and together these two fronts establish a robust foundation for underground construction safety.

Despite these important contributions, the two trajectories have advanced in parallel, and the interaction between scientific findings and industrial innovation remains underexplored. As a result, this decoupling makes the diffusion pathways through which safety-related knowledge from academic discoveries become reflected in industrial innovation poorly understood. Exploring the alignment between science and technology matters not only theoretically but also practically. In theory, the classic problem of knowledge diffusion between research institutions and firms has been a core concern for innovation studies [12,13], and translating scientific discoveries into tangible industrial facilities or protocols is the essential phase [14]. In practice, the temporal dynamics between scientific findings and industrial innovation can influence how soon novel risk mitigation devices and protocols become available at project sites [15–17]. The delayed adoption of innovative safety paradigms may increase risks in complex projects such as ultra-deep mountain tunnels and sub-sea crossing passages. In some fields, such as renewable energy [18], researchers have employed various quantitative methods to explore the co-evolutionary dynamics between science and technology to inform both industrial research and development (R&D) and governance strategies. However, similar investigations remain absent within tunneling engineering, where the interaction between scientific findings and industrial innovation remains insufficiently understood.

Among the quantitative methods revealing the science-technology co-evolutionary trajectory, two approaches are most popular: namely, the citation-based approach and the text-based approach. For the former, the key metric of knowledge flow is non-patent refer-

ences (NPRs), which are citations of academic publications within patent documents [19]. For the latter, the knowledge flow is primarily measured by keyword co-occurrence (lexical matching) between academic publications and patent documents [20]. Despite their wide adoption for macro-level analysis in various fields, both approaches have several limitations when applied to specific fields such as tunneling engineering. First, the terminology in the academic publications generally includes abstract concepts describing complex geomechanical phenomena, such as viscoelastic constitutive degradation and shear strain localization. In contrast, the patent texts are claims-oriented and contain extensive descriptions of the equipment and protocols, leading to little lexical overlap with academic literature. Consequently, simple keyword matching results in substantial semantic drift. Second, these conventional methods are only capable of capturing the static relationship between academic and patent corpora, and the dynamic co-evolution between science and technology remains unexplored. Investigating the temporal process has great practical value. For all practitioners in charge of safety for complex underground projects, knowing the dynamic relationship between scientific advancement and industrial innovation can provide an important reference for developing a proactive safety management framework. For example, identifying the evolutionary trajectory of the alignment between scientific findings and technological innovation may help project managers to anticipate emerging technology needs and reallocate resources, including safety investments and safety-related training [21]. However, the temporal interaction between scientific research and industrial innovation in tunnel construction safety remains largely unexplored.

In summary, three interrelated gaps exist in the field of underground project safety: first, the empirical analysis of dynamic co-evolution between academic findings and industrial innovation remains absent; second, traditional methods exploring knowledge transfer rely on explicit citations and keyword matching and are not suitable for fields where the lexical heterogeneity is high between academic and industrial corpora; and third, the temporal interaction and adjustment patterns between scientific findings and industrial innovation have not been systematically identified. Addressing these gaps provides multiple benefits for the whole industry in several aspects. On the one hand, the empirical findings enrich the understanding of the evolving relationship between scientific and technological knowledge among the practitioners. On the other hand, data-driven conclusions offer valuable advice on developing proactive safety management measures, such as continuously monitoring emerging areas of science-technology convergence and divergence. In addition, this study illustrates a transferable methodological pipeline that can also be applied to other similar heavy engineering domains.

To address these gaps above, this study proposes a novel research framework integrating natural language processing methods based on deep learning with temporal econometrics. In detail, simple keyword matching has been replaced by dense semantic embeddings, which project the corpora of academic publications and industrial patent claims in the field of tunnel construction safety onto a high-dimensional continuous vector space. This step aims to quantify the quarterly cross-domain similarity over a 22-year timespan. Then, the vector autoregressive (VAR) method is applied to the time series to capture the dynamic temporal interactions and response patterns among scientific activity, industrial innovation, and cross-domain semantic alignment. This integrated framework provides an empirical basis for understanding the science-technology co-evolution and constitutes a cornerstone for developing proactive underground project safety measures and advancing the sustainable development of the industry.

2. Literature Review and Theoretical Framework

2.1. Theoretical Foundations of Science–Technology Co-Evolution in Engineering Domains

The interaction between scientific discoveries and technological innovations constitutes a foundational cornerstone of innovation studies [22,23]. The classic linear model, proposed by Bush [24], argued that scientific research was an external knowledge supplier that unidirectionally propelled industrial innovation in a sequential way. Despite being criticized for oversimplification, this linear assumption reflected the postwar landscape and was supported by substantial industrial breakthroughs that large-scale investment in fundamental science had produced. However, over recent decades, contemporary socio-technical transition theory [13,25] and modern evolutionary economics [26,27] have pointed out that the relationship between science and technology is characterized by a highly nonlinear and reciprocal co-evolutionary process. This systemic perspective suggests that academic institutions and industrial firms function as two distinct but coupled knowledge-generating domains [28]. On the one hand, researchers place great emphasis on causal inference, physical principles, and theoretical models; on the other hand, industrial practitioners translate these findings into technologies, practical solutions, and commercialized applications [29].

In engineering domains where safety is a critical issue, maintaining the alignment between scientific and industrial sectors is essential. The synchronous development of theoretical models and industrial innovation provides significant advantages. By contrast, asynchronous decoupling leads to profound systemic vulnerabilities, which may result in significant safety incidents. Therefore, exploring the temporal coupling and longitudinal alignment dynamics between scientific discoveries and technological innovations is critical, particularly for safety-critical industries such as tunnel construction, which face substantial safety issues.

2.2. Methodological Shift in Measuring Science-Technology Relationships

To empirically quantify the interaction between scientific and technological domains, previous bibliometric and econometric studies have largely relied on citations appearing on the patent front pages, primarily Non-Patent References (NPRs) within patent documents [30–32]. While the tracking based on NPRs features high traceability, recent studies in the field of scientometrics have revealed critical systemic flaws of this empirical parameter:

(1) Biases Induced by Multiple Parties: The patterns underlying patent citations vary across different countries and regions. Unlike academic citations, which are voluntarily made by authors, patent citations result from the combined behavior of applicants and examiners under the legal framework, which inevitably introduces systematic bias [33,34]. For example, the United States Patent and Trademark Office (USPTO) grants examiners greater discretionary authority, while other countries' offices lay more emphasis on the prior art of the applicants. Therefore, US examiners generally insert additional references during examination, which reflect examiners' citation strategies rather than applicants' actual knowledge, and thus the citation volumes of US patents are generally higher than those in other regions like Europe. However, examiners do not engage in research and development work and thus may lack industry-level technical understanding of the patents. Furthermore, their examination habits may differ, which also introduces substantial discrepancies to the citation records [33]. Except for the distortions mentioned above, applicants tend to cite prior art to steer examination to favorable examiners, which decouples the patent citation from actual knowledge diffusion further [34].

(2) Ignorance of Informal Science-Technology Interaction: Citation relationships capture only part of the science-technology interaction. Three important types of informal interaction between science and technology have not been taken into account: (1) practical experience, instinct, and heuristics that are widely incorporated in practice but not recorded

in formal documents; (2) early conceptual insights that are disseminated through conferences and workshops prior to R&D work; and (3) shared problem-solving frameworks that become an industry consensus. Unlike the citations of formal documents, these types of informal interaction between science and technology do not leave verifiable records, which are ignored by the citation data. As a result, this conventional approach underestimates the intensity of science-technology interaction, especially for fields where practical experience plays an important role, such as civil engineering [35].

In order to address the aforementioned flaws, subsequent studies have proposed keyword-based textual analysis, co-word network analysis, and classification-based overlap mapping [36,37], which aim to quantify the alignment across different domains, such as matching the International Patent Classification (IPC) clusters with WoS subject categories [38]. Nevertheless, a critical methodological deficiency, which is the linguistic heterogeneity across different corpora, exists among these surface-level lexical approaches and manifests as follows:

(1) Disciplinary Linguistic Difference: Taking civil engineering as an example, academic scholars tend to use formal vocabulary to describe theoretical models and abstract phenomena, such as geomechanical instability (e.g., constitutive strain localization, viscoelastic rock mass degradation, and energy dissipation tensors). In contrast, the claims-oriented terms adopted by the industrial patent applicants, such as elongated anchoring rods, modular segmented shoring apparatus, and circumferential injection conduits, are lexically distinct from the formal vocabulary used by the academic publications.

(2) Inherent Defects of Keyword Matching: Traditional string-matching and co-occurrence network analysis rely on exact keyword matching or simple co-occurrence counting [36], neglecting the fact that different terms may represent the same meaning [39]. These approaches introduce two inherent defects: (1) high false-negative rates, where related documents may be missed due to the mismatch of exact keywords; (2) substantial semantic drift, where variation in the lexical context near the core concept leads to misclassification of semantically relevant corpora. Consequently, these approaches fail to recognize the interaction between an academic publication and a patent filing, which describe the same underground risk using different domain terms.

To address this inherent methodological bottleneck, contextualized embedding models, such as transformer-based semantic models, are required to be adopted. By mapping cross-domain corpora onto vector spaces, these models place semantically similar content in proximity. So, proximity measurement is invariant to lexical changes.

2.3. Research Gaps in Understanding Scientific-Technology Co-Evolution in Underground Construction Safety

The cross-domain semantic linkage frameworks have been widely explored in frontier technology fields, including artificial intelligence and biotechnology, to trace science-technology interaction [40,41]. These frameworks facilitate science-technology co-evolution within the industry and enable the detection of the evolutionary path, allowing scholars to track how scientific findings and industrial innovations inform each other. In the context of underground project safety, these frameworks remain largely unexplored. This gap is significant because complex underground projects are subject to unique challenges—including geological uncertainties, nonlinear environmental responses, and catastrophic hazards such as severe rock bursts due to extremely high in situ stress, water and mud inrushes in karst areas, and collapse of the tunnel face in weak ground [42,43]. Considering the severe risks, macro-level safety governance frameworks are essential for deep tunneling projects [44]. However, existing studies related to tunnel construction safety are overwhelmingly confined to micro-level issues. The majority of existing literature prioritizes numerical simulations,

laboratory-scale model tests, and risk assessment for specific projects [45]. These studies that focus on micro-level issues shed light on project-specific failure mechanisms and localized safety management. However, the conclusions do not address a macro-level question, which is the temporal interaction between scientific findings and industrial innovation.

Scientific findings and industrial innovation reflect two interacting knowledge domains with different objectives and terminology. Scientific activities focus on studying mechanisms, developing theories, and predicting phenomena, while technological activities lay great emphasis on practical solutions and the development of new risk-mitigation equipment. The science-technology semantic alignment reflects the conceptual relationship between the two knowledge domains. Measuring the semantic alignment makes it possible to trace the evolutionary path of the science–technology relationship. In addition, it is important to explore the temporal interactions among scientific activity, technological innovation, and cross-domain semantic alignment to reveal the characteristics of these interactions within the knowledge ecosystem.

To address these gaps and provide valuable insights into data-driven proactive safety governance, this study proposes an analytical framework that integrates cross-domain semantic vectorization and a covariance-stationary Vector Autoregressive (VAR) model to investigate the temporal interactions and co-evolutionary patterns between scientific findings and industrial innovation in the field of tunnel construction safety.

3. Methodology

3.1. Data Collection and Pre-Processing

In this study, a multi-stage strategy was adopted to build a representative and fine-grained database in the field of tunnel construction safety. The database drew on two interrelated sources: scientific findings and technological innovations.

The scientific research data was collected from the WoS Core Collection. Due to the descriptive flexibility of academic papers, the same technical concepts tend to be expressed through diverse phrasing. To ensure comprehensive coverage, a set of cluster-specific Boolean queries was designed. The query syntax is as follows:

Query = (“tunnel” OR “tunneling” OR “tunneling” OR “underground engineering” OR “underground space” OR “subway construction” OR “metro construction”) AND (“construction” OR “excavation” OR “boring” OR “shielding” OR “TBM” OR “tunnel boring machine” OR “shield operator” OR “lining” OR “jacking”) AND (“safety” OR “risk” OR “accident” OR “hazard” OR “disaster” OR “collapse” OR “collapse failure” OR “water inrush” OR “mud inrush” OR “failure mode” OR “occupational health” OR “injury” OR “fatality” OR “safety management”)

This syntax was designed to represent three fundamental thematic dimensions: (i) the target engineering domain was confined to tunnel projects; (ii) various construction methods and excavation techniques were considered; and (iii) the core concern was on various aspects of tunnel safety, such as risk, accident, and disaster. The timespan was set from January 2004 to March 2026, restricted to peer-reviewed English ‘articles’ and ‘reviews’ indexed in Science Citation Index Expanded (SCI-EXPANDED) and Social Science Citation Index (SSCI). This restriction was imposed to ensure the inclusion of only peer-reviewed publications, thereby ensuring a certain level of academic quality. Given the inherently interdisciplinary nature of tunnel construction safety, no subject category restriction was imposed to maximize thematic coverage. Duplicate records were removed by checking the unique WoS accession number (UT). Finally, 6404 relevant academic records were collected, from which the title (TI), abstract (AB), and author keywords (DE) fields were extracted to build the raw dataset.

Global patent data related to tunnel construction safety were retrieved from the patents-public-data.patents.publications repository available on Google Cloud BigQuery. This database includes key information on patents across major global patent jurisdictions. It should be noted that the query for patent information retrieval used a more compact keyword string, and the string is as follows:

Query = ("tunnel" OR "shield machine" OR "TBM") AND ("safety" OR "risk" OR "accident" OR "collapse" OR "hazard")

From a methodological standpoint, a more compact query design is necessary and a deliberate choice to prevent semantic interference from unrelated technical domains in the global patent database. Unlike academic literature, patents are classified by official hierarchical classification systems. The targeted patents in this study were restricted to two International Patent Classification (IPC) subclasses: E21D (Shafts; Tunnels; Galleries) and E02D (Foundations; Soil-shifting). These IPC subclasses serve as a domain filter and restrict the scope to tunnel and underground engineering to avoid irrelevant records from other technical domains, and the compact keyword string further concentrates on safety-related patents. Within the IPC-restricted engineering domain, a long keyword combination can introduce semantic redundancy and compromise query precision. Therefore, a streamlined keyword combination helps ensure high query precision, yielding 20,932 relevant patent documents filed between January 2004 and March 2026. For each patent, the patent title, abstract, and publication date were extracted.

Finally, both datasets were aligned to the same timeline of 89 consecutive quarters (2004-Q1 to 2026-Q1). As a result, a unified empirical database comprising 27,336 records was established for subsequent mapping into a 1024-dimensional deep semantic space. Table 1 lists the top five regions contributing to tunnel construction safety in terms of both academic publications and patent output. Notably, the divergence between the two lists reveals a geographical asymmetry: countries like Australia and the United Kingdom show strong academic output in publications but industrial patent activity remains overwhelmingly concentrated in major manufacturing powers. This asymmetry highlights a marked global imbalance between knowledge production and industrial innovations in the underground engineering sector.

Table 1. The top five most prolific jurisdictions contributing to tunnel construction safety (2004Q1–2026Q1).

Scientific Research (WoS)			Industrial Patents (Google Patents)		
Rank	**Country/Region**	**Paper Count**	**Rank**	**Country/Region**	**Patent Count**
1	China	5274	1	China	20,110
2	Australia	215	2	Japan	252
3	United Kingdom	198	3	South Korea	220
4	South Korea	129	4	WIPO (WO) *	59
5	Italy	128	5	United States	59

Note: (1) * WIPO denotes international applications filed under the Patent Cooperation Treaty. (2) The sample counts reflect unique records within the domain-specific boundaries defined in Section 3.1.

To assess whether the different retrieval strategies led to substantial differences in thematic coverage, a post-retrieval coverage validation was carried out for the cross-domain corpora of 6404 WoS publications and 20,932 patent records. Titles and abstracts were adopted for both datasets. Two sets of diagnostic concepts were evaluated independently: five common construction methods (TBM, shield tunneling, NATM (New Austrian Tunneling Method), drill-and-blast, and cut-and-cover) and seven important safety-related topics

(risk, safety, collapse/failure, hazard, monitoring/detection/inspection, water inrush/inflow, and rockburst). For each set of diagnostic concepts, the number and proportion of documents containing the corresponding expressions were calculated independently within each dataset. The results show that both corpora cover multiple construction methods and safety-related topics, but the corresponding representation varies across domains. For both corpora, documents on mechanized tunnelling are predominant, while those related to NATM and cut-and-cover account for a comparatively small proportion. More details could be found in Supplementary Table S1.

3.2. Dense Semantic Embedding of Scientific and Technical Documents

To address the lexical heterogeneity between academic terminology and industrial patent documents, a deep-learning encoding strategy was adopted to process the dataset built in Section 3.1. Let $D^{wos} = \{d_1^{wos}, d_2^{wos}, \ldots, d_N^{wos}\}$ denote the corpus of N ($N = 6404$) scientific papers, where each d_i^{wos} is constructed by concatenating the Document Title (TI), Abstract (AB), and Author Keywords (DE) of the corresponding publication. Similarly, let $D^{pat} = \{d_1^{pat}, d_2^{pat}, \ldots, d_M^{pat}\}$ denote the corpus of M ($M = 20{,}932$) industrial patents, where each d_i^{pat} is constructed by concatenating the Title and Abstract of the corresponding patent.

A high-performance dense encoder (bge-large-en-v1.5) was utilized to map each unstructured text document to a 1024-dimensional vector space [46], denoted as R^{1024}. The high-dimensional encoding helps ensure that rich connotations in the field of tunnel construction safety are mapped onto vector representations based on contextualized semantics instead of simple keyword matching. To ensure that similarity metrics are insensitive to the varying lengths of source texts, all field vectors are projected onto a unit hypersphere via L2 normalization [47]:

$$v_i^{wos} = \frac{\emptyset(d_i^{wos})}{\left\|\emptyset(d_i^{wos})\right\|_2}, \qquad v_j^{pat} = \frac{\emptyset\left(d_j^{pat}\right)}{\left\|\emptyset\left(d_j^{pat}\right)\right\|_2} \tag{1}$$

where $\emptyset(\cdot)$ is the BGE encoder, and v_i^{wos} and v_j^{pat} denote the normalized 1024-dimensional semantic coordinates for the i-th academic papers and the j-th industrial patents, respectively.

3.3. Temporal Aggregation and Cross-Domain Semantic Overlay Extraction

To transform static semantic vectors into a continuous temporal trajectory for tunnel safety knowledge fusion, a quarter-level pooling mechanism was established, with data strictly grouped by the publication date of academic papers and the publication date of industrial patents. The complete timeline was discretized into 89 consecutive quarterly bins (from 2004-Q1 to 2026-Q1) [48], represented by $\mathcal{T} = \{t_1, \ldots, t_{89}\}$. Accordingly, for any quarter t, Ω_t^{wos} and Ω_t^{pat} respectively denote the index sets of papers and patents whose publication or publication dates lie in that specific three-month interval.

To integrate different research scopes into a unified domain representation, a mean-pooling operation was conducted to obtain the collective knowledge centroids [49]:

$$S_t = \frac{1}{\left|\Omega_t^{wos}\right|} \sum_{i \in \Omega_t^{wos}} v_i^{wos}, \qquad T_t = \frac{1}{\left|\Omega_t^{pat}\right|} \sum_{j \in \Omega_t^{pat}} v_j^{pat} \tag{2}$$

Here, the Scientific Knowledge Centroid (S_t) and Technological Frontier Centroid (T_t), which are obtained via mean-pooling over the quarterly text corpora, reflect time-varying semantic centroids that represent academia and industry within the R^{1024} semantic embedding space. These centroids capture the latent semantic features of their respective knowledge domains and are designed for macro-level longitudinal analysis. Therefore, this approach provides a consistent time-series representation for the subsequent science-technology semantic overlay analysis and VAR analysis. However, mean pooling may make

small or newly emerging knowledge or technology clusters with limited document volumes less visible.

The temporal integration and cognitive proximity between the two frontiers are quantified via the Science-Technology Semantic Overlay metric ($ST_Overlay$), which measures the cross-domain semantic overlap through the cosine similarity between the two centroids [50]:

$$ST_Overlay_t = \frac{S_t \cdot T_t}{\|S_t\|_2 \|T_t\|_2} \quad (3)$$

A high $ST_Overlay$ value reflects a higher level of semantic alignment between the academic and patent-based technological corpora, rather than directly measuring knowledge transfer or technology adoption [51]. Given the data processing method, the $ST_Overlay$ value demonstrates a domain-level semantic alignment rather than measuring individual technology clusters. In addition, to assess whether the high semantic overlay values result from the mean pooling and the shared disciplinary background of the cross-domain corpora, a temporal permutation null model was constructed. This benchmark keeps the semantic characteristics and centroid structure of the corpora but disrupts the actual temporal alignment between the academic and technological corpus.

3.4. Log-Difference Transformation for Variance Stabilization and Stationarity

Let $W_t = |\Omega_t^{wos}|$ and $P_t = |\Omega_t^{pat}|$ denote the raw counts of papers and patents for quarter t, respectively. Owing to the proliferation of large underground projects, the expansion of urban subterranean zones, and advances in high-performance TBM machines, the output of both papers and patents has grown, and both variables exhibit pronounced upward trends over the sample period. Applying direct time-series regressions to these raw counts can introduce a serious econometric risk of spurious regression, in which the estimates may largely represent common temporal trends.

To stabilize variance and promote covariance stationarity (i.e., $I(0)$), a log-differencing transformation (with a + 1 offset to preserve zero counts) is applied to the volume-scale variables (W_t and P_t) [52]. A standard first difference is applied to the naturally bounded overlay similarity ($ST_Overlay_t$) to remove its trend and align its scale with those of the transformed W_t and P_t. The transformations are specified as:

$$w_t = \Delta log_{10}(W_t + 1) = log_{10}(W_t + 1) - log_{10}(W_{t-1} + 1) \quad (4)$$

$$p_t = \Delta log_{10}(P_t + 1) = log_{10}(P_t + 1) - log_{10}(P_{t-1} + 1) \quad (5)$$

$$o_t = \Delta ST_Overlay_t = ST_Overlay_t - ST_Overlay_{t-1} \quad (6)$$

Here, w_t represents the growth rate of safety-related paper output, p_t reflects the growth rate of industrial patent components, and o_t measures the period-to-period change in cross-domain semantic alignment. To verify stationarity before time-series modeling, the Augmented Dickey–Fuller (ADF) unit root test is conducted [53]:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{k=1}^{p_{ADF}} \delta_k \Delta y_{t-k} + \varepsilon_t \quad (7)$$

where the optimal lag order p_{ADF} is determined via the Akaike Information Criterion (AIC), with a constant included in the deterministic specification [54]. The ADF unit-root tests reveal that p-values are below 0.01 for all three transformed series, leading to a rejection of the null hypothesis of a unit root. This indicates the absence of a unit root and supports the treatment of all the transformed series as stationary ($I(0)$), providing a necessary precondition for subsequent Vector Autoregressive (VAR) estimation.

3.5. Vector Autoregressive (VAR) Modeling and Impulse-Response Analysis

To examine the temporal interactions among scientific activity, technological activity, and science-technology semantic alignment [55], a vector autoregressive model was adopted. The endogenous vector $\boldsymbol{Y}_t$ was defined as follows:

$$\boldsymbol{Y}_t = [\Delta ln(1 + W_t), \Delta \ln(1 + P_t), \Delta A_t]' \tag{8}$$

where W_t denotes quarterly scientific publication output, P_t represents quarterly patent output, and A_t denotes the science-technology semantic alignment metric. The transformed variables were employed to satisfy the stationary requirements of VAR modeling.

The vector autoregressive model of order p [56], denoted as $VAR(p)$, is specified for the endogenous vector $\boldsymbol{Y}_t = [w_t, p_t, o_t]^T$, and the model is as follows:

$$\boldsymbol{Y}_t = \boldsymbol{C} + \sum_{k=1}^{p_{VAR}} \boldsymbol{A}_k \boldsymbol{Y}_{t-k} + \boldsymbol{u}_t \tag{9}$$

where $\boldsymbol{C}$ is a 3×1 constant vector, $\boldsymbol{A}_k$ is a 3×3 matrix of autoregressive lag coefficients to be estimated from the data, and $\boldsymbol{u}_t \sim NID(0, \Sigma)$ is the reduced-form white-noise error vector.

The optimal lag order p_{VAR} is selected using four standard information criteria: Akaike Information Criterion (AIC) [54], Bayesian Information Criterion (BIC) [57], Final Prediction Error (FPE) [58], and Hannan-Quinn Information Criterion (HQIC) [59]. Because these criteria place different emphasis on model complexity, they generally do not suggest the same lag order for the VAR system, and thus their recommendations were compared jointly. Finally, AIC and FPE recommended lag 7 while BIC and HQIC favored shorter specifications. Because this study aims to explore the multi-period dynamic responses, the VAR(7) specification was retained for the subsequent analysis. More details can be found in Section 4.2.

To unreal the dynamic associations among the variables, the orthogonalized impulse response functions (IRFs) are computed via Cholesky decomposition of the residual covariance matrix Σ [56,60]. The recursive ordering was specified as:

$$[\Delta ln(1 + W_t), \Delta \ln(1 + P_t), \Delta A_t] \tag{10}$$

Under this identification, four response relationships were examined over a 12-quarter horizon: (1) the response of science-technology semantic alignment to a shock in scientific publication activity; (2) the response of science-technology semantic alignment to a shock in patent activity; (3) the response of patent activity to a shock in scientific publication activity; and (4) the response of scientific publication activity to a shock in patent activity.

The orthogonalized IRFs and corresponding 95% confidence intervals were obtained for four responses. Because the orthogonalized IRFs depend on the Cholesky recursive ordering of the variables, the corresponding responses should be interpreted as dynamic associations under the specific VAR system instead of casual effects.

3.6. Verification and Validation Strategy

Considering the integration of cross-domain corpora, semantic embedding, and time-series modeling, verification and validation were carried out at three levels: dataset coverage, semantic-measure robustness, and verification of the econometric model.

First, the validation of dataset coverage was conducted to evaluate whether the different retrieval strategies for scientific publications and patent records led to huge differences in terms of research themes. Five common tunnel construction methods (TBM, shield tunneling, New Austrian Tunneling Method, drill-and-blast, and cut-and-cover) and seven important safety-related topics (risk, safety, collapse/failure, hazard, monitoring/detection/inspection, water inrush/inflow, and rockburst) were examined for the cross-domain

corpora. The proportion of records containing each construction method and topic was calculated for 6404 academic publications and 20,932 patent records. This step was adopted to verify the broad coverage of research themes and to capture the differences in the relative representation of diagnostic concepts (i.e., the five construction methods and seven safety-related topics) across the two sources.

Second, the validation of semantic-measure robustness was conducted for the semantic overlay metric. The quarterly semantic overlay trajectory was recalculated based on the corpora to verify the computational reproductivity. In addition, a null model based on temporal permutation was conducted to explore whether the cosine similarities resulted from the same disciplinary background and mean pooling. The quarterly technological centroids across the 89-quarter timeline were randomly shuffled with the scientific centroids fixed, and the mean semantic overlay was recalculated 1000 times. Then, the observed statistic was compared with the empirical null distribution obtained from the permutations.

Third, the verification of the econometric model, including stationarity testing, lag-order evaluation, and stability diagnostics, was conducted. The time series after transformation were assessed using the ADF test before VAR estimation, and the lag structure was compared using four criteria. Then, the stability of the selected VAR specification was assessed through the inverse roots of the autoregressive characteristic polynomial. The whole procedure comprehensively evaluates the reproductivity, robustness, and statistical adequacy of the research pipeline used in this study.

3.7. Keyword Identification and Normalized Document-Frequency Analysis

To complement the VAR impulse analysis, a two-step keyword analysis was conducted to explore the differences and commonalities in the thematic topics of the cross-domain corpora.

In the first step, high-frequency terms were identified separately for the WoS and patent-based technological corpora. Title and abstracts were concatenated and used as the input for both datasets. After standard natural language preprocessing, including lowercasing, tokenization, minimum-length filtering, and removal of generic academic, patent-template, and domain-wide terms, was applied to both corpora, raw term-occurrence frequencies were calculated, and the 15 highest-frequency terms were obtained across academic and patent corpora. This step was designed to identify the terms with the highest frequency of occurrence.

In the second step, the 15 selected terms were evaluated using normalized document frequency (NDF). NDF was defined as follows:

$$NDF_k = \frac{n_k}{N} \times 100\% \tag{11}$$

where n_k denotes the number of documents containing term k, and N is the total number of documents in the scientific or technological corpus. The calculation process was carried out separately for the 6404 academic publications and 20,932 patent documents. Then, the 15 reselected terms for each dataset were reordered according to their NDF values.

4. Empirical Results and Insights

4.1. Co-Evolutionary Trajectory and a Three-Stage Analysis

The empirical analysis begins with a macro-level assessment of the long-term co-evolutionary dynamics of academic publications and industrial patents in the domain of tunnel construction safety. Figure 1 reveals a non-linear evolutionary trajectory by plotting the quarterly volumes of scientific papers (W_t) and industrial patents (P_t) against their cross-domain semantic similarity index ($ST_Overlay_t$). Over the 2004–2026 period, the semantic overlay curve displays a descriptive three-stage evolutionary structure, progressing from the

Exploration stage, through the Paradigm Convergence stage, and finally into a High-Level Co-Evolutionary stage.

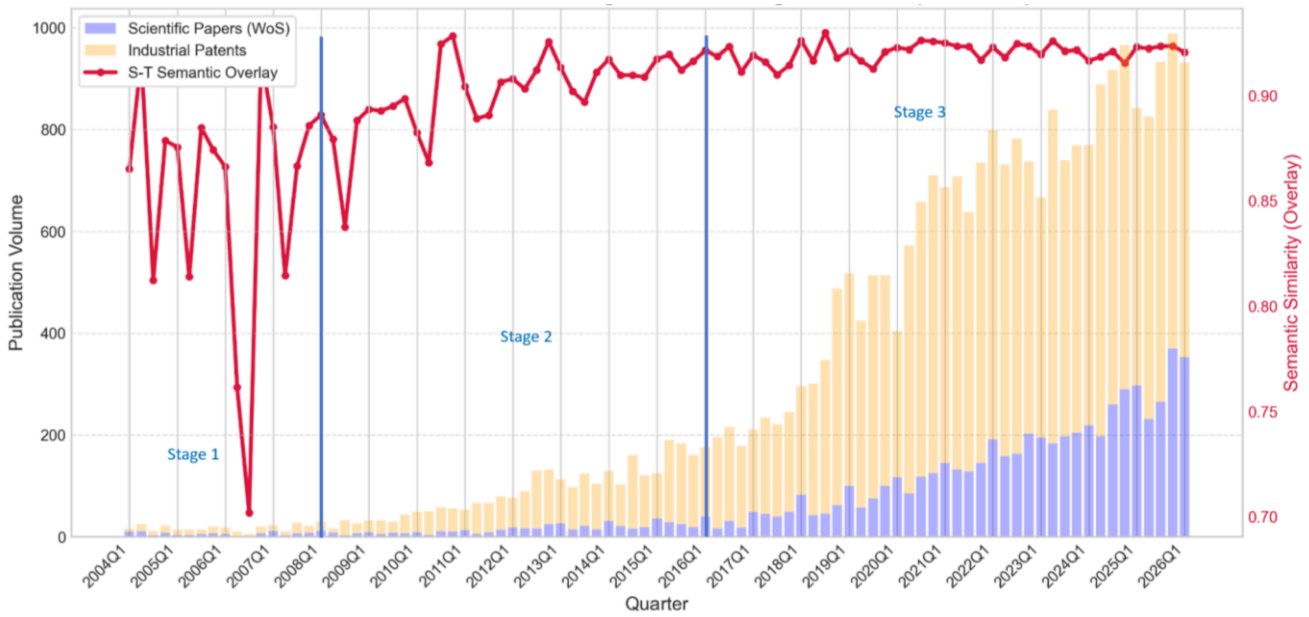


Figure 1. The macro-evolutionary trajectory of scientific papers (WoS) and industrial patents (Google Patents) alongside their dynamic semantic overlay (2004–2026).

(1) Stage 1: The Exploration stage (2004–2007): During the first stage, the tunnel safety sector produced extremely limited output in both scientific research and industrial innovations, with publication and patent counts remaining in single digits or low double digits. Importantly, the semantic overlay curve within this stage exhibits high-amplitude fluctuations, reaching its lowest value of 0.702 in 2006-Q3. This baseline volatility statistically reflects the substantial linguistic heterogeneity detected in this study's gap analysis. During this initial stage, scholarly work operated in disciplinary isolation, producing fragmented geological surveys. A representative study is that by Bilgin et al. [61], who conducted in situ investigations of a roadheader's performance in an inclined tunnel. Although their datasets were remarkably rich, their study remained strictly limited to the specific project site at Nub Cement Factory, and the authors made no attempt to extrapolate these geological parameters to any systematic safety framework or universal prediction model for tunnel construction. Meanwhile, industrial applicants submitted safety-related patents that were overly narrow and theoretically ungrounded. At this stage, Shanghai Urban Construction Design Co., Ltd. (Shanghai, China) filed a patent (CN2606180Y) for a built-in plastic valve bolt with a steel-reinforced concrete structure [62]—an improvement designed to enhance the sealing and anti-corrosion performance of a tunnel segment component. This utility model, however, contained no theoretical safety modeling or quantitative risk assessment. This component-focused patent contrasts with the macro-level academic findings during the period, and both share a common deficiency: they operated within their own isolated domains, with little interaction between scientific research and industrial innovation. The semantic gap between theoretical concepts and technological innovation was manifested as an unstable and poorly coordinated cross-domain alignment during this stage.

(2) Stage 2: The Convergence stage (2008–2016): Beginning in 2008, this field of tunnel construction safety underwent a profound structural transformation over the following eight years. The pronounced fluctuations of the red line representing the cross-domain semantic similarity index ceased, and the line showed a monotonically increasing trend, which propelled the metric, $ST_Overlay_t$, from a volatile corridor up into a high-level convergence

channel. This empirical convergence reflects the formal establishment of a unified paradigm bridging academic research and industrial innovation in tunnel construction safety. The global expansion of mechanized TBMs and deep underground urbanization drove both fields to address identical engineering bottlenecks [63,64]. Researchers increasingly focused on modeling the real-world field failure limits [65,66], such as dynamic rockbursts and karstic water inrushes. Meanwhile, patent applicants simultaneously paid more attention to strict geomechanical constitutive constraints [67]. This shared focus rapidly narrowed the cross-domain linguistic gap.

(3) Stage 3: The High-Level Co-Evolutionary stage (2017–2026): Since 2017, this period has witnessed a surge of intelligent construction technologies [68,69], featuring the widespread adoption of automated shield machines, dynamic face stability control, and integrated sensor networks for digital twinning. During this stage, the number of both research papers and industrial patents increased rapidly, with their combined quarterly total count peaking at nearly 1000, as shown in Figure 1. At the same time, the semantic overlay metric reached a stable plateau, with its value tightly bounded between 0.91 and 0.93. The high volume of publications and patents, together with the high values of the semantic overlay metric, indicates the increasing semantic convergence between scientific and technological knowledge domains [70].

As for the validation of semantic-measure robustness, the observed mean semantic overlay was 0.901366, while the mean of the 1000 temporally randomized alignments was 0.899069 (SD = 0.000378). The maximum randomized value was 0.900266 and was still lower than the observed mean value, with an empirical one-sided permutation p-value of approximately 0.001. More details could be found in Table S2. Nevertheless, the absolute difference between the observed and randomized mean values was quite small (0.002297), and it can be inferred that the disciplinary background shared by the two corpora leads to the elevated cosine similarity.

4.2. Dynamic Interactions Among Scientific Activity, Technological Innovation, and Science-Technology Semantic Alignment

To establish a rigorous foundation for VAR modeling and subsequent impulse response analysis, the three quarterly time-series metrics (WoS publications, patent applications, and cross-domain semantic overlay) were tested for unit roots using the Augmented Dickey–Fuller (ADF) test, with the lag length selected via the AIC [71]. Following the stabilization protocol proposed in Section 3.4, the ADF unit root tests were conducted on the three transformed series, and the results are listed in Table 2. The results demonstrated that the null hypothesis of a unit root was rejected for all transformed series at the 1% significance level. The ADF level statistics, which ranged from −4.99 to −8.83, fell well below the 1% critical value thresholds (−3.510 to −3.514), and the corresponding empirical p-values were below 0.001. This confirms that the transformed data form a stationary $I(0)$ system, effectively mitigating the risk of spurious regression in long-term econometric modeling. The next step is to select the optimal lag order (p) for the three-variable VAR system.

This involves a trade-off between model adequacy and estimation efficiency: too few lags may miss important dynamic linkages, while too many lags can deplete degrees of freedom. Table 3 lists the lag-selection criteria (AIC, BIC, FPE, and HQIC) considered for lags 0 to 8. For the BIC and HQIC, which impose heavy penalties on model complexity [57,59], Lag 2 (BIC = −15.23 *) and Lag 3 (HQIC = −15.72 *) are selected as the optimal choices. On the other hand, the AIC (AIC = −16.38 *) and FPE (FPE = 8.042×10^{-8} *), which prioritize explanatory capacity and overall goodness-of-fit [54], achieve their lowest values at Lag 7. This divergence represents a trade-off between parsimony and capturing multi-period dynamic relationships [60,72,73]. This study pays more attention to dynamic response

patterns, and thus Lag 7 is set to be the autoregressive order of the VAR model. Furthermore, the post-estimation stability diagnostics were conducted for the VAR(7) model to ensure its stability and the validity of subsequent simulations.

Table 2. Augmented Dickey–Fuller (ADF) unit root tests for transformed time-series metrics.

Variable	Transformation	Deterministic Term	Selected ADF Lag	ADF Statistic	1% Critical Value
WoS Publications	First difference of log (1 + x)	Constant	6	−5.12 ***	−3.514
Patent applications	First difference of log (1 + x)	Constant	2	−8.83 ***	−3.510
ST Semantic Overlay	First difference	Constant	6	−4.99 ***	−3.514

Note: This table lists the results of the ADF unit root test for three transformed metrics. The null hypothesis is that the series contains a unit root (i.e., non-stationary). The optimal lag lengths were obtained on the basis of the AIC for each ADF specification. A constant term was included in each test equation. *** means that the null hypothesis is rejected at the 1% significance level. The sample period covers 2004Q1–2026Q1.

Table 3. VAR lag order selection criteria for the three-variable VAR system.

Lag (p)	AIC	BIC (SIC)	FPE	HQIC
0	−14.79	−14.70	3.783×10^{-7}	−14.75
1	−15.59	−15.23	1.696×10^{-7}	−15.45
2	−15.86	−15.23 *	1.297×10^{-7}	−15.61
3	−16.08	−15.19	1.043×10^{-7}	−15.72 *
4	−16.09	−14.93	1.034×10^{-7}	−15.63
5	−16.08	−14.65	1.054×10^{-7}	−15.51
6	−16.28	−14.59	8.724×10^{-8}	−15.60
7	−16.38 *	−14.41	8.042×10^{-8} *	−15.59
8	−16.32	−14.08	8.746×10^{-8}	−15.42

Note: The VAR system includes three stationary transformed variables: Δlog_{10}(WoS Publications), Δlog_{10}(Patent Applications), and Δ(S-T Semantic overlay). AIC represents Akaike Information Criteria; BIC (SIC) reflects Bayesian (Schwarz) Information Criteria; FPE is Final Prediction Error; HQIC denotes Hannan-Quinn Information Criteria. Asterisks (*) mark the optimal lag order based on the corresponding criterion.

As shown in Figure 2, all 21 eigenvalues of the companion matrix lie inside the unit circle, which establishes asymptotic stability [60]. Therefore, the optimal system is determined to be Lag 7, suggesting that a seventh-order lag structure is appropriate for capturing the dynamic interrelations among the variables.

The VAR analysis explored the four impulse responses, as shown in Figure 3. In summary, the results showed that the responses are oscillatory and short-term rather than persistent and unidirectional. The details are as follows.

(a) Scientific Activity → Semantic Alignment: Following a one-standard-deviation shock in scientific publication growth, the response of semantic alignment is positive initially, subsequently turns negative at Horizon 2, and then exhibits several modest positive and negative fluctuations over the remaining horizons. This pattern indicates that changes in scientific activity are associated with the short-term oscillations in science-technology semantic alignment, and thus this response does not feature a persistent monotonic effect. Therefore, a dynamic adjustment process, rather than a science-to-technology transmission process, is supported.

(b) Patent activity → Semantic Alignment: The estimated response of semantic alignment to a one-standard-deviation shock in patent activity is quite modest and alternates

between positive and negative values over the 12-quarter horizon, with the confidence intervals largely overlapping the zero baseline. The results do not support a persistent directional response of semantic alignment to short-term changes in patent activity. Instead, patent activity and semantic alignment exhibit weak and fluctuating temporal associations.

(**c**) **Scientific Activity → Patent Activity:** A shock in WoS publication growth leads to an oscillatory response in patent growth. The response alternates between positive and negative movements across the 12-quarter horizon, and the uncertainty bands largely encompass the baseline. Although the response path exhibits short-term co-movement, there is still no robust evidence that the increase in scientific publication activity predicts an increase in patent activity.

(**d**) **Patent Activity → Scientific Activity:** The response trajectory of WoS publication growth to a shock in patent activity also shows both positive and negative movements. The estimated responses remain modest and relatively short-lived, with the confidence intervals including zero. Therefore, a stable unidirectional relationship in which patent activity consistently predicts subsequent scientific publication growth is not supported by the results.

In summary, the impulse-response analysis shows that the dynamic associations among the scientific findings, patent-based technological innovations, and their semantic alignment feature primarily short-term fluctuations and temporal adjustment rather than a persistent unidirectional relationship. It seems that the science-technology system in the sector of underground construction safety exhibits complex co-evolutionary dynamics instead of a simple linear transmission from scientific research to technological innovation or vice versa.

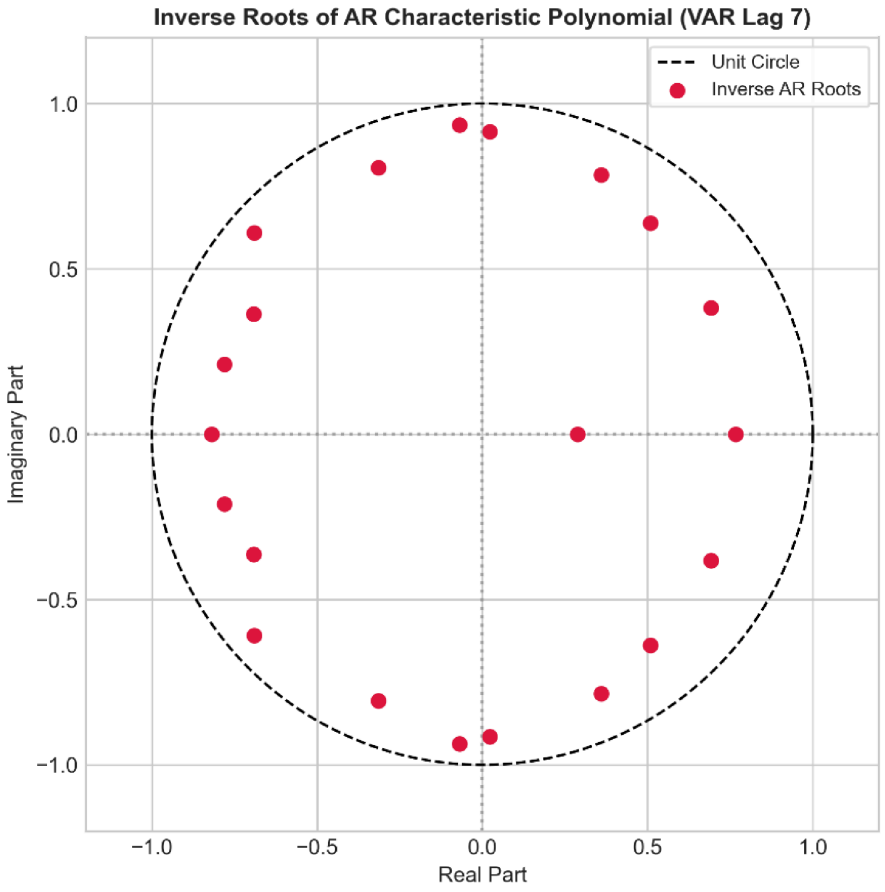
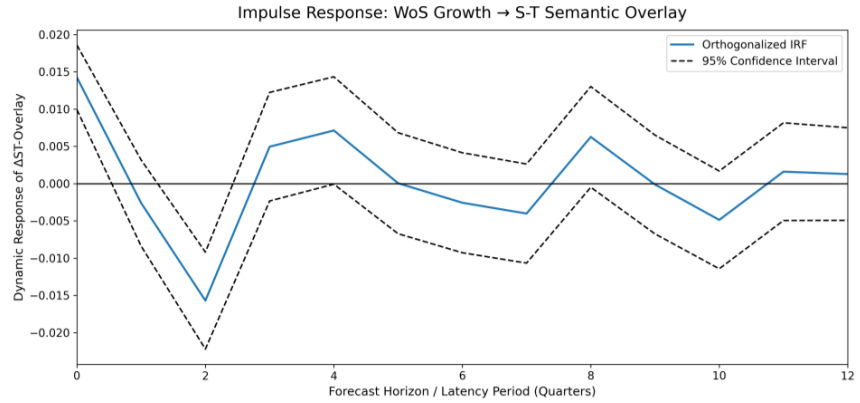
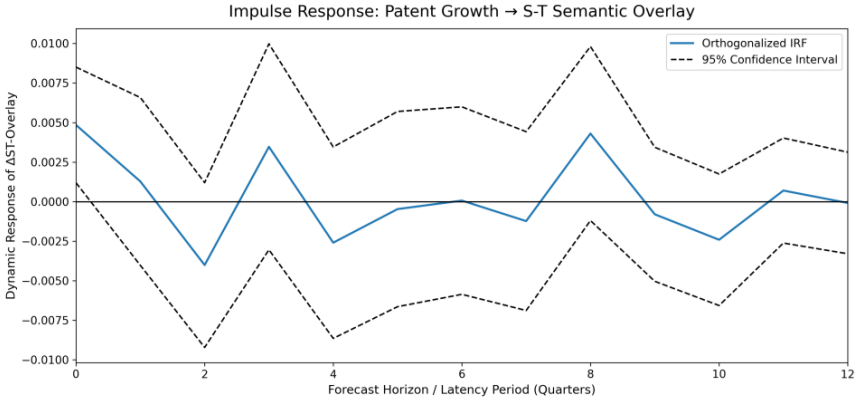


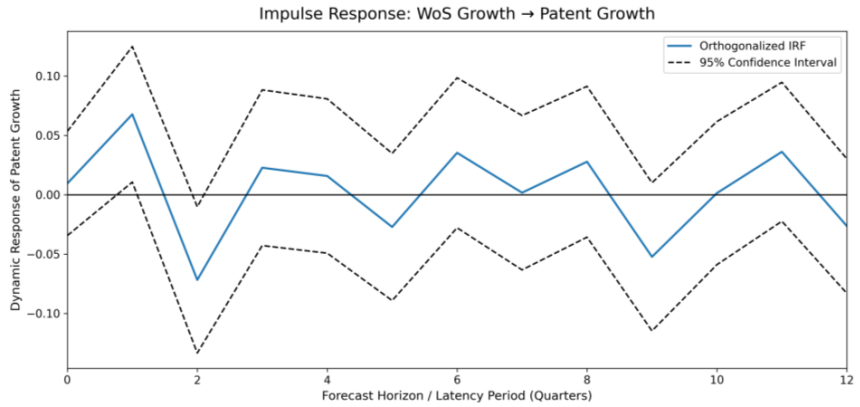
Figure 2. Inverse roots of the AR characteristic polynomial (VAR Lag 7).



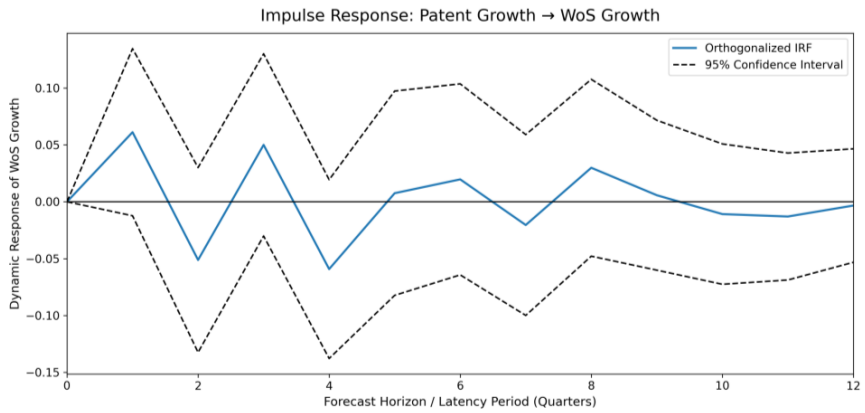
(a)



(b)



(c)



(d)

Figure 3. Orthogonalized impulse response functions for the four dynamic relationships in the VAR(7) system: (**a**) the response of semantic overlay to a one-standard-deviation shock in scientific publication

growth; (**b**) the response of semantic overlay to a one-standard-deviation shock in patent output growth; (**c**) the response of patent growth to a one-standard-deviation shock in scientific publication growth; (**d**) the response of scientific publication growth to a one-standard-deviation shock in patent growth. **Note:** The blue solid lines represent the orthogonalized impulse responses, and the black dashed lines indicate the 95% confidence intervals. The zero baseline indicates no response.

4.3. Normalized Distributions of Domain-Specific High-Frequency Keyword and Forecast Trajectories for Underground Safety Innovation (2026–2031)

To further explore the thematic characteristics between the academic literature and industrial patent corpus, a two-step keyword analysis was carried out. In the first step, the raw occurrence frequency was adopted to identify the high-frequency terms within the two datasets. After removing generic academic terms, patent boilerplate terms, and various domain descriptors, Figure 4 visualizes the distribution of the highest-frequency keywords across academic and patent corpora over the 22-year period. The scientific knowledge domain is dominated by tokens related to geomechanics and numerical modeling. High-frequency keywords include stress (4603), deformation (4365), failure (4001), surrounding (3681; semantically referring to surrounding rock), pressure (3663), damage (2771), soil (2729), underground (2512), geological (2416), monitoring (2302), stability (2298), and face (2074). In contrast, the Google Patents corpus is rich in construction-execution and apparatus-oriented terminology. High-frequency terms include steel (19,223), pipe (15,094), section (14,294), arch (13,286), wall (12,625), hole (11,715), layer (9969), grouting (9857), anchor (9713), bottom (9212), rod (8842), and concrete (8326). However, the raw occurrence counts of high-frequency keywords should not be used for quantitative cross-domain keyword comparison because the two corpora differ in size.

In the second step, the normalized document frequency was obtained for the 15 selected terms within each corpus, as shown in Tables 4 and 5. The normalized frequency provides a more comparable measure for each selected term by eliminating the impact of the document count. Moreover, this approach eliminates the bias caused by the documents in which the selected terms appear repeatedly. The normalized results show that rock, stress, deformation, and surroundings are prevalent in the scientific corpus, indicating that scholars focus on geomechanical behavior, deformation processes, and structural stability. For the patent-based technological dataset, wall, section, rock, steel, and structure- and construction-related terms exhibit high document-level prevalence, suggesting that patent applicants lay great emphasis on technological components and technical solutions.

It should be noted that rock, water, and lining appear in both term sets. The three shared terms indicate that the conditions of the surrounding rock, hazards related to groundwater, and structural safety are major concerns across the two datasets. Nevertheless, the lexical overlap should not be interpreted as direct evidence of knowledge transfer.

To connect the historical record with proactive future risk management, the Bayesian non-linear Prophet model was specified to forecast dual-domain growth boundaries through 2031 [74], as shown in Figure 5. This model incorporates historical change-point variation, quarterly seasonality at multiple scales, and saturation constraints on growth across the 2004–2026 baseline period and projects a five-year path of global academic publications and industrial patents related to underground safety.

As shown in Figure 5, the predicted curves exhibit a robust and sustained upward trend across the two domains, and the two key quantitative trajectories are described as follows:

(1) The Academic Research Momentum: The quarterly volume of scientific publications is predicted to exceed 500 by 2031. This upward trend is driven by the global expansion of ultra-deep underground infrastructure, which leads to novel research challenges related to high geostress, severe rockburst hazards, and thermal threats.

(2) **The Industrial Patenting Expansion:** The projected volume of industrial patents exhibits a markedly steeper upward trajectory. This expansion reflects the ongoing transition to smart construction paradigms, a shift that drives an increase in patent applications. In these emerging systems, conventional tunneling equipment is being enhanced with Internet of Things (IoT) sensors, real-time geological prediction tools, and automated control systems.

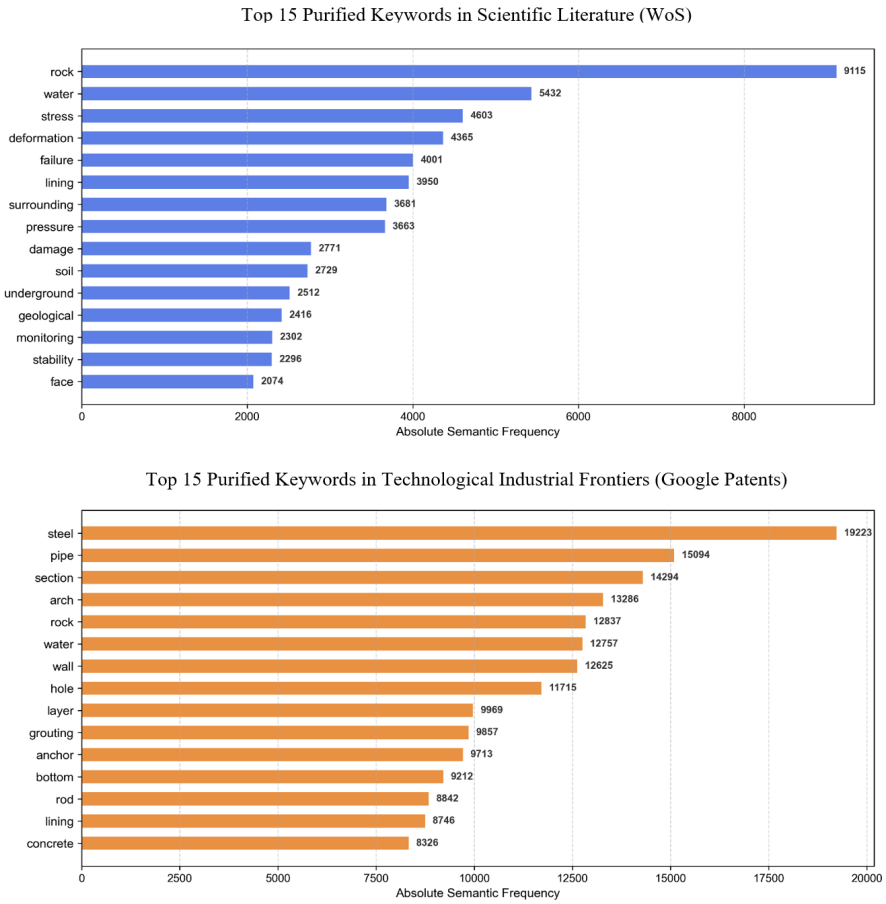


Figure 4. Domain-Specific High-Frequency Terms after Generic-Term Filtering.

Table 4. Normalized document-frequency distributions of selected high-frequency terms across the WoS corpus.

Rank	Keyword	Documents	Normalized Document Frequency
1	rock	2350	36.70%
2	stress	1638	25.58%
3	deformation	1591	24.84%
4	surrounding	1574	24.58%
5	water	1468	22.92%
6	underground	1408	21.99%
7	failure	1392	21.74%
8	pressure	1380	21.55%
9	stability	1298	20.27%
10	geological	1294	20.21%
11	lining	1228	19.18%
12	monitoring	1225	19.13%
13	damage	1132	17.68%
14	soil	1069	16.69%
15	face	831	12.98%

Table 5. Normalized document-frequency distributions of selected high-frequency terms across the Google Patent corpus.

Rank	Keyword	Documents	Normalized Document Frequency
1	wall	5236	25.01%
2	section	5185	24.77%
3	rock	4890	23.36%
4	steel	4825	23.05%
5	bottom	4573	21.85%
6	hole	4459	21.30%
7	pipe	4276	20.43%
8	water	4126	19.71%
9	arch	3619	17.29%
10	concrete	3493	16.69%
11	layer	3295	15.74%
12	rod	3001	14.34%
13	lining	2963	14.16%
14	grouting	2957	14.13%
15	anchor	2832	13.53%

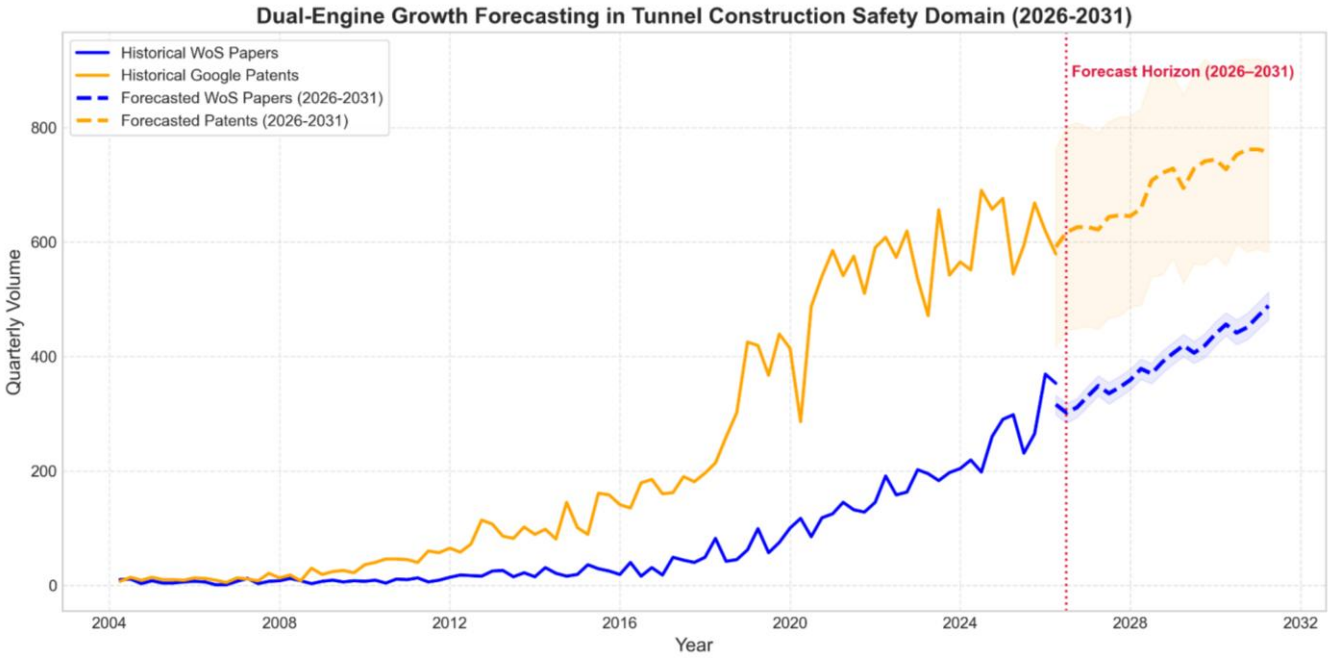


Figure 5. Dual-Domain Growth Trajectory for WoS Publications and Patent Applications Using the Prophet Model (2026–2031). (Note: Dashed lines denote point estimates for WoS publications and patent applications, with the semi-transparent shaded bands representing their respective 95% posterior predictive uncertainty intervals).

Notably, the 95% posterior predictive intervals, shown as semi-transparent shaded bands in Figure 5, remain structurally stable over the five years. Statistically, this constrained variance suggests that the co-evolutionary parameters linking academic publications with industrial patent applications have achieved structural stability.

5. Discussion and Management Implications

This empirical study reveals the co-evolutionary trajectories and dynamic temporal associations among scientific research, industrial technological innovation, and science-technology semantic overlay in tunnel construction safety. These findings, along with the

detection of normalized high-frequency terms, provide insights into how scientific and technological knowledge domains evolve within tunnel construction safety. The identified cross-domain semantic patterns and temporal dynamics offer valuable guidance to researchers, engineering organizations, and regulatory stakeholders involved in complex underground infrastructure projects.

5.1. Understanding Dynamic Interactions Between Scientific and Technological Knowledge Evolution

The long-term evolution of the science-technology semantic overlay reveals the relationship between scientific research and industrial innovation in underground construction safety. The evolutionary trajectory of the cross-domain semantic similarity demonstrates a descriptive three-stage. First, the semantic overlay curve features high-amplitude fluctuations, and its values remain consistently low. Then, the curve exhibits an upward trend. Finally, the values of semantic alignment remain consistently high, ranging from 0.91 to 0.93. However, the high cosine similarity value does not necessarily indicate knowledge transfer or technology adoption. Part of the observed similarity results from the shared disciplinary concepts and terminology because both corpora originate from the same domain. The temporal permutation benchmark also confirms this.

The VAR-based impulse-response analysis indicates that the relationships among scientific activity, technological innovation activity, and science-technology semantic overlay exhibit dynamic adjustment patterns rather than a persistent unidirectional pathway. The response of semantic alignment to scientific activity features a short-term adjustment pattern, indicating that scientific findings may initially expand conceptual boundaries. Gradually, the semantic alignment adjusts through the coordination between scientific and technological knowledge structures. Similarly, technological activity shocks also lead to temporal changes in semantic alignment, but the changes reflect a gradual adjustment process.

The response of technological innovation to scientific activity shocks shows that the scientific advances lead to subsequent adjustments in technological innovation, but a sustained directional response is absent. This pattern is consistent with the characteristics of engineering innovation, where scientific findings require adaptation and integration before they influence practical solutions. Meanwhile, the response of scientific activity to technological activity shocks also indicates that industrial innovation may influence scientific development. New construction equipment and operational protocols can give rise to new research questions. This bidirectional interaction challenges the assumption that industrial innovation is merely the recipient of scientific discoveries [75].

Overall, the four impulse-response analyses show the characteristics of tunnel construction safety innovation. Scientific findings and technological innovation influence each other, but this mutual influence exhibits dynamic adjustment behavior.

5.2. Understanding Differences and Complementarities Between Scientific and Technological Knowledge Domains

The asymmetry in terms of high-frequency keywords across the dual-domain further clarifies the divergence of cognitive focus between the scientific and patent-based technological corpora. The scientific corpus provides diagnostic and explanatory insights into the stress–strain field and failure criteria through laboratory-scale models and numerical simulations, with particular attention to the deformation of the structure and surrounding rock under special conditions [5,7]. In contrast, the patent applicants place great emphasis on construction equipment and supporting measures (such as steel arches, pipe umbrellas, reinforced lining, and facilities related to grouting) and the associated procedures. Crucially, three keywords, which are rock, water, and lining, frequently appear in the corpora

across the two domains, indicating that the surrounding rock stability, underground water control, and structural safety are major concerns for both corpora circles in the field of tunnel construction safety [76].

To bridge the gap between these two complementary domains and eliminate cross-disciplinary barriers, it is recommended to incorporate the following two integrative mechanisms into the tunnel construction safety management procedure:

(1) Automated Scientific and Technological Knowledge Tracking: Safety management organizations should regard scientific research and industrial innovation as interconnected information sources rather than separate knowledge systems. As a result, timely tracking of the latest academic findings and technological innovation is highly valuable for tunnel construction safety management, and an AI-driven natural language processing (NLP) pipeline is recommended to monitor global academic research and technological innovation. Many scientific findings deserve attention from patent applicants, especially those concerning new high-strength materials, failure modes, and constitutive models, which may provide valuable insights for developing new safety equipment or modifying existing facilities [76]. Meanwhile, the new practical solutions, including new equipment or construction protocols, could introduce new research questions for researchers [77–79]. Consequently, this automated pipeline could promote the coordination between research institutions and engineering organizations.

(2) Academia–Industry Collaboration for the Development of New Equipment and Workflows: Many complex underground projects may encounter challenging geohazards, including high in situ stress and water-rich strata, and these conditions may pose great risks to construction safety. For these cases, it is advisable for the construction team to establish a collaborative team consisting of academic researchers, equipment manufacturers, and on-site engineers to tackle these problems. In this way, each party is capable of leveraging its own strengths, and theoretical analysis along with multi-scale models could then directly support the industrial innovation. At the same time, real-time field data, such as thrust, torque, penetration rate, and other indicators of the TBM, could also be used to calibrate the theoretical models. This closed-loop process supports a continuous co-development cycle in which scientific findings and engineering practices can mutually inform each other.

5.3. Policy Framework for Proactive Safety Management in Tunnel Construction

According to the Prophet forecast shown in Figure 5, the tunnel construction safety field is experiencing a sustained innovation surge, driven by the deployment of autonomous TBM control, real-time vision-based excavation monitoring, and other emerging intelligent technologies [70]. This trend is accompanied by a rapid increase in academic publications and patent applications.

This innovation surge presents both a challenge and an opportunity for the industry. Regulatory agencies and safety oversight bodies must transition from reactive, post-incident investigations to proactive, data-driven management to respond to this change:

(1) Timely Safety-Related Standard Updating: Traditionally, tunnel safety management standards lagged years behind validated scientific findings and industrial innovation. However, regulators can leverage the current high-level semantic convergence identified in this trajectory analysis to update the standards in a timely manner. For example, advanced industrial innovation, such as AI-driven geological forecasting, which is widely recognized as a critical enabler for tunnel construction, have attracted considerable attention from both academia and industry. Regulators could integrate these technologies into the national safety codes as compulsory baselines.

(2) Targeted Co-Innovation Subsidies: National research funding bodies should prioritize multi-party research teams that not only focus on scientific problems but also

place great emphasis on industrial innovations. This kind of team is capable of translating on-site problems into theoretical frameworks and multi-scale models, which can be examined by robust research methods. The research findings, in turn, could provide valuable insights for industrial innovation that directly addresses the aforementioned issues. This closed-loop paradigm stems from field problems, is examined in the laboratory, and ultimately feeds back into engineering practice. This cycle not only maximizes the impact of research funding but also advances tunnel construction safety management worldwide.

5.4. Limitations and Future Research Directions

Although this study is based on a cross-domain corpus spanning more than two decades and employs several robust methods to obtain reliable results, certain boundary conditions nevertheless remain, providing pathways for subsequent investigation:

Although the data was obtained from global databases, Chinese publications and patents account for approximately 82% and 96% of the scientific and technological corpora, while the representation of different construction methods is also uneven. Therefore, the findings should not be interpreted as a geographically balanced global or method-specific comparison. Future studies could retrieve country-and method-specific datasets to explore how institutional heterogeneity and construction methods influence the science-technology co-evolution.

This research is carried out at the macro level and relies on quarterly industrial aggregates, which fail to capture the dynamics at the firm level or technology level. To fill this research gap, micro-level data from various types of companies and documents associated with emerging technology could be incorporated into the subsequent research.

Although the VAR modeling and impulse-response analysis reveal the temporal dependencies and dynamic associations, the casual relationship between the scientific activity and patent-based technological activity has not been established. Future studies should incorporate the firm-level and project-level data to explore the casual relationship between scientific findings and industrial technology.

6. Conclusions

This study proposes a comprehensive framework to explore the temporal interactions and co-evolutionary patterns between global academic findings and industrial technological innovations in tunnel construction safety. To ensure the reliability of research conclusions, the empirical dataset comprises a cross-domain corpus of 6404 WoS publications and 20,932 Google Patents covering 22 years (2004–2026). Methodologically, more robust models and quantitative methods, such as the bge-large-en-v1.5 model and VAR, are employed to process the textual data instead of simple keyword-based approaches. The econometric and semantic analysis yield several data-driven empirical conclusions, which can provide a reference for proactive safety governance in complex underground projects:

(1) The temporal interaction, featuring the cross-domain semantic similarity between scientific findings and industrial innovations, reveals a descriptive three-stage trajectory over the 22-year period. The cross-domain semantic similarity experienced pronounced fluctuations prior to 2008 and then exhibited a steadily growing trend until 2017. Finally, the semantic alignment crystallized into a sustained plateau, stabilizing within the range of 0.91 to 0.93, indicating the increasing semantic convergence between scientific and technological knowledge domains in the sector of underground engineering safety. However, the temporal permutation benchmark indicates that a portion of the similarity results from the shared domain of the two corpora.

(2) The VAR analysis shows that the relationships among the scientific activity, patent-based technological activity, and their semantic alignment feature temporal fluctuations

and dynamic adjustment. The absence of persistent directional responses indicates that the science-technology co-evolution in tunnel construction safety is more complex than a simple science-push or technology-pull pattern. Instead, the two knowledge domains exhibit asymmetric and temporally varying interactions, indicating that scientific advancement, engineering requirements, and technological solutions continuously interact.

(3) The analysis of normalized document-frequency keywords uncovers a divergence in intellectual priorities between scientific publications and patent-based technological corpora. Academic scholars focus on risk detection and mechanical mechanisms, with terms such as stress, deformation, failure, and monitoring ranking as the most frequent. However, terms related to support and reinforcement measures, such as steel, piping, arching, and grouting, are the most frequent in the patent corpus, indicating that patent applicants concentrate on the physical implementation. The shared keywords reveal that both corpora pay great attention to surrounding rock stability, underground water control, and structural safety. Bayesian Prophet simulations exhibit a stable, predictable upward trend in the volume of scientific publications and patent applications.

These findings should nevertheless be interpreted in light of the limitations discussed in Section 5, particularly those concerning geographic and construction methodological coverage, macro-level aggregation, and the interpretation of temporal relationships.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/su182110808/s1, Table S1: Post-retrieval coverage validation of the scientific and industrial corpora; Table S2: Temporal permutation benchmark for the semantic overlay.

Author Contributions: Conceptualization, G.Y.; methodology, G.Y.; software, G.Y.; validation, G.Y.; formal analysis, G.Y.; investigation, G.Y.; data curation, G.Y.; writing—original draft preparation, G.Y.; writing—review and editing, G.Y. and L.Y.; supervision, L.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data generated or analyzed, models, or code used during the study are available from the corresponding author by request.

Acknowledgments: The authors gratefully appreciate all the support.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pandey, B.; Brelsford, C.; Seto, K.C. Rising infrastructure inequalities accompany urbanization and economic development. *Nat. Commun.* **2025**, *16*, 1193. [CrossRef] [PubMed]
2. Shakeri, J.; Ghorbani, E.; Taheri, A. A review of deep mining challenges, hazard mitigation, design approaches, and support methods. *Undergr. Space* **2026**, *28*, 86–136. [CrossRef]
3. Shayan, F.; Uromeihy, A.; Hassanpour, J. Classification framework for engineering geological challenges in fault zones during TBM tunneling: Insights from mechanized projects in Iran. *Eng. Geol.* **2025**, *358*, 108382. [CrossRef]
4. Wang, M.B.; Kuang, D.M.; Chen, L.; Song, Z.B.; Peng, C. Refined Inversion Analysis of 3D Ground Stress Field Based on Random Field. *Geotech. Geol. Eng.* **2025**, *43*, 105. [CrossRef]
5. Huang, X.; Wang, T.; Xu, Z.H.; Zhu, Z.G.; Guo, J.Q.; Lin, P.; Tian, Y.C. Experimental Study on the Mechanism of Intermittent Water and Mud Inrush Caused by Dredging and Rainfall Considering Sedimentary Characteristics in Karst Tunnels. *KSCE J. Civ. Eng.* **2023**, *27*, 2016–2029. [CrossRef]
6. Gao, Y.X.; Li, J.Y.; Hao, J.; Yang, X. Application of constitutive model in simulation analysis of tunnel construction. *J. Appl. Sci. Eng.* **2021**, *24*, 791–797.
7. Xiao, B.; Li, X.; Huang, L. Numerical Simulation of Tunnel Face Support Pressure in Layered Soft Ground. *Appl. Sci.* **2024**, *14*, 2943. [CrossRef]

8. Peng, Y.; Zhang, S.; Su, L.; Yao, Z. Risk Prediction of Water Inrush in Diversion Tunnel Crossing Water-Rich Fault Based on NRBO-XGBoost Algorithm. *Appl. Sci.* **2026**, *16*, 3831. [CrossRef]
9. Chen, W.Z.; Shen, K.; Tan, X.J.; Gao, H.; Zheng, P.C.; Lan, Z.J.; Zhang, C.X.; Liu, Y.B. Three-dimensional yielding support physical modelling test system for underground engineering large deformation. *Tunn. Undergr. Space Technol.* **2026**, *168*, 107230. [CrossRef]
10. China Railway Tunneling Equipment Co., Ltd. Cutting Device with Function of Changing Diameter in Small Range for Soft Rock Shield Machine. WO2012024938A1, 1 March 2012.
11. Shandong University. Method and System for Real-Time Prediction of Jamming in TBM Tunneling. US-2023167739-A1, 1 June 2023.
12. Nonaka, I. A dynamic theory of organizational knowledge creation. *Organ. Sci.* **1994**, *5*, 14–37. [CrossRef]
13. Geels, F.W. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Res. Policy* **2002**, *31*, 1257–1274. [CrossRef]
14. Agarwal, R.; Moeen, M.; Shah, S.K. Athena's Birth: Triggers, Actors, and Actions Preceding Industry Inception. *Strateg. Entrep. J.* **2017**, *11*, 287–305. [CrossRef]
15. Yu, Y.; Wu, J.; Shui, D.C.; Dong, X.L. Asynchrony as information: Predicting technology transfer opportunities through science-technology knowledge lag. *Inf. Process. Manag.* **2026**, *63*, 104906. [CrossRef]
16. Liu, Z.; Dong, T.; Lv, Y.; Qi, X.; Shui, Z. Internal moisture evolution and humidity-dependent shrinkage behavior of metakaolin-modified ultra-high-performance concrete. *Constr. Build. Mater.* **2026**, *541*, 147873. [CrossRef]
17. Luo, J.; Zhu, S.; Zhai, W. An advanced train-slab track spatially coupled dynamics model: Theoretical methodologies and numerical applications. *J. Sound Vib.* **2021**, *501*, 116059. [CrossRef]
18. Hötte, K.; Pichler, A.; Lafond, F. The rise of science in low-carbon energy technologies. *Renew. Sustain. Energy Rev.* **2021**, *139*, 110654. [CrossRef]
19. Narin, F.; Hamilton, K.S.; Olivastro, D. The increasing linkage between U.S. technology and public science. *Res. Policy* **1997**, *26*, 317–330. [CrossRef]
20. Leydesdorff, L.; Rafols, I. Local emergence and global diffusion of research technologies An exploration of patterns of network formation. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 846–860. [CrossRef]
21. Yan, H.; Li, H.; Xie, H.; Xiao, M.; Pei, J.; Liu, H.; He, J. A review on rockburst prediction and prevention to shape an ontology-based framework for better decision-making for underground excavations. *Deep Undergr. Sci. Eng.* **2025**, *5*, 331–349. [CrossRef]
22. Mansfield, E. Technical Change and the Rate of Imitation. *Econometrica* **1961**, *29*, 741–766. [CrossRef]
23. Dasgupta, P.; David, P.A. Toward a New Economics of Science. *Res. Policy* **1994**, *23*, 487–521. [CrossRef]
24. Bush, V. *Science, the Endless Frontier: A Report to the President on a Program for Postwar Scientific Research*; U.S. Government Printing Office: Washington, DC, USA, 1945.
25. Kemp, R.; Schot, J.; Hoogma, R. Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technol. Anal. Strateg. Manag.* **1998**, *10*, 175–198. [CrossRef]
26. Nelson, R.R.; Winter, S.G. *An Evolutionary Theory of Economic Change*; Belknap Press of Harvard University Press: Cambridge, MA, USA, 1982.
27. Dosi, G. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Res. Policy* **1982**, *11*, 147–162.
28. Stokes, D.E. *Pasteur's Quadrant: Basic Science and Technological Innovation*; Brookings Institution Press: Washington, DC, USA, 1997.
29. Gittelman, M.; Kogut, B. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Manag. Sci.* **2003**, *49*, 366–382. [CrossRef]
30. van Raan, A.F.J. Patent Citations Analysis and Its Value in Research Evaluation: A Review and a New Approach to Map Technology-Relevant Research. *J. Data Inf. Sci.* **2017**, *2*, 13–50. [CrossRef]
31. Tijssen, R.J. Global and domestic utilization of industrial relevant science: Patent citation analysis of science–technology interactions and knowledge flows. *Res. Policy* **2001**, *30*, 35–54. [CrossRef]
32. Meyer, M. Does science push technology Patents citing scientific literature. *Res. Policy* **2000**, *29*, 409–434. [CrossRef]
33. Roach, M.; Cohen, W.M. Lens or prism Patent citations as a measure of knowledge flows from public research. *Manag. Sci.* **2013**, *59*, 504–525. [CrossRef] [PubMed]
34. Bikard, M. Made in academia The effect of institutional origin on inventors' attention to science. *Organ. Sci.* **2018**, *29*, 818–836. [CrossRef]
35. Lensjø, M. Vocational Teachers' Craft Knowledge and Working-life Experiences in Building and Construction: A Narrative Study of Embodied and Tacit Learning. *Vocat. Learn.* **2024**, *17*, 333–350. [CrossRef]
36. Ding, Y.; Chowdhury, G.G.; Foo, S. Bibliometric cartography of information retrieval research by using co-word analysis. *Inf. Process. Manag.* **2001**, *37*, 817–842. [CrossRef]
37. Kay, L.; Newman, N.; Youtie, J.; Porter, A.L.; Rafols, I. Patent overlay mapping: Visualizing technological distance. *J. Assoc. Inf. Sci. Technol.* **2014**, *65*, 2432–2443. [CrossRef]

38. Poege, F.; Harhoff, D.; Gaessler, F.; Baruffaldi, S. Science Quality and the Value of Inventions. *Sci. Adv.* **2019**, *5*, eaay7323. [CrossRef] [PubMed]
39. Wu, G.; Tan, Y.; Zeng, J.; Zheng, J.; Zhang, R.; Liu, Y.; Yue, S. Laboratory tests of an eccentrically loaded strip footing above single underlying void. *J. Build. Eng.* **2025**, *111*, 113211. [CrossRef]
40. Abramo, G.; D'Angelo, C.A. Inter-and intra-domain knowledge flows Examining their relationship with impact at the field level over time. *J. Informetr.* **2025**, *19*, 101614. [CrossRef]
41. Zheng, Z.J.; Ma, Y.X.; Ba, Z.C.; Pei, L. Tree knowledge structure for better insight Capturing biomedical science-technology knowledge linkage with MeSH. *J. Informetr.* **2024**, *18*, 101568. [CrossRef]
42. Xie, H.; Lu, J.; Li, C.; Li, M.; Gao, M. Experimental study on the mechanical and failure behaviors of deep rock subjected to true triaxial stress: A review. *Int. J. Min. Sci. Technol.* **2022**, *32*, 915–950. [CrossRef]
43. Lin, C.; Zhang, M.; Zhou, Z.; Li, L.; Shi, S.; Chen, Y.; Dai, W. A new quantitative method for risk assessment of water inrush in karst tunnels based on variable weight function and improved cloud model. *Tunn. Undergr. Space Technol.* **2020**, *95*, 103136. [CrossRef]
44. Wu, X.; Sun, P. Dynamic Analysis and Temporal Governance of Safety Risks: Evidence from Underground Construction Accident Reports. *Sustainability* **2024**, *16*, 8531. [CrossRef]
45. Rössler, K.; Franzén, G.; Galler, R.; Grunicke, U.; Höfer-Öllinger, G.; Jäger, J.; Marulanda Escobar, A. Understanding tunnel failures during excavation Causes, limitations, and recommendations. *Geomech. Tunn.* **2025**, *18*, 599–607. [CrossRef]
46. Xiao, S.; Liu, Z.; Zhang, P.; Muennighoff, N.; Lian, D.; Nie, J.-Y. C-Pack Packed Resources For General Chinese Embeddings. In *Proceedings of the 47th International ACM SIGIR Conference, Washington, DC, USA, 2024*; ACM: New York, NY, USA, 2024; pp. 641–649.
47. Banerjee, A.; Dhillon, I.S.; Ghosh, J.; Sra, S. Clustering on the unit hypersphere using von Mises-Fisher distributions. *J. Mach. Learn. Res.* **2005**, *6*, 1345–1382.
48. Hamilton, W.L.; Leskovec, J.; Jurafsky, D. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2016*; ACL: Stroudsburg, PA, USA, 2016; Volume 1, pp. 1489–1501.
49. Arora, S.; Liang, Y.; Ma, T. A Simple but Tough-to-Beat Baseline for Sentence Embeddings. In *Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)*; OpenReview.net: Toulon, France, 2017.
50. Zhelezniak, V.; Savkov, A.; Shen, A.; Hammerla, N.Y. Correlation Coefficients and Semantic Textual Similarity. In *Proceedings of the 2019 Conference of the North American Chapter of th e Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019)*; ACL: Minneapolis, MN, USA, 2019; Volume 1, pp. 951–962.
51. Rafols, I.; Porter, A.L.; Leydesdorff, L. Science Overlay Maps: A New Tool for Research Policy and Library Management. *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 1871–1887. [CrossRef]
52. Lawless, P.A. Bounded Logarithm Scale for Particle Count Data. *Aerosol Sci. Technol.* **1996**, *24*, 217–219. [CrossRef]
53. Dickey, D.A.; Fuller, W.A. Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root. *Econometrica* **1981**, *49*, 1057–1072. [CrossRef]
54. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
55. Granger, C.W.J. Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica* **1969**, *37*, 424–438. [CrossRef]
56. Sims, C.A. Macroeconomics and Reality. *Econometrica* **1980**, *48*, 1–48. [CrossRef] [PubMed]
57. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
58. Akaike, H. Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.* **1969**, *21*, 243–247. [CrossRef]
59. Hannan, E.J.; Quinn, B.G. The determination of the order of an autoregression. *J. R. Stat. Soc. Ser. B* **1979**, *41*, 190–195. [CrossRef]
60. Lütkepohl, H. *New Introduction to Multiple Time Series Analysis*; Springer Science & Business Media: Dordrecht, The Netherlands, 2005.
61. Bilgin, N.; Dincer, T.; Copur, H.; Erdogan, M. Some Geological and Geotechnical Factors Affecting the Performance of a Roadheader in an Inclined Tunnel. *Tunn. Undergr. Space Technol.* **2004**, *19*, 629–636. [CrossRef]
62. Shanghai Urban Construction Design Co., Ltd. Reinforced Concrete Structure Embedded Plastic Valve Plug. CN2606180Y, 10 March 2004.
63. Bäppler, K. New developments in TBM tunnelling for changing grounds. *Tunn. Undergr. Space Technol.* **2016**, *57*, 18–26. [CrossRef]
64. Bobylev, N. Underground space in the Alexanderplatz area, Berlin: Research into the quantification of urban underground space use. *Tunn. Undergr. Space Technol.* **2010**, *25*, 495–507. [CrossRef]
65. Yan, P.; Lu, W.; Chen, M.; Shan, Z.; Chen, X.; Zhou, Y. Energy release process of surrounding rocks of deep tunnels with two excavation methods. *J. Rock Mech. Geotech. Eng.* **2012**, *4*, 160–167. [CrossRef]
66. Gong, Q.; Yin, L.; Ma, H.; Zhao, J. TBM tunnelling under adverse geological conditions: An overview. *Tunn. Undergr. Space Technol.* **2016**, *57*, 4–17. [CrossRef]

67. Habib, P. Method for Constructing an Underground Tunnel or Hole to Create an Impervious Plug for the Storage of Hazardous, Particularly Radioactive, Waste. US2011116868A1, 19 May 2011.
68. Pan, Q.; Zhang, Y.; Tang, C.; Luo, W.; Yang, Z.; Lei, M. Digital twin modelling of shield tunneling undercrossing existing tunnels: A case study. *Tunn. Undergr. Space Technol.* **2026**, *177*, 107890. [CrossRef]
69. Chen, X.; Liu, L.; Ren, S.; Cheng, P.; Liu, M.; Liu, Y.; Chen, E.J. Physical modelling of face stability and ground settlement in EPB shield tunnelling incorporating dynamic cutterhead–soil interaction. *Tunn. Undergr. Space Technol.* **2026**, *170*, 107332. [CrossRef]
70. Zhou, J.; Qi, H.; Peng, K.; Zhang, Y.; Khandelwal, M. Comprehensive review and future perspectives on prediction and mitigation of tunnel-induced ground settlement: A bibliometric analysis and methodological overview (2002–2022). *Tunn. Undergr. Space Technol.* **2024**, *154*, 106081. [CrossRef]
71. Weber, C.E. F-tests for lag length selection in augmented Dickey-Fuller regressions some Monte Carlo evidence. *Appl. Econ. Lett.* **2001**, *8*, 455–458. [CrossRef]
72. Ivanov, V.; Kilian, L. A practitioner's guide to lag order selection for VAR impulse response analysis. *Stud. Nonlinear Dyn. Econom.* **2005**, *9*. [CrossRef]
73. Hugo, I.P.; Rafael, D. Public investment, infrastructure and private investment in Brazil: Is there a crowding-in effect? *EconomiA* **2024**, *25*, 289–308. [CrossRef]
74. Taylor, S.J.; Letham, B. Forecasting at Scale. *Am. Stat.* **2018**, *72*, 37–45. [CrossRef]
75. Morris, Z.S.; Wooding, S.; Grant, J. The answer is 17 years, what is the question: Understanding time lags in translational research. *J. R. Soc. Med.* **2011**, *104*, 510–520. [CrossRef] [PubMed]
76. Zhou, X.; Liu, W.; Liang, J.; Yu, Q. Research on Time-Dependent Buoyancy Characteristics of Shield Tail Grouting Slurry. *Sustainability* **2026**, *18*, 6489. [CrossRef]
77. Guo, Z.; Yang, B.; Xie, J.; Wang, T.; Yang, J.; Niu, B. A multi-task integrated question-answering framework for highly generalizable fault and RUL assessments based on LLMs with a post-training and fine-tuning strategy. *Reliab. Eng. Syst. Saf.* **2026**, *278*, 113402. [CrossRef]
78. Tang, B.; Yang, Z.; Yao, D.; Yang, B.; Chen, L.; Bi, R.; Zhao, Y.; Xie, J.; Guo, Z. LoFT-CLIP: Few-shot anomaly detection for railway fasteners based on large vision-language models. *Pattern Recognit.* **2026**, *180*, 114099. [CrossRef]
79. Wu, G.; Tan, Y.; Xu, Z.; Yuan, Z.; Zhang, J.; He, Y.; Wang, C. XAI-Aided Method for Accurately Assessing the Stability of Unlined Double Tunnels in Rock Masses. *Transp. Geotech.* **2026**, *61*, 102052. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.