ICAS X 2026
8–10 July 2026 • Kraków, Poland

Smart Farming Innovations using Patent Data

Hannah Gerits[1]
UNU-MERIT Maastricht University, Maastricht, The Netherlands - hannah.gerits@student.maastrichtuniversity.nl

Tommaso Ciarli
UNU-MERIT Maastricht University, Maastricht, The Netherlands - ciarli@merit.unu.edu

Bruno Martorano
UNU-MERIT Maastricht University, Maastricht, The Netherlands - bruno.martorano@maastrichtuniversity.nl

Abstract

Digital technologies are increasingly embedded within agricultural production systems, yet robust measurement frameworks for tracing their structural integration remain limited. This paper proposes a network-based approach to identify and characterize smart farming innovation trajectories using patent citation data from the European Patent Office (EPO) over the period 1980–2025. We construct a full citation network and extract the Network of Main Paths (NMP) using the Search Path Node Pair (SPNP) metric to identify structurally important knowledge flows. Rather than defining smart farming at the patent level, we classify trajectories as smart farming when agricultural and digital technologies co-occur along the same citation path, capturing temporal knowledge integration within agri-food innovation systems. Applying this framework, we identify 26,885 smart farming trajectories, representing 3.7% of domain-relevant paths. These trajectories are structurally longer, exhibit higher centrality weights, and display distinct sequencing patterns compared to single-domain trajectories. Digital-first integration dominates, with a median lag of 14 years before agricultural application. Climate-related technologies identified through Cooperative Patent Classification (CPC) Y02 and Y04S codes are increasingly embedded within these hybrid trajectories. The proposed framework offers a first set of trajectory-based indicators for monitoring technological convergence, digital transformation, and climate responsiveness within agri-food innovation systems, providing an empirical foundation for further methodological refinement and policy-relevant analysis.

Keywords: Smart Farming; Patent Citation Networks; Technological Trajectories; Digital Agriculture; Knowledge Integration; Climate Technology; Main Path Analysis.

[1]The authors confirm that the text and materials in this manuscript are free from any copyright violations.

1 Introduction

Agricultural production systems are undergoing profound transformation driven by digitalization, environmental pressures, and the need for greater resilience. Technologies such as precision sensing, robotics, artificial intelligence, and networked communication systems are increasingly embedded within farming operations and agri-food value chains, while climate change and resource constraints are reshaping technological priorities. Understanding how digital and agricultural technologies converge and evolve over time is therefore essential for monitoring structural change in agri-food systems.

Patent data offer harmonized, longitudinal, and internationally comparable records of technological development. Their structured classification systems and citation relationships allow researchers to trace knowledge flows across technological domains [1, 2]. However, conventional patent-based indicators - such as counts by classification or citation-based centrality measures - primarily capture static technological attributes and do not reveal how knowledge from distinct domains converges dynamically over time.

Smart farming exemplifies this measurement challenge. It does not correspond to a single patent class or clearly delineated field; rather, it emerges at the intersection of agricultural and digital technologies [3]. Identifying such hybrid domains requires moving beyond patent-level classification toward trajectory-based analysis of knowledge flows: instead of asking whether an individual invention is "digital" or "agricultural", the relevant question becomes whether digital and agricultural knowledge combine along the same technological pathways.

This paper proposes a network-based framework to identify and characterize smart farming innovation trajectories using complete patent citation networks from the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO) covering 1980 - 2025. Using the Network of Main Paths (NMP) methodology and the Search Path Node Pair (SPNP) metric [4, 5], we extract structurally important source-to-sink knowledge-flow paths and classify trajectories as smart farming when agricultural and digital technologies co-occur along the same citation sequence.

The paper advances data-driven monitoring of agri-food innovation in three ways. First, it introduces a reproducible trajectory-level approach for identifying hybrid innovation domains. Second, it provides structural indicators of technological convergence, sequencing, and knowledge integration along citation paths. Third, by incorporating climate-related technologies identified through CPC Y02 and Y04S codes, it offers a framework for tracking environmental responsiveness within digital agriculture. From a statistical measurement perspective, the proposed indicators can be computed annually from publicly available patent data, offering national statistical offices and international organizations a scalable complement to existing innovation surveys and research and development expenditure statistics for monitoring digital transformation in agri-food systems. Together, these contributions support more informed analysis of sustainability transitions in agri-food innovation systems.

2 Literature Review

2.1 Digital technologies in agriculture: measurement challenges

The integration of digital technologies into agricultural production has attracted growing policy and research attention. International organizations have increasingly emphasized the role of data-driven decision systems, precision sensing, and digital advisory services in improving productivity, sustainability, and resilience [6, 7]. Systematic reviews of the emerging literature identify a wide range of technologies — including the Internet of Things, cloud computing, robotics, and artificial intelligence - that are reshaping farm management cycles and agri-food value chains [3, 8]. Empirical research on digital agriculture adoption has relied primarily on farm-level surveys, case studies, and technology diffusion models, which provide important evidence on adoption determinants and outcomes but are limited in their ability to capture the systemic, cross-sectoral character of technological change [8]. Smart farming — broadly understood as the application of digital technologies to agricultural production and management within a cyber-physical systems framework [3] - does not correspond to a single technological field. It emerges at the intersection of developments in sensing, communication, computation, and agricultural science, making it difficult to delineate using conventional classification boundaries. Patent landscape analyses of the agri-food sector have documented the breadth of this innovation space, spanning a broad variety of technological sub-domains across agricultural technology and food technology, with substantial growth in automation and IoT-related areas [9]. However, such analyses primarily track innovation intensity through patent counts and classification distributions. More broadly, patent-based indicators—while offering important advantages in terms of availability, quantitative comparability, and disaggregation [1, 10] — are less equipped to measure how knowledge flows between distinct technological domains converge over time. Addressing this gap requires analytical approaches capable of tracing temporal knowledge integration across classification boundaries.

2.2 Patent citation networks and trajectory analysis

Patent citation networks offer a structured framework for studying cumulative technological development. By representing patents as nodes and citations as temporally directed edges, these networks capture how inventions build sequentially on prior knowledge [1], [2]. Connectivity-based traversal measures have been validated as recovering meaningful technological development patterns rather than arbitrary link structures [11]. Main path analysis, introduced by Hummon and Dereian [4], identifies the most significant knowledge-flow routes within citation networks using traversal metrics such as the Search Path Node Pair (SPNP) measure, which weights citation links by their global structural importance. Efficient algorithms for computing these metrics in large networks have been developed by Batagelj [12], while applications to specific technological domains — including fuel cells [13] and data communications [14] — have demonstrated that extracted paths correspond to recognisable development trajectories. The Network of Main Paths (NMP) framework proposed by Nomaler and Verspagen [5] extends this approach by retaining multiple co-existing optimal trajectories, making it particularly suited to domains characterised by parallel experimentation and recombination. Their application to green technologies [15] provides a direct conceptual foundation for the present study: rather than pre-defining a technological domain, they identify it through the co-occurrence of attributes along structurally important citation paths.

2.3 Climate-related patent classification

The Cooperative Patent Classification (CPC) Y02 and Y04S tagging scheme provides a standardised mechanism for identifying technologies relevant to climate change mitigation and adaptation [16]. Maintained and updated by the European Patent Office [17], this classification has been widely used in empirical research to study green innovation dynamics. Within the broader OECD framework for measuring environmental innovation using patent data, Y02-based indicators have proven particularly effective for tracking technology development, international co-invention, and technology diffusion in climate-relevant domains [10]. Integrating Y02 classifications within citation-based trajectory analysis enables a structural assessment of whether climate-related technologies are embedded within the dominant knowledge-flow pathways of agri-food innovation or remain peripheral to them.

3 Methodology

Smart farming is not a clearly delineated technological field; it emerges at the intersection of agricultural and digital technologies, making it difficult to define *a priori* which patents belong to this domain. Following [5], who studied green technologies as a hybrid domain, we conceptualize smart farming as innovations that are both "agricultural" and "digital". We employ the Network of Main Paths (NMP) framework [5], which identifies multiple co-existing technological trajectories rather than a single dominant path - a property particularly relevant for domains characterized by branching, parallel experimentation, and recombination.

Our empirical strategy proceeds in three stages: (i) construction of the patent citation network, (ii) identification of the Network of Main Paths using flow-based connectivity weights, and (iii) extraction of smart farming trajectories through domain-specific filtering.

To clarify the terminology used throughout: a *trajectory* is a unique maximum-SPNP source-to-sink path (deduplicated by exact node sequence across focal nodes); a *focal patent* is the center node used in the node-centered trajectory construction; and a *domain-relevant trajectory* is one passing through at least one agricultural or digital patent.

We construct a comprehensive citation network for the EPO by extracting first-filing patent applications over the period 1980 - 2025 from the PATSTAT Database along with their citation relationships [18]. Following [5], three types of citations are included: direct EPO-to-EPO citations, indirect citations mediated through other authorities with one degree of separation (i.e., an EPO patent citing a non-EPO patent that itself cites another EPO patent), and citations derived from DOCDB patent families representing technological continuity. The citation network forms a directed graph where nodes represent patent applications and edges represent citations oriented from cited to citing patents. To ensure the directed acyclic graph (DAG) structure required for path analysis, a citation is retained only if the filing date of the cited patent strictly precedes that of the citing patent; same-day filings are excluded.

To quantify the structural importance of citation links, we employ the Search Path Node Pair (SPNP) metric [4]. SPNP assigns to each citation link a weight proportional to the number of times that link appears on paths connecting source nodes (patents with no incoming citations) to sink nodes (patents with no outgoing citations). Because SPNP values compound multiplicatively along citation chains, we apply a logarithmic transformation to enable efficient identification of optimal trajectories using dynamic programming [5, 12].

The NMP is constructed by treating every patent as a potential focal point of technological development. For each node, we identify the highest-weighted path from any source node to the focal patent (backward trajectory) and the highest-weighted path from the focal patent to any sink node (forward trajectory). The concatenation of these segments forms the maximum-SPNP trajectory through that node. The NMP is then defined as the union of all edges belonging to at least one such trajectory. As multiple paths may achieve identical maximum weights, all tied optimal edges are retained.

Domain restrictions are deliberately not imposed during network construction or NMP identification. Instead, smart farming is identified *ex post* through patent-level classification, and trajectories are extracted from the pre-computed global NMP. This strategy preserves both upstream technological roots and downstream applications [1, 7].

Each patent is assigned one or more labels - "agricultural", "digital", or "other" - using a structured identification system combining patent classification codes - International Patent Classification (IPC) and Cooperative Patent Classification (CPC) - and semantic queries applied to titles and abstracts. A trajectory is classified as a smart farming trajectory if and only if it contains at least one agricultural patent and at least one digital patent along the same path. This definition captures the temporal and structural integration of agricultural and digital knowledge, including cases where digital technologies are introduced upstream and later applied in agricultural contexts, or vice versa.

To evaluate whether smart farming trajectories reflect structured knowledge integration rather than arbitrary recombination, we measure semantic coherence using cosine similarity between patent embeddings generated by PaECTER [19], a patent-specific document encoder fine-tuned on patent texts. Embeddings are derived from concatenated titles and abstracts; patents with missing abstracts are embedded using titles only, and all vectors are L2-normalized.

For each trajectory, we compute local step similarity (mean cosine similarity between consecutive patents) and global drift (cosine similarity between the first and last patents). We benchmark observed coherence against two counterfactual baselines: shuffled within-trajectory paths and year-matched cross-trajectory pairs. To quantify the magnitude of differences between distributions without assuming normality, we report Cliff's δ [20], interpreted as the probability that a randomly drawn true-trajectory similarity exceeds a randomly drawn baseline similarity. Finally, trajectories are further classified as agricultural-first, digital-first, or simultaneous based on the temporal order of domain entry.

Climate-related technologies are identified using the CPC Y02 and Y04S tagging scheme [16, 17], which provides a standardized cross-sectoral classification of technologies relevant to climate change mitigation and adaptation. Each patent is assigned a binary climate flag if it carries at least one Y02 or Y04S prefix. A trajectory is classified as climate-responsive if at least one patent along its path carries such a tag.

To assess whether smart farming trajectories integrate climate technologies differently from single-domain paths, we compute climate embedding rates by trajectory category, decompose Y02 sub-section composition, and examine the normalized position of climate entry within trajectories. A trajectory-level logistic regression estimates the association between climate embedding and structural trajectory characteristics, including length, SPNP centrality weight, and sequencing type.

4 Results

4.1 Patent Citation Networks and Main Path Extraction

Table 1 summarizes the patent citation network and its Network of Main Paths, alongside USPTO statistics for reference. The USPTO network is substantially larger and denser, averaging 19.9 citations per patent compared to 3.7 for EPO - reflecting the US "duty of disclosure" requirement versus examiner-driven EPO citations. This density difference directly affects NMP selectivity: the SPNP metric retains only 8.4% of USPTO edges as structurally important, compared to 35.0% for EPO. Both networks achieve 100% node coverage, meaning every patent connects to a complete source-to-sink trajectory.

Table 1: Patent citation network and NMP statistics.

Metric	USPTO	EPO
Patent applications (nodes)	10,265,025	3,469,815
Citation edges (after temporal filtering)	203,905,946	12,756,143
Citations per patent (avg.)	19.9	3.7
NMP edges (% of total)	17,212,730 (8.4%)	4,470,169 (35.0%)
Nodes with non-unique trajectories	75.5%	39.8%

The remainder of this paper presents results for the EPO network. The sparser EPO citation structure, driven by examiner-only citations, produces an NMP that is easier to interpret substantively, as each retained edge carries a more targeted knowledge-flow signal; the denser USPTO network produces a more aggressively filtered NMP (8.4%) but with greater path redundancy. USPTO results, which show broadly consistent patterns, are reserved for future comparative analysis. Note that because EPO and USPTO citations reflect different institutional practices - examiner-driven versus applicant-disclosed - they are not directly equivalent signals of knowledge flow, and cross-system comparisons require caution.

Patents were classified as agricultural using IPC code A01 (Agriculture; Forestry; Animal Husbandry; Hunting; Trapping; Fishing), yielding 91,803 agricultural patents. This boundary excludes food processing and supply-chain technologies outside A01; we adopt it for comparability with prior literature while acknowledging the scope limitation. Digital technologies were identified through a hybrid strategy combining World Intellectual Property Organization (WIPO) search definitions for frontier technologies with OECD J-Tag definitions for information and communication technologies (ICT) [21, 22], identifying 385,788 digital patents. The largest digital categories are high-speed network technologies (150,497), mobile communication (68,307), artificial intelligence (30,070), and robotics (23,449).

4.2 Smart Farming Trajectories

From the EPO Network of Main Paths, we extracted 719,054 unique maximum-SPNP trajectories passing through patents classified as "agricultural", "digital", or "both" (Table 2). The majority (75.5%) contain only digital patents, reflecting ICT dominance in the overall patent landscape. Smart farming trajectories - those containing at least one agricultural and at least one digital patent - number 26,885, representing 3.7% of domain-relevant paths.

Table 2: Distribution of trajectories by technological category (EPO).

Trajectory Category	Count	Share
Digital only	542,855	75.5%
Agricultural only	149,314	20.8%
Smart farming (both)	26,885	3.7%
Total	719,054	100%

Smart farming trajectories are structurally distinct from single-domain paths. They are systematically longer (mean length 9.5 patents versus shorter peaks for digital-only and agricultural-only categories), reflecting the longer citation chains required to bridge distinct technological communities. The average smart farming trajectory contains two to three digital patents and three to four agricultural patents, with approximately one patent classified as both. Smart farming trajectories also carry higher SPNP weights than single-domain trajectories, indicating that digital-agricultural knowledge integration occurs along the most structurally central citation links rather than peripheral connections.

4.3 Temporal Dynamics and Knowledge Integration

Digital-first trajectories - where digital technologies enter the citation path before agricultural applications - account for 50.5% of smart farming trajectories, with a median lag of 14 years before agricultural integration (Table 3). Agricultural-first trajectories (37.5%) show a median lag of 11 years before digital entry. The remaining 12% exhibit simultaneous entry of both knowledge types.

Table 3: Sequencing of knowledge entry in smart farming trajectories (EPO).

Sequence Type	Count	Share	Median Lag (years)
Digital-first	13,578	50.5%	+14
Agri-first	10,078	37.5%	-11
Simultaneous	3,229	12.0%	0

Fig. 1 shows how technological composition evolves along the normalized trajectory lifecycle. Patents classified as "other" (neither agricultural nor digital) dominate early trajectory positions, comprising over 55% of nodes in the first decile. This share declines steadily toward trajectory endpoints, where the agricultural share rises from approximately 20% to over 45%. The digital share remains relatively stable at 15-25% throughout. This pattern indicates that smart farming trajectories typically originate outside the agricultural and digital domains before progressively incorporating domain-specific innovations.

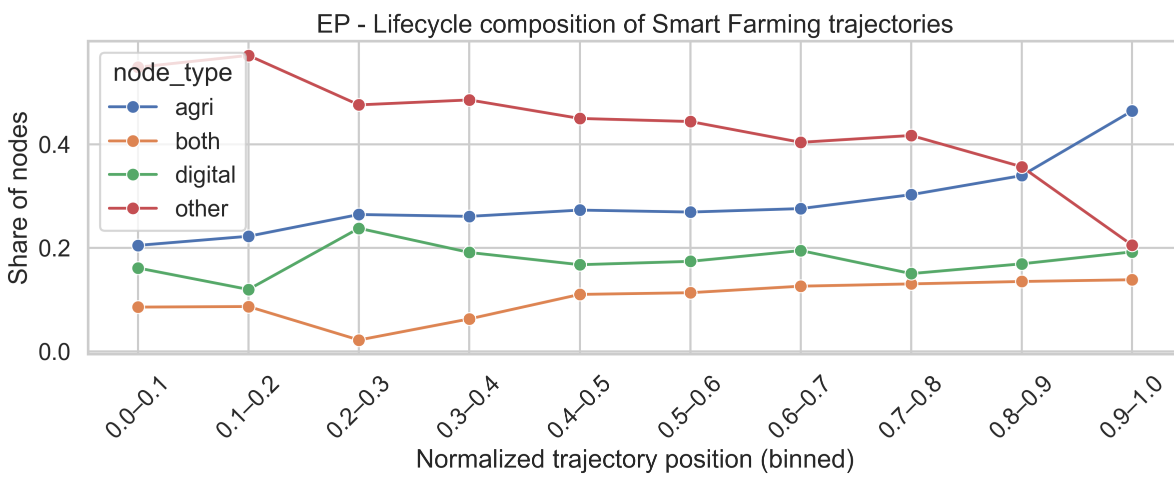

Figure 1: Technological composition along the normalized trajectory lifecycle (EPO). Smart farming trajectories progressively incorporate agricultural and digital innovations toward their endpoints.

Among digital technologies, autonomous systems are the most common first-entry point (5,241 trajectories), followed by high-speed networks (3,209) and robotics (2,535). WIPO frontier technologies such as AI (980) and IoT (14) appear less frequently as initial entry points, reflecting their more recent emergence. Foundational ICT infrastructure integrates with agriculture more rapidly (often under 5 years) than frontier technologies, which predominantly show lags of 10+ years.

Semantic coherence analysis confirms that these trajectories reflect structured knowledge integration rather than arbitrary recombination. True within-trajectory similarity (mean cosine similarity = 0.931) substantially exceeds both shuffled ordering (0.905, same patents in randomized sequence) and year-matched cross-trajectory baselines (0.865, patent pairs drawn from different trajectories but the same filing years), with a large effect size

(Cliff's $\delta \approx 0.91$, indicating that in 91% of random pairings a true-trajectory similarity exceeds the baseline). Digital-first trajectories exhibit greater semantic drift than agricultural-first paths, consistent with the longer adaptation process implied by the 14-year median integration lag.

4.4 Climate Technology Integration

Of the 26,885 smart farming trajectories, 7,031 (26.2%) contain at least one Y02- or Y04S-tagged patent. For comparison, 25.3% of agricultural-only and 20.1% of digital-only trajectories contain climate-tagged patents ($\chi^2 = 2212.9$, $df = 2$, $p < 0.001$). The near equality of smart farming and agricultural-only headline rates is misleading, however: decomposition by Y02 sub-section reveals that the two categories achieve similar overall rates through structurally different channels.

Table 4: Y02 sub-section composition by trajectory category (EPO). Values are row-normalized shares among climate-tagged patents.

CPC Section	Smart Farming	Agri-only	Digital-only
Y02A (adaptation)	0.270	0.479	0.022
Y02B (buildings)	0.083	0.032	0.094
Y02C (carbon capture)	0.001	0.011	0.001
Y02D (ICT efficiency)	0.058	0.000	0.471
Y02E (energy)	0.107	0.085	0.124
Y02P (agri/industry)	0.195	0.299	0.130
Y02T (transport)	0.215	0.033	0.074
Y02W (waste)	0.015	0.060	0.004
Y04S (smart grids)	0.055	0.001	0.080

As Table 4 shows, smart farming trajectories are heavily weighted toward Y02T (transport and machinery emissions, 21.5%) and Y02D (ICT energy efficiency, 5.8%) - categories virtually absent from agricultural-only paths - as well as Y04S (smart grids, 5.5% versus 0.1%). Agricultural-only climate integration, by contrast, is concentrated in Y02A (biological adaptation, 47.9%) and Y02P (production efficiency, 29.9%). The digital component of smart farming, therefore, does not increase the overall probability of climate relevance but redirects climate integration toward digitally-enabled decarbonization pathways - principally machinery electrification, network energy efficiency, and smart grid connectivity - that are structurally absent from single-domain agricultural innovation.

Climate technologies also enter trajectories predominantly at later stages. As Fig. 2 shows, approximately 50% of first climate patent entries occur in the late or terminal zone of trajectories (normalized position ≥ 0.75), and 21% appear only at the terminal node itself. Climate entry follows both digital and agricultural domain entry in the majority of cases (69.7% and 57.4% respectively), reinforcing the interpretation of climate technologies as downstream applications of digital-agricultural convergence rather than constitutive drivers of it. A patent-level logistic regression confirms that the probability of carrying a Y02 tag increases significantly with normalized trajectory position ($\beta = 1.78$, $z = 54.6$, $p < 0.001$).

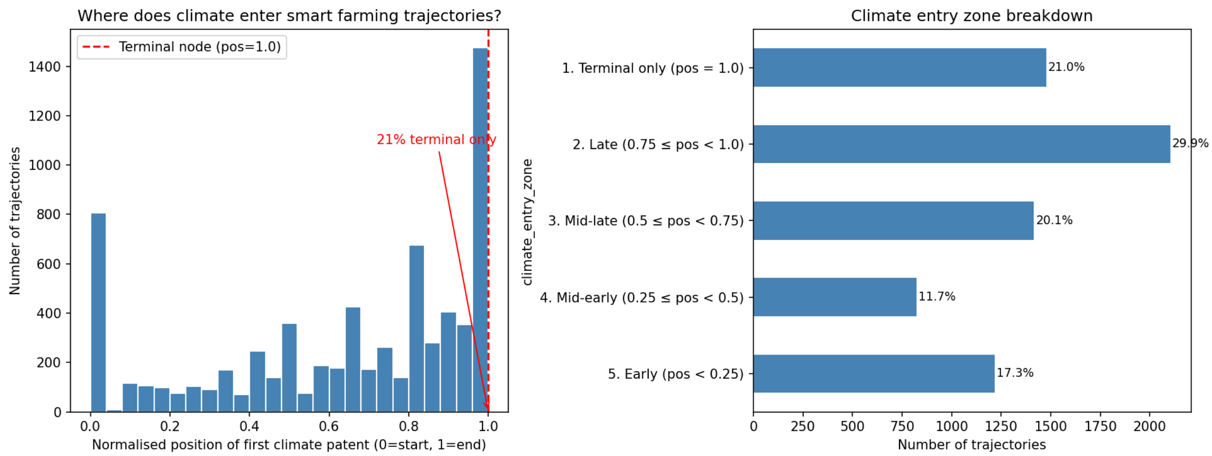

Figure 2: Position of first climate-tagged patent within smart farming trajectories (EPO). Left: distribution across normalized trajectory positions, with 21% appearing only at the terminal node. Right: breakdown by entry zone.

Climate embedding has also grown over time: among trajectories whose focal patent was filed before 1990, 12.1% contain a Y02-tagged patent, rising to 29.3% among 2020-2024 cohorts. Trajectory-level logistic regression confirms that digital-first sequencing is the strongest predictor of climate embedding (coefficient = 0.633, $z = 20.2$, $p < 0.001$), with digital-first trajectories exhibiting the highest climate embedding rate (31.6%) compared to agricultural-first (19.6%) and simultaneous (23.4%) types.

5 Conclusion

This paper proposes a trajectory-level framework for identifying and characterizing smart farming innovation using patent citation networks. Rather than defining smart farming through individual patent classifications, we identify it through the co-occurrence of agricultural and digital technologies along structurally important knowledge-flow paths extracted from the Network of Main Paths. Applied to the EPO patent system over 1980-2025, this approach identifies 26,885 smart farming trajectories - 3.7% of domain-relevant paths - that are structurally longer, carry higher centrality weights, and display distinct compositional patterns compared to single-domain trajectories.

Three substantive findings emerge. First, digital-agricultural knowledge integration is predominantly asymmetric: digital technologies enter trajectories before agricultural applications in over half of cases, with a median diffusion lag of 14 years, suggesting substantial adaptation rather than direct transfer. Second, foundational ICT infrastructure - particularly autonomous systems and high-speed networks - serves as the primary entry point for digital knowledge into agricultural trajectories. These established technologies function as bridging nodes that connect general-purpose citation chains to agricultural applications, whereas more recent frontier technologies, such as AI and IoT, tend to appear further along already-hybrid paths. Third, smart farming trajectories follow a characteristic lifecycle: they originate outside the agricultural and digital domains and progressively accumulate domain-specific agricultural and digital knowledge toward their endpoints. This gradient - visible in Fig. 1 - suggests cumulative, path-dependent knowledge recombination rather than isolated hybrid inventions. A fourth finding concerns climate integration: while smart farming and agricultural-only trajectories show comparable overall climate embedding rates, they achieve this through structurally different channels, with smart farming distinguished by machinery electrification (Y02T) and ICT energy efficiency (Y02D) rather than biological adaptation (Y02A). Climate technologies enter predominantly at the later stages of trajectories, consistent with their role as downstream applications of digital-agricultural convergence.

The framework has limitations. The trajectory-level definition captures sequential knowledge integration but does not distinguish active recombination from incidental co-occurrence. Patent data exclude non-patented innovations and knowledge flows outside formal IP systems - particularly important in agricultural contexts where public research institutions play central roles. Despite these limitations, the proposed indicators offer a scalable approach for monitoring technological convergence in agri-food systems. For national statistical offices and international organizations, trajectory-based metrics provide complementary evidence to innovation surveys, capturing how knowledge flows across technological boundaries. Future work will extend the comparative analysis across patent systems and explore how trajectory-level indicators relate to downstream outcomes in agricultural productivity and sustainability.

References

[1] Griliches Z. Patent statistics as economic indicators: a survey part I. Cambridge, MA: NBER; 1990.

[2] Hall BH, Jaffe AB, Trajtenberg M. The NBER patent citation data file: lessons, insights and methodological tools. Cambridge, MA: NBER; 2001.

[3] Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big data in smart farming—a review. *Agric Syst.* 2017;153:69–80.

[4] Hummon NP, Dereian P. Connectivity in a citation network: the development of DNA theory. *Soc Netw.* 1989;11(1):39–63.

[5] Nomaler Ö, Verspagen B. Greentech homophily and path dependence in a large patent citation network. Maastricht: UNU-MERIT; 2019.

[6] Trendov NM, Varas S, Zeng M. Digital technologies in agriculture and rural areas. Rome: FAO; 2019.

[7] Schroeder K, Lampietti J, Elabed G. What's cooking: digital transformation of the agri-food system. Washington, DC: World Bank Publications; 2021.

[8] Klerkx L, Jakku E, Labarthe P. A review of social science on digital agriculture, smart farming and agriculture 4.0: new contributions and a future research agenda. *NJAS Wagening J Life Sci.* 2019;90(1):1–16.

[9] World Intellectual Property Organization (WIPO). Patent landscape report - Agrifood. Geneva: WIPO; 2024.

[10] Haščič I, Migotto M. Measuring environmental innovation using patent data. OECD Environment Working Papers, No. 89. Paris: OECD Publishing; 2015.

[11] Barberá-Tomás D, Jiménez-Sáez F, Castelló-Molina I. Mapping the importance of the real world: the validity of connectivity analysis of patent citations networks. *Res Policy.* 2011;40(3):473–86.

[12] Batagelj V. Efficient algorithms for citation network analysis. arXiv preprint cs/0309023. 2003.

[13] Verspagen B. Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research. *Adv Complex Syst.* 2007;10(1):93–115.

[14] Fontana R, Nuvolari A, Verspagen B. Mapping technological trajectories as patent citation networks: an application to data communication standards. *Econ Innov New Technol.* 2009;18(4):311–36.

[15] Nomaler Ö, Verspagen B. Patent landscaping using 'green' technological trajectories. Maastricht: UNU-MERIT Working Papers; 2021.

[16] Angelucci S, Hurtado-Albir FJ, Volpe A. Supporting global initiatives on climate change: the EPO's "Y02-Y04S" tagging scheme. *World Pat Inf.* 2018;54:S85–92.

[17] European Patent Office (EPO). Updates on Y02 and Y04S [Internet]. Munich: EPO; 2023 [cited 2026 Feb 15]. Available from: https://www.epo.org/en/news-events/in-focus/classification/classification/updatesYO2andY04S

[18] European Patent Office (EPO). PATSTAT – EPO Worldwide Patent Statistical Database [Internet]. Vienna: EPO; 2025 [cited 2026 Feb 15]. Available from: https://www.epo.org/searching-for-patents/business/patstat.html

[19] Ghosh M, Erhardt S, Rose ME, Buunk E, Harhoff D. PaECTER: patent-level representation learning using citation-informed transformers. arXiv preprint 2402.19411. 2024.

[20] Cliff N. Dominance statistics: ordinal analyses to answer ordinal questions. *Psychol Bull.* 1993;114(3):494–509.

[21] World Intellectual Property Organization (WIPO). The direction of innovation. World Intellectual Property Report 2022. Geneva: WIPO; 2022.

[22] Inaba T, Squicciarini M. ICT: a new taxonomy based on the international patent classification. OECD Science, Technology and Industry Working Papers, No. 2017/01. Paris: OECD Publishing; 2017. Available from: https://doi.org/10.1787/ab16c396-en.