M@n@gement

From Network Effects to Data Network Effects: Enabling Ecosystemic Innovation for Sustainability

Wim Vanhaverbeke*

University of Antwerp, Antwerpen, Belgium

Abstract

This paper distinguishes between traditional network effects and data network effects, showing how the latter become central levers of sustainability transitions by enhancing efficiency, equity, and resilience across ecosystems. We argue that data network effects provide a more dynamic and sustainable basis for value creation, not only by enhancing firm-level performance but, more importantly, also by enabling ecosystemic innovation. They allow distributed actors across industries to cocreate novel solutions and adapt jointly to complex challenges, especially in ecosystems undergoing digital transformation … Drawing on case studies in precision agriculture, digital healthcare, and decentralized energy, we show how these effects improve performance, foster ecosystem-wide learning, and support sustainability objectives such as resource efficiency, equity, and resilience. A comparative framework clarifies key differences in value drivers, saturation risks, and competitive dynamics. We further explore how orchestrators – such as distribution system operators, healthcare consortia, and agri-data platforms – can become stewards of shared intelligence by aligning innovation strategies with ethical, open, and interoperable data governance. This paper concludes by outlining future research priorities, including sustainability metrics for data ecosystems and policy frameworks that enable responsible and inclusive platform evolution.

The growth of digital platforms over the last two decades has been closely associated with the concept of network effects. Platforms such as Facebook, Uber, and Tinder have thrived because their value increased as more users joined, creating powerful feedback loops. These traditional network effects are rooted in user growth – each additional participant makes the platform more valuable to others, either directly or indirectly. However, as artificial intelligence (AI) technologies and real-time data flows become embedded in the core operations of industrial sectors, a distinct and increasingly important mechanism has emerged: the data network effect. Or to use an analogy, traditional network effects are like a party: the more people who join, the more fun and useful it becomes. By contrast, data network effects are more like a(n online) library that (potentially) grows smarter as more books and experiences are added. The system becomes better at answering questions, even if the number of visitors does not grow. Unlike traditional models, data network effects do not depend primarily on expanding the user base. Instead, they derive value from the richness, diversity, and timeliness of data generated by users, sensors, and connected systems. These data are used to train algorithms that continuously learn and improve, generating a feedback loop in which platform performance enhances over time, independent of user count. As Govindarajan and Venkatraman (2024) argue, these data-driven feedback loops are becoming the defining source of advantage in industrial ecosystems. This aligns with broader research in digital platform studies: Hagiu and Wright (2023) emphasize how learning from usage data enhances service quality over time, while Gregory et al. (2021) and Parker et al. (2016) highlight the growing interplay between traditional and data network effects in hybrid platform architectures.

Yet, much of the existing literature focuses on data network effects as a firm-level asset or technological phenomenon. In this paper, we take a broader perspective by examining how data network effects can drive sustainability-oriented ecosystem innovation. Building on this insight, we argue that data network effects are not just mechanisms for improving firm

*Corresponding author: Wim Vanhaverbeke, Email: wim.vanhaverbeke@me.com

performance, but they are also core enablers of ecosystemic innovation. Because data-driven learning loops transcend firm boundaries, they allow distributed actors in an ecosystem to cocreate solutions that no single organization could develop alone. These ecosystem-wide learning loops reposition data network effects as enablers of collective innovation. In digitally transforming sectors, it is no longer the single firm that generates breakthrough solutions but the coordinated adaptation of multiple actors within an ecosystem. Data network effects thus become the connective tissue through which ecosystems experiment, learn, and innovate together. Moreover, the limitations of traditional network effects — plateaus, inequality, and congestion — contrast sharply with the sustainability potential of data network effects. Their ability to embed efficiency, ensure equitable participation, and build resilience positions them as essential mechanisms for achieving sustainability transitions. While related ideas such as data-enabled learning (Hagiu & Wright, 2023) or hybrid platform architectures (Gregory et al., 2021) exist, our contribution is to frame data network effects explicitly as enablers of ecosystemic innovation for sustainability. This shifts the focus from firm-level competitive dynamics to cross-sector innovation processes, which is new both conceptually and empirically.

To clarify the distinctions between network and data network effects — and why it matters — I present a comparative framework and sectoral examples from three industries undergoing deep digital transformation: precision agriculture, digital healthcare, and peer-to-peer energy systems. These cases illustrate how data network effects reshape strategic priorities — not by maximizing user acquisition but by enhancing collective intelligence and platform adaptiveness. The resulting insights inform not only platform strategy but also ecosystem governance and sustainability transitions in digitally mediated sectors.

The contributions of this paper are threefold. Theoretically, it clarifies how data network effects differ from and extend traditional network effects by emphasizing ecosystemic innovation. Empirically, it shows across agriculture, healthcare, and energy how these effects drive sustainability transitions. Practically, it provides guidance for orchestrators and policymakers on how to align data governance with efficiency, equity, and resilience.

Traditional network effects and their limits

Network effects occur when a product or service becomes more valuable as more people use it. The most cited examples are communication networks, where each additional user increases the utility for all others. Platform businesses extend this concept to two-sided or multi-sided markets, where indirect network effects arise between distinct groups (e.g., users and developers, riders, and drivers).

Traditional network effects have fueled the growth of iconic digital firms, from eBay to Airbnb. Their effectiveness lies in generating positive feedback loops where each new user adds value for others, making early-mover advantages especially strong (see Table 1). In transportation networks like Uber, for instance, an increase in drivers improves availability and reduces wait times for riders; simultaneously, more riders attract more drivers. Such dynamics help these platforms scale rapidly.

However, traditional network effects have critical limitations, particularly when considering the transition toward sustainability. First, value creation tends to plateau when market saturation is reached. Adding more users eventually delivers diminishing marginal returns, especially in saturated urban markets or overserved digital sectors. Second, these effects often emphasize scale over quality or innovation, reinforcing 'winner-takes-all' dynamics that may stifle diversity, inclusivity, and ecological resilience (Evans & Schmalensee, 2016).

Furthermore, traditional network effects prioritize transactional volume, not necessarily long-term ecosystem health. In sectors such as urban mobility or housing, platform scale can exacerbate congestion, inequality, and resource depletion. Airbnb's impact on housing affordability (Nieuwland & van Mellik, 2020) or Uber's contribution to traffic congestion (Li et al., 2022) illustrates these unintended consequences. As sustainability becomes a strategic priority, firms and regulators must ask whether network growth aligns with social and environmental outcomes.

Finally, competitive advantage derived from traditional network effects is often fragile. It is rooted in market dominance, switching costs, and brand recognition rather than continuous improvement. While this model may create strong short-term moats, it lacks the adaptability required in ecosystems where innovation must simultaneously serve sustainability transitions.

Data network effects: Concept and mechanism

In contrast, data network effects occur when products or services improve as more data are collected and used to train AI or other analytical systems. The key distinction is that improvement is not necessarily tied to the number of users but to the feedback loop between data acquisition, learning, and enhanced performance. Two core conditions underpin data network effects: (1) one user's data must improve the experience for other users, (2) these improvements must occur in real time or near real time, becoming perceivable to end users. In other words, my data should help improve your experience, and the improvements should be visible quickly enough that people notice the benefit.

This creates a fundamentally different logic of value creation. Instead of scale-driven feedback loops, firms harness learning loops. As Gregory et al. (2021) argue, these loops rely not on direct interaction between users but on the system's capacity

to learn from the full spectrum of digital traces. Platforms improve by detecting patterns, outliers, and contextual nuances from diverse user environments. This leads to personalization, automation, and new types of data-driven services.

Crucially, data network effects scale differently. In highly heterogeneous contexts – such as climate-sensitive agriculture or individualized healthcare – the richness and diversity of data are more valuable than the number of users alone. Each data point gathered from a unique condition (e.g., a specific microclimate, genetic profile, or energy usage pattern) adds to the system's predictive capacity. This expands the potential for hyper-personalized, adaptive, and environmentally responsive services.

In addition, the sustainability implications are profound. As firms integrate AI and data-driven optimization into core operations, they can enable more resource-efficient processes, reduce waste, and enhance resilience. For example, predictive maintenance powered by sensor data reduces material inputs and energy use in industrial operations. Adaptive logistics platforms reduce empty miles in freight, lowering emissions. Data network effects thus not only generate competitive advantage but can actively contribute to sustainability transitions (Bag et al., 2021).

Govindarajan and Venkatraman (2024) emphasize the role of 'datagraphs' – sensor-rich digital twins that integrate operational data across time and space. These models continuously evolve as they ingest new information from connected devices, enabling real-time optimization. In agriculture, such models help tailor irrigation to local soil moisture levels; in energy, they forecast consumption based on behavioral data and weather patterns. This real-time intelligence becomes an institutional asset that underpins both innovation and sustainability.

Similarly, Varian (2014), Schäfer and Sapi (2020), and recent work by Pentland et al. (2021) point to the emergence of algorithmic systems as central mediators of value. These systems discover latent user needs and optimize system-level outcomes through continuous experimentation. In platforms for food delivery, smart thermostats, or e-mobility sharing, algorithms incrementally adapt supply and demand, promoting efficiency at scale. When designed intentionally, these effects can align with ecological priorities – for example, minimizing peak energy loads or reducing per capita resource consumption.

Furthermore, new studies highlight how data network effects foster innovation ecosystems that align digital transformation with planetary boundaries. For instance, AI-powered industrial platforms in the chemical sector reduce greenhouse gas emissions through process optimization and closed-loop production (Saggar & Nigam). In the context of the circular economy, Kristoffersen et al. (2020) explain how data flows enable lifecycle monitoring, predictive reuse, and modular design. These insights extend the relevance of data network effects beyond competitiveness into the realm of long-term environmental stewardship.

Unlike traditional network effects, which often plateau or even generate negative externalities at scale, data network effects offer compounding, path-dependent advantages that reinforce both competitiveness and sustainability. Their effectiveness depends less on user numbers and more on a firm's ability to extract, process, and act upon data in diverse, evolving contexts. This shift favors organizations that invest in interoperable data infrastructure, open innovation partnerships, and transparent AI governance.

Data network effects thus represent not only a new model of platform strategy but a powerful lever for ecosystem-wide sustainability transformation. As ecosystems become increasingly digitalized, the capacity to generate shared intelligence from distributed data becomes central to long-term value creation, regulatory compliance, and societal legitimacy.

Yet, these effects are not automatic. If access to data is restricted, if data quality is uneven, or if algorithms overfit to biased inputs, the benefits may stall or even generate harmful outcomes. Overreliance on automated decision systems can also reduce human oversight, creating risks in safety-critical domains like healthcare.

Precision agriculture: From user networks to data-driven learning

Precision agriculture illustrates the power of data network effects in a sector traditionally characterized by product-based competition. Initially, digital platforms in agriculture emerged to support machinery sharing or market coordination among farmers. These early platforms benefited from traditional network effects: more users meant better resource utilization, more robust community features, and economies of scale in equipment services.

However, as digital technologies have penetrated farming practices more deeply, a structural shift has occurred. IoT-enabled devices, such as drones, soil sensors, and GPS-guided tractors, now generate high-resolution data on weather patterns, soil health, pest dynamics, and machinery performance. Platforms developed by companies like John Deere, Bayer's Climate FieldView, or IBM's Watson Decision Platform for Agriculture harness these data flows to provide AI-driven insights on when to plant, irrigate, fertilize, or harvest (Carbonell, 2016; Eastwood et al., 2019; Klerkx et al., 2019).

These platforms demonstrate a clear data network effect: as more farmers use the platform and contribute data, the algorithms become more accurate and context-sensitive. Importantly, this improvement is not simply additive; it is exponential when data diversity increases. A farmer operating in a novel microclimate or using alternative cropping practices adds a new node to the system's learning architecture. The model adapts and refines its recommendations based on collective use, creating a feedback loop that benefits the entire user base.

Sustainability outcomes arise when this learning loop is steered toward resource efficiency and ecological resilience. Precision agriculture platforms can significantly reduce overuse of fertilizers and pesticides, optimize water use, and minimize fuel consumption by mapping efficient field paths. For example, the use of variable rate application systems, informed by real-time data, has been shown to reduce nitrogen runoff and improve soil carbon retention (Zilberman et al., 2018). The integration of remote sensing data from satellite imagery with farm-level sensor inputs allows for dynamic adaptation to weather and climate conditions, enhancing resilience in the face of droughts or floods.

Moreover, data network effects support peer learning and benchmarking in regenerative agriculture. Platforms like Yara's Atfarm or the Global Open Data for Agriculture and Nutrition (GODAN) initiative promote transparent, standardized data sharing. This enables comparisons of soil organic matter levels, cover crop efficacy, and carbon sequestration across farms, creating collective intelligence for sustainability. As more farms participate, the precision and relevance of recommendations improve not only for high-tech commercial farms but also for smallholders practicing low-input, agroecological methods.

Data-driven precision farming is also increasingly embedded in ecosystem-based governance. Public-private partnerships, such as the EU's SmartAgriHubs or USDA's Climate Hubs, orchestrate data sharing across agronomic research institutions, input suppliers, and farmers. This ecosystem coordination amplifies data network effects across organizational boundaries and links them directly to regional sustainability strategies, including climate adaptation, biodiversity enhancement, and sustainable land use planning (Kanter et al., 2018).

Finally, ethical considerations and inclusivity must be addressed to realize the full sustainability potential of data network effects in agriculture. Issues of data ownership, interoperability, and algorithmic transparency influence who benefits from shared intelligence. Collaborative governance models and data trusts are emerging to ensure equitable participation and fair value distribution (Bronson & Knezevic, 2016).

In summary, precision agriculture demonstrates how data network effects not only transform productivity and competitiveness but also enable agricultural ecosystems to evolve toward greater sustainability. These effects are strongest when platforms foster cross-farm learning, integrate diverse data types, and embed sustainability objectives into the very structure of their optimization logic.

This illustrates how data network effects extend beyond enhancing yields for individual farms. They underpin an innovation dynamic in which farmers, suppliers, policymakers, and research institutes collaboratively generate practices that regenerate soils and ecosystems. In this sense, the value of the data network effect is inseparable from its capacity to enable ecosystemic innovation.

Digital healthcare: Diagnostic accuracy and ecosystem sustainability through shared learning

Healthcare is another domain where data network effects are becoming critical. In traditional models, telemedicine platforms or electronic health record (EHR) systems benefited from network effects: the more doctors and patients onboarded, the more useful the system became (Adler-Milstein & Jha, 2017; Agarwal et al., 2010).

But a more transformative shift is occurring through AI-enabled diagnostics. Consider a radiology platform that uses machine learning to analyze imaging data. As more scans are processed, the algorithm becomes better at detecting abnormalities. The value of the system improves not because more people are using it simultaneously, but because the data feeding the system enhance its accuracy and reduce diagnostic error (Erickson et al., 2017; Topol, 2019). Put simply, each scan makes the system smarter. Even if you are the only radiologist using it today, you benefit from all the scans analyzed before.

Importantly, these improvements benefit all users. A radiologist using the system today gets better decision support because of data collected from previous users. This form of shared learning across users and contexts is what gives data network effects their compounding power (Jiang et al., 2017; Kelly et al., 2019).

Beyond diagnostic imaging, data network effects are increasingly applied in clinical decision support systems, disease surveillance, and personalized medicine. Platforms like Tempus and Flatiron Health integrate genomic, clinical, and real-world data to predict treatment responses in oncology. As patient cohorts diversify and data sources increase in resolution, these platforms continuously improve their predictive models, accelerating the adoption of evidence-based and precision therapies (Rajpurkar et al., 2022).

Moreover, data-driven platforms play a key role in pandemic resilience and population health management. During COVID-19, health systems using real-time dashboards and federated AI systems could track intensive care unit (ICU) capacity, ventilator usage, and contagion clusters with greater accuracy, allowing for rapid policy responses and resource allocation (Rieke et al., 2020). These systems benefited not from user volume but from heterogeneous data inputs spanning geographies, demographics, and clinical settings.

From a sustainability perspective, data network effects reduce redundant procedures, optimize resource use, and help avoid overdiagnosis and overtreatment. For example, AI-supported triage systems in emergency rooms have been shown to reduce unnecessary admissions and imaging, lowering healthcare's environmental footprint (Secinaro et al., 2021). Diagnostic platforms that continuously learn from diverse patient data can also help lower disparities by tailoring interventions to underrepresented populations, thus promoting health equity.

Equally important is the role of federated learning in preserving patient privacy while enabling shared intelligence across institutions. In projects such as the EXAM study and Swarm Learning initiatives, models are trained on decentralized hospital data without transferring sensitive information, creating a secure data network effect that respects regulatory frameworks (Kaissis et al., 2020). This model has become central in EU healthcare AI strategy, where trust, interoperability, and sustainability are interlinked objectives.

Finally, data network effects can reinforce collaborative innovation across healthcare ecosystems. Shared learning platforms such as the Observational Health Data Sciences and Informatics (OHDSI) network connect universities, hospitals, and pharma companies to codevelop treatment pathways based on real-world evidence. As more partners share structured data using a common framework, algorithmic accuracy and treatment personalization improve collectively.

Digital energy: Data-driven orchestration in P2P markets

In energy markets, especially with the rise of decentralized and peer-to-peer (P2P) trading, DSOs (distribution system operators) are beginning to evolve from passive grid managers to active data coordinators. Traditional network effects in energy platforms occur as more prosumers join a P2P exchange, increasing liquidity and transaction potential (Parag & Sovacool, 2016; Sousa et al., 2019).

However, the real value shift comes from real-time data collection and AI-powered orchestration. Smart meters, weather forecasts, and battery sensors generate continuous streams of data. DSOs can use these data to forecast local demand, optimize grid usage, and anticipate congestion (Burger & Weinmann, 2019; Gautier et al., 2019). As more households provide granular consumption and production data, forecasting models become more precise. This benefits all participants by reducing volatility, improving pricing accuracy, and enhancing grid stability. Again, data network effects operate independently of user count. What matters is the richness

and timeliness of the data being collected and processed (Schittekatte et al., 2021; Schwanitz et al., 2017).

Crucially, these effects also create the basis for sustainability-oriented orchestration. Smart energy platforms like Power Ledger (Australia) and Vandebron (Netherlands) exemplify how data network effects can optimize local matching of renewable supply and demand. As prosumers generate real-time data on solar or wind energy availability, platforms can dynamically allocate resources to neighboring users, reducing transmission losses and supporting decarbonization targets.

AI-driven energy management systems further enhance efficiency by predicting demand peaks and coordinating distributed energy resources (DERs), including home batteries and electric vehicles. These systems can automatically shift consumption away from carbon-intensive periods or prioritize clean energy use, helping cities and regions meet climate action commitments (Entezari et al., 2023). In Denmark, grid operators use AI to balance wind power variability with industrial energy loads in real time, showcasing the deep integration of predictive data analytics into sustainability transitions.

Moreover, data network effects in energy platforms can support social equity. Equity-focused algorithms allow platforms to identify and compensate low-income households that provide flexibility to the grid. For instance, Time-of-Use tariffs that respond to behavioral data can reward users who shift consumption away from peak times, lowering both grid pressure and household bills. Some platforms even enable peer donations of surplus renewable energy to energy-poor households, linking data-driven efficiency with inclusive energy access (Georgarakis et al., 2021).

These developments are increasingly embedded within regional and national energy strategies. The EU's Clean Energy Package emphasizes data transparency, interoperability, and consumer empowerment as essential to the energy transition. Projects like InterConnect and OneNet aim to build interoperable data ecosystems that support smart grid integration, enhance resilience, and enable algorithmic coordination across the European electricity network.

Table 1. Comparative framework

Dimension	Traditional network effects	Data network effects
Value driver	User base growth	Continuous learning from diverse data
Saturation risk	High (plateaus at critical mass)	Low (improves with greater data heterogeneity)
Competitive advantage	Scale, lock-in effects	Model performance, learning curves
Sustainability impact	Often negative or neutral	Enables resource efficiency, equity, resilience
Example in Agriculture	Machinery sharing platforms	AI-driven precision and regenerative farming
Example in Healthcare	Telemedicine network size	Diagnostic AI and population health management
Example in Energy	More prosumers on P2P platforms	Smart, equitable energy orchestration

Source: own elaboration.

In summary, the transition to sustainable, decentralized energy systems hinges not only on hardware innovation but also on the strategic use of data. Data network effects enable smarter, fairer, and more adaptive energy ecosystems. Their success depends on collaboration across prosumers, DSOs, regulators, and platform providers – and on aligning algorithmic intelligence with ecological and societal objectives.

These examples underline that data network effects in energy are not just technical optimizers but enablers of ecosystemic innovation: they help orchestrators, prosumers, and regulators codesign pathways toward sustainable energy transitions.

Comparative framework

This comparative framework highlights the conceptual, strategic, and sustainability-related differences between traditional network effects and data network effects. These differences are not only academic in nature but also have practical implications for platform design, ecosystem governance, and policy formulation.

First, the value driver in traditional network effects lies in user base growth. Platforms thrive on interaction volume: the more users participate, the more others benefit from direct or indirect effects. However, data network effects shift the locus of value from user numbers to the continuous enrichment of data. Value accrues from diversity, granularity, and contextual richness of data, making heterogeneity more important than scale.

Second, saturation risk is higher in traditional models. Once critical mass is achieved, marginal benefits taper off, and negative externalities may emerge (e.g., congestion, platform bloat, or inequity). In contrast, data network effects improve over time if the system continues learning from novel data – for example, rare diseases in healthcare, microclimates in agriculture, or shifting demand patterns in energy systems.

Third, competitive advantage under traditional network effects often depends on early dominance and user lock-in, strategies that can entrench market power without necessarily improving the underlying service. Data network effects offer a more dynamic form of advantage: platforms that learn faster become more accurate, efficient, and tailored to user contexts, thus reinforcing adoption.

Fourth, the sustainability impact diverges sharply. Traditional platforms may unintentionally contribute to overconsumption, exclusion, or environmental strain – as seen in gig economy congestion or the resource-intensity of some social platforms. In contrast, data network effects – when properly aligned – can actively support sustainability by reducing waste (e.g., predictive farming), enhancing resilience (e.g., energy demand response), and promoting inclusivity (e.g., equitable diagnostics or smart subsidies).

The revised examples from agriculture, healthcare, and energy illustrate how data network effects drive ecosystem-wide benefits. In agriculture, aggregated field data train models that guide regenerative practices. In healthcare, distributed data improve diagnostics while protecting patient privacy. In energy, data-rich orchestration tools balance the grid while ensuring participation from low-income prosumers. In each sectoral case, sustainability arises not incidentally but directly from data network effects: precision farming optimizes inputs and safeguards soil health (efficiency), federated healthcare learning reduces disparities across patient groups (equity), and adaptive energy orchestration enhances grid stability against shocks (resilience).

Taken together, these contrasts suggest a shift in the design logic of platforms: from maximizing interactions to cultivating collective intelligence. This shift calls for new capabilities in data governance, interoperability, and cross-sector collaboration. Policymakers and ecosystem orchestrators must foster environments where such learning-based, sustainability-aligned platforms can thrive.

Ecosystem innovation for sustainability: A data-driven perspective

This section extends the firm-level perspective by showing how data network effects create interdependencies that structure ecosystem-wide innovation dynamics. Understanding data network effects solely through the lens of firm-level advantage misses a crucial implication: these effects can be designed and governed to catalyze ecosystem innovation for sustainability. Ecosystem innovation involves the coordinated transformation of products, services, and institutional arrangements across multiple organizations to achieve shared sustainability goals (Bocken et al., 2014; Boons & Lüdeke-Freund, 2013).

Three ecosystem-level mechanisms link data network effects to sustainability-oriented innovation. First, the continuous feedback loops inherent in data network effects enable ecosystems to experiment, learn, and adapt dynamically. This is a core requirement for complex sustainability transitions. Second, these effects create interdependencies that reward cross-sector collaboration, such as integrating agriculture and energy data to optimize water use and carbon emissions. Third, data-driven ecosystems can develop collective intelligence infrastructures, such as open data standards or federated learning models, which democratize access to insights while preserving data sovereignty.

Effective ecosystem orchestration for sustainability thus hinges on both technological interoperability and institutional governance. Orchestrators must develop data-sharing frameworks that align private incentives with public goods creation, often by introducing novel incentive designs or

sustainability metrics. In agriculture, for instance, shared data platforms can support regenerative farming by enabling peer benchmarking of soil health and carbon sequestration (Jayasinghe et al., 2023). In healthcare, federated AI models that preserve patient privacy can still enhance diagnostic accuracy across hospitals, aligning efficiency with ethical care (Teo et al., 2024). And in energy, prosumer platforms can incorporate equity-focused algorithms that ensure fair compensation for low-income contributors to the green grid (Georgarakis et al., 2021).

These sectoral examples not only validate the concept but also show the operational diversity of ecosystem innovation. In precision agriculture, multi-actor platforms coordinate data flows between farmers, seed suppliers, agronomists, and public agencies, enabling smarter crop planning and climate adaptation. In digital health, interoperability standards like HL7 FHIR (HL7 International, 2023) underpin secure data exchange between diagnostic AI providers, public health agencies, and hospitals. In decentralized energy markets, orchestrators like DSOs and energy service companies manage real-time data exchange among prosumers, grid operators, and local authorities to ensure fair, efficient, and low-carbon energy distribution.

These arrangements reflect a shift *from firm-centric value capture to distributed, purpose-driven value creation*. They also necessitate new forms of governance – such as data trusts, public-private data partnerships, and algorithmic transparency mandates – that reinforce accountability and inclusiveness. As such, sustainability-focused data ecosystems are not just technological infrastructures but institutional innovations that reconfigure power, access, and responsibility.

To realize this potential, ecosystem partners must treat data as a shared strategic asset. This requires co-investment in infrastructure that supports open innovation, ethical AI, and adaptive governance. It also demands long-term commitment from public actors to create enabling environments for sustainability-aligned platforms through regulatory support, funding schemes, and participatory design.

Discussion

Understanding the difference between traditional and data network effects is essential for strategy in sectors that are digitally transforming. For organizations, the key implication is that user acquisition is no longer the only pathway to value. Instead, building capabilities for data integration, real-time analytics, and AI governance becomes central. This requires a shift in how digital initiatives are conceived and managed. Rather than simply scaling user participation, firms must focus on creating high-quality data loops, investing in infrastructure that supports interoperability and deploying AI systems that improve over time. The accumulation of valuable data and the ability to learn from it become a core strategic asset.

Moreover, the ability to harness data network effects depends not only on technical capabilities but also on organizational learning and cross-functional alignment. Data scientists, domain experts, and digital strategists must collaborate to identify where learning loops can be created, and how they can be amplified. In this sense, strategic control shifts from market share acquisition to algorithmic performance and adaptation. Companies that master this shift are better positioned to capture long-term value, resist commoditization, and respond flexibly to external shocks.

Crucially, data network effects challenge organizations to reconceive their role in broader ecosystems. Firms are no longer just service providers or platform operators – they are *stewards of shared intelligence*. In agriculture, firms cocreate value by contributing to collaborative soil health models. In healthcare, hospital consortia use federated learning to enhance diagnostic performance without compromising patient privacy. In energy, local utilities optimize decentralized supply and demand in ways that reflect both carbon goals and affordability. These examples illustrate that competitive advantage now coexists with systemic value creation. By shifting the unit of analysis from firms to ecosystems, data network effects thus reframe competition itself: value stems less from protecting proprietary advantage and more from contributing to ecosystemic learning loops. This reconceptualization highlights why understanding them as enablers of ecosystemic innovation is critical.

From a regulatory perspective, data network effects raise *new challenges*. The compounding nature of learning from data can create highly entrenched positions, especially when the data are proprietary. This can reinforce barriers to entry and enable dominant players to entrench their market leadership not just through scale but through superior models that continuously self-improve. This necessitates discussions about data-sharing mandates, open standards, and ethical oversight. Regulators must address the risk of algorithmic opacity and concentration of control over critical learning systems, particularly in sectors like healthcare and energy where societal interests are at stake. Effective policy will need to balance innovation incentives with the need for open, fair, and transparent ecosystems.

For ecosystem orchestrators such as DSOs or healthcare platforms, the shift suggests a role not just as service providers but as data custodians who facilitate collective learning while maintaining trust and compliance. These actors are uniquely positioned to govern access to shared data infrastructures and to design incentive structures that encourage data sharing without compromising individual rights or competitive neutrality. Their role increasingly resembles that of digital infrastructure stewards – setting the rules, maintaining interoperability, and ensuring resilience of learning loops across organizational boundaries. To do so effectively, they must invest *in both governance models and technical*

architectures that can accommodate diverse stakeholders, evolving technologies, and shifting policy frameworks.

As a consequence, the strategic, organizational, and regulatory implications of data network effects require a rethinking of how digital transformation is understood and managed. They challenge traditional notions of scale, competition, and control and demand new models of collaboration, capability-building, and policy design. Emerging concepts such as 'data collaboratives', 'digital commons', and 'algorithmic governance' reflect a broader reorientation toward ecosystem-level strategy and distributed accountability. Moreover, measuring data network effects remains challenging. Potential indicators include the speed of accuracy improvements in predictive models, reductions in resource use per data increment, or equity gains in service access across user groups. Developing standardized sustainability metrics for data ecosystems remains an urgent research agenda.

The concept of data network effects also offers a novel lens through which to understand sustainability transitions. Unlike static technology adoption models, ecosystems that cultivate learning loops across stakeholders can accelerate the emergence of sustainable practices. For instance, real-time feedback from energy prosumers or precision farmers can be integrated into distributed intelligence systems that continuously adapt to environmental conditions, consumer behavior, and regulatory shifts. This adaptability is vital for addressing the dynamic nature of sustainability challenges, from climate volatility to aging populations. Furthermore, data network effects can facilitate the alignment of economic incentives with environmental performance – for example, through dynamic pricing models that reward energy conservation or regenerative farming practices.

Ecosystem orchestrators thus face a dual challenge: *they must govern for both data quality and sustainability outcomes*. This means balancing innovation with ethics, openness with control, and short-term efficiency with long-term systemic resilience. The orchestrator's success will depend on its capacity to embed data governance in institutional design and to shape data flows in ways that support inclusive, adaptive, and sustainability-oriented innovation.

By embedding sustainability parameters directly into algorithms – such as carbon intensity thresholds, equity weights, or resilience indices – ecosystems can ensure that each new cycle of data-driven learning actively contributes to sustainability transitions. This makes sustainability a design principle rather than a by-product. From a policy standpoint, enabling ecosystem innovation for sustainability through data network effects requires targeted interventions. Governments can play a catalytic role by funding interoperable data infrastructures, mandating open standards for data sharing in key sectors, and incentivizing data collaboratives that align with Sustainable Development Goals (SDGs). These interventions should be codesigned with ecosystem stakeholders to ensure legitimacy, adaptability, and long-term impact.

Conclusion

While network effects have been a cornerstone of platform strategy, they are no longer the sole mechanism of value creation in digital markets. Data network effects offer a more dynamic and sustainable form of competitive advantage, but more importantly, they are enablers of ecosystemic innovation for sustainability transitions. Recognizing and leveraging these effects requires a reorientation of strategy, governance, and innovation practices.

As demonstrated in precision agriculture, digital healthcare, and P2P energy, data network effects can be instrumental in fostering ecosystem-wide learning, innovation, and sustainability. They represent not only a new logic of competition but also a new paradigm of collective value creation.

Future research should explore how data network effects can be designed to serve sustainability transitions. Promising avenues include developing metrics to measure their environmental and social impacts at the ecosystem level, examining trade-offs between data sharing and competitive positioning and identifying policy frameworks that foster open yet responsible data collaboration. Understanding the governance of learning loops, and how these can be ethically and inclusively scaled, will be central to leveraging data network effects for a just and sustainable digital transformation. At the same time, recognizing their limits – data access, quality, bias, and governance – is essential for harnessing data network effects responsibly.

In summary, this paper contributes by (1) introducing a comparative framework distinguishing traditional and data network effects, (2) demonstrating their role as enablers of ecosystemic sustainability innovation, and (3) offering some actionable insights for orchestrators and policymakers. Together, these contributions extend current debates on platforms, ecosystems, and sustainability transitions.

References

Adler-Milstein, J. & Jha, A. K. (2017). HITECH Act drove large gains in hospital electronic health record adoption. *Health Affairs*, *36*(8), 1416–1422. https://doi.org/10.1377/hlthaff.2016.1651

Agarwal, R., Gao, G., DesRoches, C. & Jha, A. K. (2010). Research commentary – The digital transformation of healthcare: Current status and the road ahead. *Information Systems Research*, *21*(4), 796–809. https://doi.org/10.1287/isre.1100.0327

Bag, S, Pretorius, J. H. C., Gupta, S. & Dwivedi, Y. K. (2021). Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities, *Technological Forecasting and Social Change*, *163*, 120420. https://doi.org/10.1016/j.techfore.2020.120420

Bocken, N. M. P., Short, S. W., Rana, P. & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production*, *65*, 42–56. https://doi.org/10.1016/j.jclepro.2013.11.039

Boons, F. & Lüdeke-Freund, F. (2013). Business models for sustainable innovation: State-of-the-art and steps towards a research agenda. *Journal of Cleaner Production*, 45, 9–19. https://doi.org/10.1016/j.jclepro.2012.07.007

Bronson, K. & Knezevic, I. (2016). Big data in food and agriculture. *Big Data & Society*, 3(1). https://doi.org/10.1177/2053951716648174

Burger, C. & Weinmann, J. (2019, May). Business model innovation of blockchain-based energy trading platforms. In *Local energy, global markets: 42nd IAEE international conference*, May 29–June 1, 2019. International Association for Energy Economics.

Carbonell, I. M. (2016). The ethics of big data in big agriculture. *Internet Policy Review*, 5(1), 1–11. https://doi.org/10.14763/2016.1.405

Eastwood, C., Klerkx, L., Ayre, M. & Dela Rue, B. (2019). Managing socio-ethical challenges in the development of smart farming: From a fragmented to a comprehensive approach for responsible research and innovation. *Journal of Agricultural and Environmental Ethics*, 32(5), 741–768. https://doi.org/10.1007/s10806-017-9704-5

Entezari, A., Aslani, A., Zahedi, R. & Noorallahi, Y. (2023). Artificial intelligence and machine learning in energy systems: A bibliographic perspective. *Energy Strategy Reviews*, 45, 101017. https://doi.org/10.1016/j.esr.2022.101017

Erickson, B. J., Korfiatis, P., Akkus, Z. & Kline, T. L. (2017). Machine learning for medical imaging. *Radiographics*, 37(2), 505–515. https://doi.org/10.1148/rg.2017160130

Evans, D. S. & Schmalensee, R. (2016). *The new economics of multi-sided platforms*. Harvard Business Review Press.

Gautier, A., Hoet, B., Jacqmin, J. & Van Driessche, S. (2019). Self-consumption choice of residential PV owners under net-metering. *Energy Policy*, 128, 648–653. https://doi.org/10.1016/j.enpol.2019.01.055

Georgarakis, E., Bauwens, T., Pronk, A.-M. & AlSkaif, T. (2021). Keep it green, simple and socially fair: A choice experiment on prosumers' preferences for peer-to-peer electricity trading in the Netherlands. *Energy Policy*, 159, 112615. https://doi.org/10.1016/j.enpol.2021.112615

Govindarajan, V. & Venkatraman, N. (2024). *Fusion strategy: How real-time data and AI will power the industrial future*. Harvard Business Review Press.

Gregory, R. W., Henfridsson, O., Kaganer, E. & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review*, 46(3), 534–551. https://doi.org/10.5465/amr.2019.0178

Hagiu, A. & Wright, J. (2023). Data-enabled learning, network effects and competitive advantage. *The RAND Journal of Economics*, 54(4), 638–667. https://doi.org/10.1111/1756-2171.12453

HL7 International. (2023). *HL7 FHIR: Fast healthcare interoperability resources*. Retrieved from https://www.hl7.org/fhir/

Jayasinghe, S. L., Thomas, D. T., Anderson, J. P., Chen, C. et al. (2023). Global application of regenerative agriculture: A review of definitions and assessment approaches. *Sustainability*, 15(22), 15941. https://doi.org/10.3390/su152215941

Jiang, F., Jiang, Y., Zhi, H., Dong, Y. et al. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Kaissis, G. A., Makowski, M. R., Rückert, D. & Braren, R. F. (2020). Secure, privacy-preserving and federated machine learning in medical imaging. *Nature Machine Intelligence*, 2, 305–311. https://doi.org/10.1038/s42256-020-0186-1

Kanter, D. R., Musumba, M., Wood, S. L. R., Palm, C. et al. (2018). Evaluating agricultural trade-offs in the age of sustainable development. *Agricultural Systems*, 163, 73–88. https://doi.org/10.1016/j.agsy.2016.09.010

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. et al. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1), 195. https://doi.org/10.1186/s12916-019-1426-2

Klerkx, L., Jakku, E. & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. *NJAS: Wageningen Journal of Life Sciences*, 90–91(1), 1–16. https://doi.org/10.1016/j.njas.2019.100315

Kristoffersen, E., Blomsma, F., Mikalef, P. & Li, J. (2020). The smart circular economy: A digital-enabled circular strategies framework for manufacturing companies. *Journal of Business Research*, 120, 241–261. https://doi.org/10.1016/j.jbusres.2020.07.044

Li, Z., Liang, C., Hong, Y. & Zhang, Z. (2022). How do on-demand ridesharing services affect traffic congestion? The moderating role of urban compactness. *Production and Operations Management*, 31(1), 239–258. https://doi.org/10.1111/poms.13530

Nieuwland, S. & van Melik, R. (2020). Regulating Airbnb: How cities deal with perceived negative externalities of short-term rentals. *Current Issues in Tourism*, 23(7), 811–825. https://doi.org/10.1080/13683500.2018.1504899

OECD. (2021). *Enhancing access to and sharing of data: Reconciling risks and benefits for data re-use across societies*. OECD Publishing.

Parag, Y. & Sovacool, B. K. (2016). Electricity market design for the prosumer era. *Nature Energy*, 1. https://doi.org/10.1038/nenergy.2016.32

Parker, G. G., Van Alstyne, M. W. & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. W. W. Norton & Company.

Pentland, A., Lipton, A. & Hardjono, T. (2021). *Building the new economy: Data as capital*. MIT Press. https://doi.org/10.7551/mitpress/13991.001.0001

Rajpurkar, P., Chen, E., Banerjee, O. & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28(1), 31–38. https://doi.org/10.1038/s41591-021-01614-0

Rieke, N., Hancox, J., Li, W., Milletarì, F. et al. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3, 119. https://doi.org/10.1038/s41746-020-00323-1

Saggar, A. & Nigam, B. (2023). Maximising net zero in energy-intensive industries: An overview of AI applications for greenhouse gas reduction. *Journal of Climate Change*, 9(1), 13–23. https://doi.org/10.3233/JCC230003

Schäfer, M. & Sapi, G. (2020). *Learning from data and network effects: The example of internet search* (DIW Discussion Papers, No. 1894). Deutsches Institut für Wirtschaftsforschung.

Schittekatte, T. & Meeus, L. (2020). Flexibility markets: Q&A with project pioneers. *Utilities Policy*, 63, 101017. https://doi.org/10.1016/j.jup.2020.101017

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V. et al. (2021). The role of artificial intelligence in healthcare: A structured literature review. *BMC Medical Informatics Decision Making*, 21, 125. https://doi.org/10.1186/s12911-021-01488-9

Sousa, T., Soares, T., Pinson, P., Moret, F. et al. (2019). Peer-to-peer and community-based markets: A comprehensive review. *Renewable and Sustainable Energy Reviews*, 104, 367–378. https://doi.org/10.1016/j.rser.2019.01.036

Teo, Z. L., Jin, L., Liu, N., Li, S. et al. (2024). Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. *Cell Reports Medicine*, 5(2), 101419. https://doi.org/10.1016/j.xcrm.2024.101419

Topol, E. J. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.

Varian, H. R. (2014). Beyond big data. *Business Economics*, 49(1), 27–31. https://doi.org/10.1057/be.2014.1

Zilberman, D., Gordon, B, Hochman, G. & Wesseler, J. (2018). Economics of sustainable development and the bioeconomy. *Applied Economic Perspectives and Policy*, 40(1), 22–37. https://doi.org/10.1093/aepp/ppx051