

Value Creation and Value Capture in AI: A Triple Helix Model

Antoni Lorente

King's College London
antoni.lorete_martinez@kcl.ac.uk

Abstract

The idea of 'value chain' faces limitations as an effective concept to address the complex mechanisms of value creation and value capture mechanisms in Artificial Intelligence. Starting with a review of Michael Porter's original concept, this paper explores the use of the term in the AI context, highlighting the need for a more robust theoretical foundation to prevent reductionist accounts of value. To address the theoretical gap and make the complexity of value creation and value capture in AI tractable, we introduce the triple helix model of AI—a comprehensive analytical framework that decomposes AI into its three fundamental strands: data, architecture, and computational power. With this, AI can be understood at the intersection of these three strands in each social context, capturing different value dynamics. Our methodology advances a three-phase approach to thinking about the triple helix: (1) an ontology or list that identifies key elements and their relationships; (2) a kinematic account which portrays—in a historical way—the evolution of these components; and (3) a dynamic account, which allows the identification of the driving forces of such evolution. This allows the drawing-up of a comprehensive and faithful account of AI stagnations and booms while opening the door to reading AI evolution as a story of power.

Introduction

In *Competitive Advantage: Creating and Sustaining Superior Performance* (Porter 1985/2011), Michael Porter introduced the concept and terminology of the 'value chain' as a strategic tool for diagnosing and enhancing firm-level competitive advantage. The conceptual power of Porter's value chain framework lies in its capacity to disaggregate a firm into its discrete activities (i.e., design, marketing, production, distribution and support of its product) thereby capturing the company's history and strategy (Porter 1985/2011, p. 36).

The value chain functions most effectively when applied to a firm's activities within a specific industry or business unit (Porter 1985/2011). Within this scope, it becomes

possible to identify value-adding steps that contribute to a final product, enabling firms to identify and exploit their competitive advantages. However, most industries operate through networks of concatenated players, each possessing distinct value chains where one firm's output serves as another's input and so on. Consider, for example, cement manufacturers who transform stone into powder, which construction companies use as material input, ultimately providing the foundation for real estate development.

When the value chains of all the firms partaking in each supply chain or network are aggregated (including nested projections into supply and delivery, until reaching the raw components and the final consumer respectively), the result constitutes what's known as a 'value system'. Building on Porter's seminal work, different frameworks and tools have been derived, including the Value-Stream Mapping technique (Rother and Shook 2003) and frameworks for analysing the aggregation and distribution of value chains at a global level with global value chains (GVC)—such as the ones developed and applied by the OECD (OECD 2013).

Industry evolution and technological progress, driven by increasing globalisation, decentralisation, and technological innovation, have underscored the need to develop industry-level value chains or value systems to better represent the processes to produce goods and services. However, the technological transformation of the first quarter of the 21st century has fundamentally altered not only how we live but also the analytical models used to understand, organise, and even commercialise products and services. According to a 2017 OECD study, communication technologies will continue fostering the globalisation of value chains while simultaneously shortening certain chains through technological innovation and efficiency gains, particularly through the integration of AI and advanced robotics across different production stages.

Considering this, this paper develops an alternative framework for understanding value creation and capture in

AI systems through three interconnected analyses. First, I explore the conceptual limitations of value chains in AI, underscoring the need for a more sophisticated framework. Second, I introduce the triple helix model of AI, which decomposes AI systems into their three fundamental strands—data, architecture, and computational power—demonstrating how these interact in non-linear ways within specific social contexts. Third, I present a three-phase analytical approach: an ontological mapping that identifies key elements and relationships; a kinematic account that traces historical evolution; and a dynamic investigation that uncovers driving forces shaping development patterns. Through this framework, I discuss how AI value generation extends beyond economic metrics to include social, political, and ethical dimensions that traditional value chain analyses obscure. By revealing these additional value dynamics, this paper explores how a permissive use of the concept of an AI value chain hinders the exploration of value creation and value capture in AI.

‘The AI Value Chain’: A Concept Without Definition

In March 2018, the European Commission announced the creation of an AI expert group and a European AI Alliance. Two years later, it published a white paper that would become the seed of Regulation 2024/1689, commonly known as the ‘EU AI Act’. Since the European Commission introduced the concept of ‘the AI value chain’ in this first white paper, all subsequent regulatory drafts, including the final version, have referenced the AI value chain, while discussing its complexity or the need to distribute responsibilities across its components.

However, the notion of an ‘AI value chain’ has not undergone a rigorous conceptual development comparable to Porter’s seminal work on the generic value chain framework. While some organisations, such as the OECD, have adopted the concept in their publications (Muñoz 2022; Tiedrich 2023), the AI value chain remains a poorly defined construct, frequently conflated with the machine learning (or ML) value chain, which focuses on the integration of machine learning into operations processes in business.

The increasing diversity of what contemporary AI applications, encompassing a myriad of techniques and approaches, combined with emerging business opportunities and genuine advances towards generality (broadly understood as the showcasing of different capabilities across tasks or modalities) have exposed the limitations of assuming a predetermined, singular value chain for AI.

Both the regulatory pressure from the AI Act and the technological advances in multi-purpose and increasingly general AI have underscored the need for a definition of generality. To facilitate governance, Gutiérrez et al. provide

a qualitative definition of General-Purpose AI Systems as “[systems] that can accomplish or be adapted to accomplish a range of distinct tasks, including some for which [they were] not intentionally and specifically trained” (Gutiérrez et al. 2022).

Hernández-Orallo et al. (2021) adopt a more systematic approach by distinguishing between the notions of generality and capability. Generality captures the level of specialisation (the broader the range of tasks an agent can perform, up to a level of difficulty, the more general it is). Capability captures the level of difficulty (more capable agents can perform increasingly challenging tasks). In this context, the authors propose that “independently of its overall capability, an agent can only be called fully general if it covers all tasks up to an equivalent level of difficulty, determined by the resources that are needed for them” (Hernández-Orallo et al. 2021, emphasis their own). This framework disentangles the concept of general intelligence by introducing direct measures of generality and capability for individuals, independent of population comparisons, and suggests these could serve as quantitative measures for both artificial and natural intelligence. But more importantly, this analysis reveals the fundamental problem to define AI.

Within the European regulatory context, co-legislators have distinguished between narrow and general-purpose AI (Regulation 2024/1689), yet the complexity of the value chain remains unresolved. To define the AI value chain, Engler and Renda state that

“the AI value chain [can be defined] as the organisational process through which an individual AI system is developed and then put into use (or deployed). Proactively considering which organisations may execute on which parts of the AI value chain is necessary to ascertain which organisation is best placed to conform with regulatory requirements, and more generally to ensure safe and responsible function of the AI system. This focus on the organisational process distinguishes the AI value chain from algorithmic development, in which the focus is more narrowly on the technological steps of building an AI system.” (Engler and Renda, 2022)

The authors identify several developmental stages: problem definition, data collection and pre-processing, model training, model retraining, model testing and evaluation, integration into software, and model deployment. They also categorise two primary stakeholder types: providers and users (Engler and Renda 2022). Their report documents various value chain configurations that reflect how different stages and services can be distributed across one, two, and multiple entities.

However, Engler and Renda’s definition fails to capture the complexity of value generation and value capture in AI. Returning to Porter’s original discussion of the value chain

and how this concept can be extended along the supply chain to build the notion of the ‘value system’ illuminates how the discrete activities required to develop and deploy fully functional AI systems cannot be adequately contained within traditional value chain frameworks designed for other industries. While the notion of a ‘value system’ affords a higher level of complexity, capable of reflecting some of the relationships between the different actors involved in the process of developing and deploying AI systems, it fails to capture the interactions between the three strands of AI—which operate in parallel rather than sequentially—and the multiple forms of value beyond purely economic considerations that can be generated around AI systems and captured by different stakeholders.

As capital flows into AI, and strategic partnerships proliferate (OpenAI 2023; Microsoft 2023), it becomes increasingly apparent that AI development targeting generality has become a promising business model without the need to put forward clearly defined problem statements—beyond that of developing powerful AI. And while application-agnostic technological development is not inherently problematic, the mechanisms determining where value resides remain unclear.

Following ChatGPT’s release, social media feeds and traditional newspapers were inundated with articles, memes, opinion pieces, and statements about the model and AI more broadly. Sam Altman, OpenAI’s CEO, tweeted on December 5: “ChatGPT launched on Wednesday [November 30th]. today it crossed 1 million users!” (Altman 2022). The remarkable scale at which ChatGPT was made freely available and used may have allowed OpenAI to capture part of the market while monitoring the use of its product as an extensive beta test before integration into one of Microsoft’s products—a new form of strategy development that could change the way AI companies shape and relate with the market itself for good.

The partnership between Microsoft and OpenAI was operationalised through cloud computing services, with Microsoft providing exclusive computational resources in exchange for privileged access to cutting-edge AI-powered functionalities developed by OpenAI. This arrangement highlights some of the fundamental problems related to the ambiguity of ‘(general purpose) AI value chains’. The traditional conception of AI beyond its constituent strands is no longer adequate for comprehending and modelling value generation by the different actors, who do not operate along sequential value chain steps but rather provide the essential elements necessary for value generation across several steps.

The Triple Helix Model

Artificial Intelligence emerges at the intersection of three fundamental strands: architecture, data, and compute. Beneath the terminological complexity that characterises AI discourse, AI ultimately results from automating data (or meta-data) patterns via algorithms that require computational power. AI manifests across a broad spectrum of implementations, ranging from a simple linear regression models to sophisticated expert system or a more complex transformer-based architectures. While each form imposes distinct requirements across the three dimensions, all necessitate the integration of architecture, data, and compute to function effectively.

Value generation in AI can thus be better modelled by decomposing the research, development and deployment processes of AI systems into the contributions along each one of its three strands. This analytical approach offers both a powerful and versatile framework while reducing the complexity inherent in identifying value generation across AI artifacts, a complexity which only intensifies given the radically different nature of each strand.

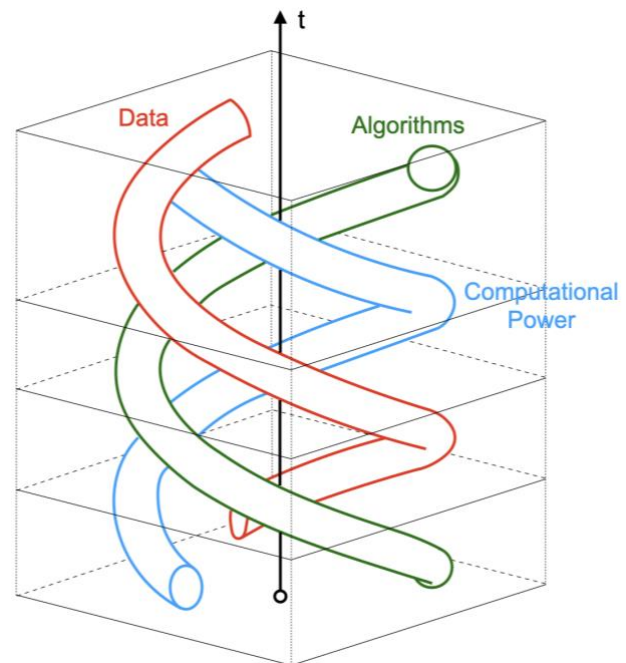


Figure 1: Triple Helix Model of AI. This is a qualitative representation of the triple helix model of AI, with the dotted box representing the AI space and the vertical axis, the time vector. (Drawing inspired by Figure 1 in Sonsilphong and Wongkasem, 2013)

The triple helix model of AI is designed to reduce the complexity and clarify the mechanisms of value generation

and capture in AI. This approach of examining each strand separately has gained increasing recognition in recent scholarship. Sastry et al. (2024) exemplify this trend in their research on AI capability development, where they identify data, algorithms, and compute as the "three key technical inputs to AI.

While this structural similarity is noteworthy, the analytical power of the triple helix model lies in its capacity to accommodate diverse forms of value that transcend purely economic or monetisation metrics. Moreover, it facilitates exploration of the multiple pathways through which different value types are perceived, interpreted, and capitalised upon by various stakeholders within the AI ecosystem.

Action

Artificial Intelligence - A Modern Approach by Stuart Russell, Peter Norvig and Ernest Davis (Russell, Norvig and Davis 2010) holds a privileged status within the AI literature. This seminal work offers an unparalleled comprehensive treatment of AI, spanning from accessible introductions to the more cumbersome philosophical or mathematical discussions around AI development. Most significantly, Russell et al. present a straightforward yet enduring taxonomy that classifies agent learning based on three types of feedback mechanisms that remain relevant today.

Unsupervised learning occurs when an agent identifies and extracts patterns from a given input without explicit feedback or guidance. Reinforcement learning occurs when an agent learns or infers a pattern based on a set of rewards provided by the environment via a reward function. Supervised learning enables agents to discern correlation patterns from input and output pairs that are given to it (Russell, Norvig and Davis 2010). Additionally, deep learning refers to the depth of the network or the number of layers it has and can both refer to supervised and unsupervised learning. Lastly, semi-supervised learning has also emerged as an effective training technique, leveraging a small subset of labelled data within a larger set of unlabeled data.

These approaches collectively constitute what is commonly understood as Machine Learning (ML). However, when the concept of AI was first proposed in Dartmouth in 1956, competing views about how intelligence should be conceptualised and artificially implemented clashed. One camp advocated for high-level logical reasoning as the key to unlocking AI, while others defended that natural intelligence should inspire the underlying computing architectures (Mitchell 2020). These opposing views were exacerbated by a pattern of boom-bust cycles derived, in part, from the competitiveness of raising funds for research—with promises that could not be met by a

technology that was volatile and immature (Mitchell 2020)—in part due to the asynchronies among AI's fundamental strands. As a result, algorithmic innovation was hindered by data scarcity and insufficient computational power. In AI's historical records, these up and down phases are known as AI 'winters' and 'springs'.

Within each learning paradigm, different algorithms serve distinct applications. Supervised learning predominantly employs classification and regression techniques, with algorithms ranging from linear regression to support vector machines or neural networks. Conversely, unsupervised learning encompasses techniques including clustering, dimensionality reduction, anomaly detection or neural network development through algorithms such as k-means, auto encoders, or generative adversarial networks.

Given this landscape, it is important to keep in mind the rigour underpinning the mathematical and computational use of the term 'algorithm' as a clearly defined process or set of rules, and in opposition to heuristics, to underscore that algorithms alone do not constitute the sole driver of value generation in AI.

Data

Algorithms—clearly defined processes or sets of rules—serve to organise and extract meaning and patterns from data which, in this context, refers to digital representations that can be processed and analysed and that convey information about the world. In the digital age, humans function analogously to cyanobacteria during 'the great oxidation event', when the production of oxygen by these microorganisms fundamentally altered Earth's atmosphere and enabled aerobic life to emerge (Aiyer 2022). Similarly, our digital presence generates vast quantities of data that can be harnessed for AI development.

These data generation processes, however, are not neutral. Suresh and Gutttag (2021) present a framework for understanding unintended consequences in machine learning by identifying various sources of bias throughout the ML pipeline, organising them into two primary stages: data generation and model building.

This framework provides researchers and practitioners with a systematic approach to identifying critical junctures in the ML pipeline where bias infiltrates the model. The first source of bias is what is commonly known as 'historical bias', or the recognition that even with perfect sampling and feature selection, data inevitably reflects existing structural imbalances in the world. Combined with representation bias (or the bias that derives from how a population is defined and sampled), these factors fundamentally determine *what* information is conveyed by the data.

Moreover, data generation processes can be endogenous or exogenous relative to a given AI. Exogenous data consists of data representations of the world (or a part of it) that are

processed by AI to create a model of it. This encompasses training, validation, and test datasets, as well as data that informs the logic of non-ML AI. These data can be considered exogenous because while they substantially influence and define the model, the model's existence remains independent of the data's existence. Endogenous data refers to data which derives from the relationship between the system and its environment—the feedback loops between the AI and its context. This data is endogenous because it both originates within and influences the network of which the AI is forms a part, informing decisions about the AI's relationship to the rest of the network.

Crucially, the relational value of data is inherent and not fully controllable by data subjects or those who collect it. Pellungrini et al. (2018) illustrate this complexity through mobility data, which simultaneously contains valuable information about human movement patterns and collective dynamics while reflecting individual sensitive information as well, the use of which can entail privacy risks for such individuals. In consequence, data not only reflects—even if only partially—existing structural inequalities but also embodies the interdependencies between individual and collective observations.

Compute

If algorithms provide the logical structure and architecture for AI, and data supplies the information and building blocks, neither of them is sufficient to produce artificial intelligence without computational power and its underlying infrastructure. Compute can be measured through various proxies (processing units, FLOP, electric consumption, CO2 emissions) but it ultimately denotes the hardware dimension of AI—the place or space in which AI is conceived, generated, and stored.

Together, algorithms, data, and compute constitute the three primary drivers of AI progress—that is, the arms of the triple helix—which are in turn embedded within a multi-faceted context encompassing political, social, economic, and cultural dimensions populated by diverse actors and stakeholders under power dynamics at various levels.

The triple helix should thus not be understood as an isolated structure but as an interconnected framework that explains the emergence of AI as the intersection of three distinct technical domains: data, algorithms, and computational power. This convergence results from diverse driving forces generated by multiple stakeholders across different contexts. Consequently, the evolution of the AI landscape is determined both by the dynamic interactions among the three helix arms and by the evolving context within which AI development occurs.

Value Beyond Monetisation

The technical complexity of developing highly functional AI architectures, combined with the substantial costs of both conceptual design and practical implementation, creates driving forces for large-scale model scarcity and concentration among a limited number of companies capable of attracting the necessary talent and capital for frontier AI. In pursuit for returns on investment, providers of General-Purpose AI (GPAI) models are exploring diverse pricing and access models.

Proposals like the “windfall clause” (O’Keefe et al. 2020) highlight the potentially unprecedented advantages that companies developing the most powerful AI products may encounter. This clause or mechanism is designed to ensure that entities developing Transformative Artificial Intelligence (or TAI, broadly understood as artificial intelligence that brings about a significant development with a great economic impact) bindingly agree to donate a considerable portion of their benefits if they achieve earnings at an unprecedented scale, such as capturing 1% of global economic output through advanced AI development (O’Keefe et al. 2020).

Regardless of the mechanisms through which AI-powered products generate monetary value, this technology is positioned to impact society far beyond economic transformation. However, mechanisms like the windfall clause overemphasise AI's economic dimensions. While this focus may be appropriate for the clause's specific purposes, it contributes to an understanding of AI-generated value—particularly from Transformative AI—as predominantly monetary.

In “The Impact of Artificial Intelligence: A Historical Perspective” (2022), Garfinkel claims that “artificial intelligence is starting to emerge as a general-purpose technology” (a class of advancements that are characterised by being unusually pervasive, triggering complementary innovations and bearing the potential for high technical improvement) (Garfinkel 2022). Some of these technologies can also be revolutionary—i.e., supporting a fundamental transformation in the nature of economic production (Ibid., p. 2).

Garfinkel evaluates AI against these definitions by comparing it to historical general-purpose technologies, including plants and animals, the wheel, three-masted sailing ships, and the invention of the computer (Lipsey et al. 2005). From this catalog, he identifies crop domestication and the invention of the steam engine as innovations that triggered fundamental changes to the nature of economic production, thereby qualifying as revolutionary. These transformations, among other things, modified the ways in which energy was captured, shaped health or gender relations, and altered war-waging and political capabilities such as collective decision-making or

power centralisation (Garfinkel 2022). Garfinkel concludes that from a historical perspective, and measured against this benchmark, AI represents a good candidate for a general-purpose technology that could potentially usher in a new revolution—though this possibility remains unrealised.

This analysis is compelling because the technological innovations listed by Richard Lipsey and referenced by Garfinkel affected not merely the economic underpinning of their respective social contexts but fundamentally transformed how humans organised themselves and, even, related to the environment. Plant and animal domestication, for example, reconstituted the relationship between humans and nature—shifting from a foraging and hunting paradigm where humans sought shelter from weather and defended against other species, to one characterised by dominance over animals, landscapes, and even natural elements. Domestication created new species that had not previously existed and rendered previously inhospitable climates habitable through technological intervention.

The value of animal and plant domestication manifested not only in new possibilities for economies of scale, trade, and other socioeconomic transformation but also reconstituted human's relationship with land—transitioning from nomadic and hunter-gathering dependency relationship to one where roots could be set for diverse forms of cultural capital to emerge.

Yet this productive organisation of the environment involves a crucial communicative dimension. In the 1940s, Harold Innis revolutionised the study of the relationship between communications and economic development, emphasising their spatial and temporal dimensions across historical contexts, from Ancient Egypt and Rome through modern imperial power (Innis, 2007). Consider the telegraph, a foundational technology of modern electronic communication, that generated substantial economic gains for inventors, commercialising companies, infrastructure builders and, ultimately, users. The compression of space through quasi-immediate electronic communication across vast distances enabled a new coordination of time: standardisation.

The invention of the mechanical clock in the fourteenth century achieved localised uniformity, enabling consistent timekeeping within specific locations, yet coordination with other locations remained virtually impossible. However, by the late nineteenth century, the telegraph had made coordinating a global network of local times feasible. From military planning to policy implementation, time standardisation provided benefits extending far beyond purely economic considerations. By 1884, standardised time was adopted globally at the International Meridian Conference, having first been proposed in Canada in 1879 by Sanford Fleming, and adopted in 1883 in the United States (Bartky 2007; Kern, 2003).

From this, two corollaries stand out. First, the telegraph enabled real-time global coordination for the first time, creating the inaugural artificially synchronous global network. Second, time and space became disentangled in ways that changed how humans experienced these physical realities: before instant communication, distances were described in units of time—e.g., 'two days ride'. The deep transformations triggered by the telegraph and, ultimately, its value, cannot be measured exclusively in terms of capital gains. From a cultural to a political or social perspective, the value of the telegraph was constructed and captured along multiple chains.

AI Complexity Through Value Diversity

Although value chains were originally conceived as tools for economic improvement that enabled business to identify the value-producing steps for competitive differentiation, the notion of a value chain can be applied more broadly to encompass diverse definitions of value. The triple helix approach facilitates this expansion by distinguishing value creation across AI's different strands (each possessing distinct supply chains and value systems), while fostering a more flexible, sociotechnical approach to value identification. This framework serves not merely as a tool for economic impact assessment but as a conceptual baseline for comprehensive approaches to value alignment.

By adopting multifaceted understanding of value that accommodates multiple and less rigid definitions, analysing AI's value chain becomes a powerful baseline for value alignment. This expanded conception encompasses not only a traditional approach to value alignment (i.e., how to align the technical object with a set of principles or expected behaviours) but also mechanisms for measuring, fostering, and restricting certain behaviours that, given their economic, social, political, or other forms of value, serve to align AI as a node within sociotechnical networks.

Identifying and understanding the economic incentives underpinning AI releases is crucial for defining alignment mechanisms required to mitigate downstream harms and address potential friction with implementation contexts. This approach thus extends the classic value alignment problem into new dimensions, transcending purely technical and normative considerations to encompass the political implications of technology development and adoption while facilitating a more comprehensive analysis of value generation to adequately assess AI's impact.

Examining the data strand's role in GPAI model development, however, allows for a provocative reflection on value measurement methodologies. Gross Domestic Product (GDP) represents one of the most widely used macro-indicators for value measurement, defined as "a monetary measure of the market value of all the final goods and services produced and sold in a specific time period by

countries.” (Wikipedia, Gross Domestic Product). However, Connor Leahy, CEO of Conjecture, proposes an illuminating thought experiment: consider two identical societies existing in parallel worlds. One has developed Wikipedia; the other has not. Since GDP measures the market value of final goods produced and sold within specific timeframes, and Wikipedia functions as a free online encyclopedia written and maintained by volunteer communities through open collaboration, both societies register identical GDP measurements (Future of Life Institute, 2023).

This scenario challenges market value as a proxy for broadly conceived value and, more specifically, for capturing value generated around GPAI and AI systems. Wikipedia contributes nothing to GDP because it operates outside market mechanisms. At a more fundamental level, the relationship between GPAI-related value and GDP is problematic because GDP is designed to capture value from specific product categories (assessed through production, income, or expenditure approaches), favouring products whose value does not rapidly change. For example, as software proliferates and becomes more affordable, its contribution to the GDP decreases, while the inflation of raw materials increases its contribution. Consequently, the value generated by AI (as software-based technology) will likely not be primarily captured through GDP or traditional value measurement approaches.

Ontology, Kinematics, Dynamics: A Method for Value Identification

A central claim of this article is that economic value (and the framing of value chains along this variable) is insufficient to understand the potential for transformation that AI poses. The triple helix and the notion of transformative technologies have provided a starting point for a broader understanding of the concept of the value chain in AI. However, the triple helix alone fails to provide a framework to understand how value generation and value capture occur in AI. In what follows, I propose a method to assess value generation and capture mechanisms based on three phases. First, an ontology, understood as a list of actors and relationships among them, capturing the spaces of opportunity for value generation. Second, a kinematic account or a historicist overview of how the actors and the relationships among them have evolved. Lastly, a dynamic account that seeks to describe the forces driving the evolutions of the diverse set of value-generating mechanisms.

An Ontology

A value-generation ontology refers to the list of elements that are necessary to produce value in AI. Wirtz et al. (2022)

conduct a comprehensive analysis of the AI ecosystem to derive an integrative framework for the governance of AI. The authors propose an economic-ecosystem understanding (which contrasts with the idea of an AI value chain) to capture all process-influencing actors, including those that would fall out of linear value creation processes. In doing so, value can be understood to arise in a network of actors participating in this process of value creation.

The triple helix model instantiates such an ecosystem, departing from a chain-like approach to value generation in AI and GPAIs more specifically through the concept of a value-generation ontology. However, and in contrast with Wirtz et al.’s understanding of AI as “the capability of a computer system to show human-like intelligent behaviour characterised by certain core competencies, including perception, understanding, action, and learning” (Wirtz and Müller 2019), the triple helix approach rests on an understanding of AI as the intersection of its three strands in a given context. This allows the identification of relevant aspects such as access to data, traceability, or legality (Wirtz, Langer and Weyerer, 2022), while capturing the complexity and dynamism of the AI ecosystem understood as the intersection of different strands and driving forces that allow for multiple forms of value to arise.

The ecosystem perspective is conceived to identify key governance challenges (into six categories identified by Wirtz) for AI systems, and it does so in a novel and promising way. But understanding AI as the intersection of the three strands that configure the triple helix and not as a set of features within technological artefacts that compose an ecosystem is what allows identifying multiple forms of value-generation and value capture.

This does not mean that, to effectively govern AI, all aspects related to AI (including big data, cloud computing, and the other enabling technologies) must be governed, but that only by virtue of a more holistic understanding of AI (as a set of complex processes which can be captured and formalised in a triple helix model) it may be effectively governed. The same holds for value identification and value capture: AI’s value cannot be understood, either in a linear chain or an ecosystem-like network, from the perspective of a system presenting certain capabilities, but as the situated product of the intersection of AI’s strands. With this, the ecosystem is naturally contained within the ontology, and complemented by neighbouring spaces such as data, computational power and algorithms unrelated to AI or more sociotechnical ones. This facilitates identifying and historically tracing how the development of AI contributes to value generation and how such value is captured and organised by different actors beyond monetisation.

Figure 2 qualitatively represents the ‘AI ontology’ (or different components of Wirtz et al.’s ecosystem) for a given moment of time ($t = x_i$). This representation is obtained by intersecting the triple helix with a given time plane. In it, the

grey area represents the context (social, political, academic, economic, etc.) in which the triple helix and its driving forces occur. Each arm of the triple helix corresponds to a different strand and is depicted with varying sizes—both relative to past and future time slices, and to each other—to capture the historical evolution of each strand over time. Within each strand, there is a subset (depicted in orange) that corresponds to the corresponding data, algorithms, and computational power that constitute AI. Hence, the data helix, for example, comprises the data used to develop AI systems but also all the other data unrelated to AI. Lastly, in the middle of each time slice there is a yellow triangle depicting the space of AI, as well as the forces affecting it (such as funding, academic resources, or policy measures among others).

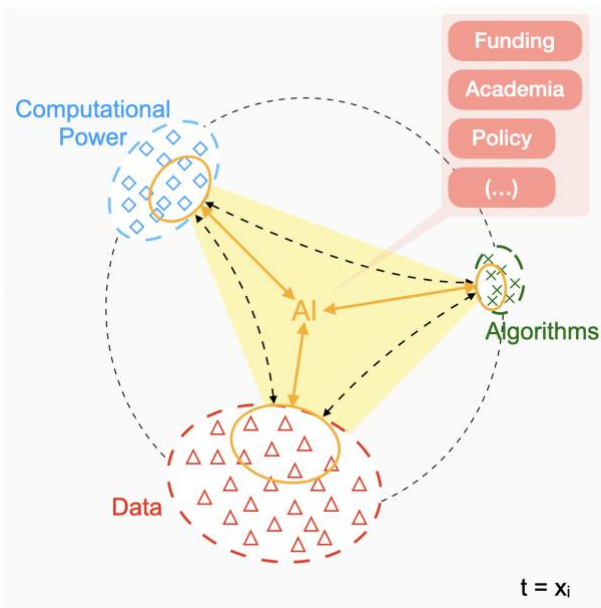


Figure 2. Time slice view of the triple helix: a value generation ontology. Time-slice of the triple helix, or an AI ontology for a given moment in time $t = x_i$. The yellow area represents the space of AI as the result of the sub-set of data, computational power and architectures dedicated to it given the context of available funding, academic interests, policy initiatives, etc. that configure and constitute the relevant stakeholders.

The value-generation ontology consists of a conceptual exercise of listing and identifying all relevant actors, relationships, and sub-spaces that compose the ‘AI ecosystem’ to grasp the complexity of value generation, governance, alignment, and so on. With this, the realisation that stakeholders beyond market dynamics are affected and are willing to effect change in how AI is defined, developed, and incorporated into society, the different ways in which

value can be generated also become more obviously apparent. An ontology allows the possibility for an ethnography of value, in which subjective and alternative forms of value generation and value capture can exist.

For some, value can be captured via exit and funding operations—see, for example, Basole (2021) analysing more than 30,000 funding and exit activities around AI. For others, the adoption of AI represents a hand-off of certain capabilities that have consequences even in how ethics is defined around AI. There are some others for whom the development of GPAIs reflects their gains in political capital to influence regulation in a given territory. Value generation mechanisms are complex, and the construction of a value-agnostic AI ontology facilitates identifying novel ways in which AI and GPAI transform the network they are part of.

A Kinematic Account

The triple helix is directional, in the sense that it represents all the ‘time-slices’ or snapshots of AI’s ontology over time. AI can thus be conceptualised via the triple helix in terms of the actors that constitute the space of AI and the relationships established among them, as well as through a historicist account of how AI’s different strands and context have evolved over time. In this sense, Figure 3 captures two key elements. On the left-hand side, the time vector is represented, to reflect how any AI ontology made at a given point in time is, in fact, a snapshot—or a building block of the kinematic account. Conversely, the centre and right-hand side diagrams show two different ways to understand the historicity of AI. Column A represents how, for different times, the ontology was evolving, whereas column B represents how a given strand has changed over time.

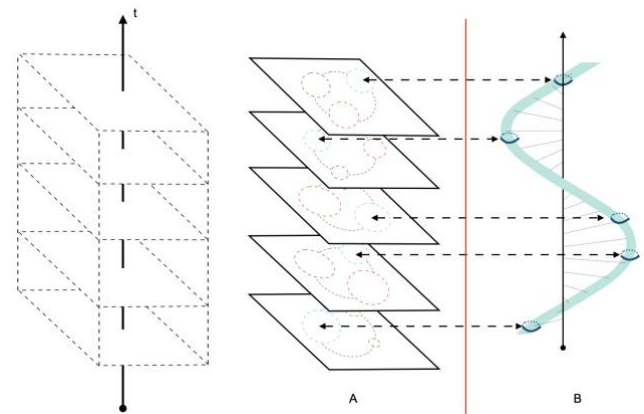


Figure 3. Kinematic account of the triple helix. The vertical axis (arrowhead) depicts the time vector. On the left-hand side, the stacked slices capture the evolution of the context. On the right-hand side, two different ways to observe the historical evolution (column A and B) are depicted: A shows

how different ontologies stack onto each other to configure a discrete history of the field, whereas B shows the continuous tracking of a single strand

This kinematic view captures both the historical baggage of each of the AI strands as well as their contextual and relative strength. Innovations (along each strand, but also revolutions, social acceptance and market success, or other relevant vectors affecting the AI space) are accounted for and reflected in the varying size for each one of the strands along the different time slices.

Within this framework, funding dynamics responding to failures and successes are captured by shrinking and expanding AI spaces respectively—with relatively weaker and stronger arms of the triple helix and varying importance of different contextual elements as their causal explanations—which are depicted by the central yellow area in Figure 2. This allows understanding the complex interactions not only between computational power, data availability and algorithmic innovation which are, indeed, key to AI progress, but also with the political landscape, economic availability or social acceptance and demand for a given technology.

The kinematic account allows explaining AI winters (or historical failures in AI development) not as the consequence of an algorithmic innovation plateau or compute or data scarcity necessarily, but as the consequence of several driving forces of AI progress, such as the social acceptability of AI in a given context (either due to lack of demand within a market of products or due to explicit rejection of certain innovations for ethical or safety reasons).

Moreover, the kinematic account affords a contextual and historically informed perspective of AI progress, whereby the relative importance of different factors can be identified (e.g., computational power or capital interest) throughout AI progress. But while the actors and relationships are captured by the ontology and the history of their effects are reflected in the kinematic account, it is in the dynamic account that the overall effects of the actors over AI progress can be observed.

A Dynamic Account

The kinematic account characterises the interaction and coordination mechanisms between AI's strands and their historical context, putting into perspective some of the ex-post effects of each gradual step towards AI progress. This progress, however, is driven by a set of forces that constitute the dynamics behind algorithmic innovation, product development (Basole 2021), computational capacity development, or even policy windows to publicly discuss the need to regulate or incorporate AI into different domains

of life. Considering this, the dynamic account can be understood from two different perspectives.

On the one hand, and by analysing how different elements of the value-generation ontology evolve over time—such as scaling laws for algorithmic and model performance (Kaplan et al. 2020), as well as for data (Villalobos et al. 2022), Moore's law on the evolution of transistor density, regulatory pressure, or funding interest—the dynamic account allows posing hypothesis about the key drivers of AI progress. On the other hand, observing how different driving forces have impacted AI progress can serve as a prescriptive tool to define effective governance and alignment mechanisms.

The triple helix model affords a tentative overview of the causal effects and balance of power behind AI progress, allowing for more detailed views of specific spaces (such as the geopolitical tensions around the semiconductor supply chain, or power and raw material scarcity directly affecting compute hardware availability and geopolitical stances of corporations and countries with vested interests) (CSET 2021).

Through the composition of a value-generation ontology, a kinematic description of its evolution throughout time and the posing of hypotheses about the driving forces, the roles of different actors and contexts in AI development and adoption are surfaced, allowing in turn the breakdown and allocation of accountability throughout the ecosystem.

Acknowledgments

I thank Prof. Mark Coté, Dr. Ana Valdivia, and Dr. Roberto Pellungri for their valuable feedback on an earlier version of this draft, which forms part of my doctoral dissertation.

References

- Aiyer, K., 2022. The Great Oxidation Event: How Cyanobacteria Changed Life. American Society for Microbiology. <https://asm.org/Articles/2022/February/The-Great-Oxidation-Event-How-Cyanobacteria-Change> Accessed: 2024-11-13
- Bartky, I.R. 2007. *One time fits all: the campaigns for Global Uniformity*. Stanford University Press.
- Basole, R. 2021. *Visualizing the Evolution of the AI Ecosystem*. Center for Security and Emerging Technology et al. 2021. *The Semiconductor Supply Chain: Assessing National Competitiveness*. Center for Security and Emerging Technology.
- Engler, A. C., & Renda, A., 2022. Reconciling the AI Value Chain with the EU's Artificial Intelligence Act. CEPS.
- Future of Life Institute., 2023. *Connor Leahy on AI Progress, Chimps, Memes, and Markets* [Video]. YouTube. <https://www.youtube.com/watch?v=2RjuJzmafAA&t=3232s>
- Garfinkel, B. 2022. The Impact of Artificial Intelligence: A Historical Perspective. *The Oxford Handbook of AI Governance*. J.B. Bullock et al., eds. Oxford University Press. 86–108.

- Gutierrez, C.I. et al. 2023. A Proposal for a Definition of General Purpose Artificial Intelligence Systems. *Digital Society*. 2, 3 (2023), 36. DOI:<https://doi.org/10.1007/s44206-023-00068-w>.
- Hernández-Orallo, J. et al. 2021. General intelligence disentangled via a generality metric for natural and artificial intelligence. *Scientific Reports*. 11, 1 (Nov. 2021), 22822. DOI:<https://doi.org/10.1038/s41598-021-01997-7>.
- Innis, H.A. et al. 2022. *Empire and communications*. University of Toronto Press.
- Kaplan, J. et al. 2020. Scaling Laws for Neural Language Models. arXiv.
- Kern, S. 2003. *The culture of time and space, 1880-1918: with a new preface*. Harvard University Press.
- Lipsey, R. 2006. *Economic Transformations: General Purpose Technologies and Long-term Economic Growth*. Oxford University Press, Incorporated.
- Microsoft., 2023. Microsoft and OpenAI extend partnership. *Microsoft Corporate Blogs*. <https://blogs.microsoft.com/blog/2023/01/23/microsoftandopenaiextendpartnership>
- Mitchell, M. 2020. *Artificial intelligence: a guide for thinking humans*. Pelican, an imprint of Penguin Books.
- Muñoz, C., 2022. Responsible AI Licenses: a practical tool for implementing the OECD Principles for Trustworthy AI. *The AI Wonk*. OECD. <https://oecd.ai/en/wonk/rails-licenses-trustworthy-ai>
- OECD 2013. *Interconnected Economies: Benefiting from Global Value Chains*. OECD.
- O’Keefe, C. et al. 2020. The Windfall Clause: Distributing the Benefits of AI for the Common Good. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (New York NY USA, Feb. 2020), 327–331.
- OpenAI., 2023. OpenAI and Microsoft extend partnership. *OpenAI*. <https://openai.com/index/openai-and-microsoft-extend-partnership>
- Pellungrini, R. et al. 2018. Analyzing Privacy Risk in Human Mobility Data. *Software Technologies: Applications and Foundations*. M. Mazzara et al., eds. Springer International Publishing, 114–129.
- Porter, M. E. 2011. *Competitive advantage of nations: creating and sustaining superior performance*. Simon and schuster (1985).
- Rother, M. and Shook, J. 2018. *Learning to see: value-stream mapping to create value and eliminate muda*. Lean Enterprise Inst.
- Regulation 2024/1689. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689
- Russell, S.J., Norvig, P., and Davis, E., 2010. *Artificial intelligence: a modern approach*. Prentice Hall.
- Sastry, G. et al. 2024. Computing Power and the Governance of Artificial Intelligence. arXiv.
- Sonsilphong, A. and Wongkasem, N. 2013. Low Loss Circular Birefringence in Artificial Triple Helices. *Progress In Electromagnetics Research M*. 29, (2013), 267–278. DOI:<https://doi.org/10.2528/PIERM13012510>.
- Suresh, H. and Gutttag, J.V. 2021. A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. arXiv.
- Villalobos, P., Ho, A., Sevilla, J., Besiroglu, T., Heim, L., Hobbhahn, M., 2024. Will we run out of data? Limits of LLM scaling based on human-generated data. arXiv.
- Wikipedia. 2025. Gross Domestic Product. https://en.wikipedia.org/wiki/Gross_domestic_product. Accessed: 2025-01-22.
- Wirtz, B.W, Langer F.L., and Weyerer, J.C., 2022. An Ecosystem Framework of AI Governance. *The Oxford Handbook of AI Governance*. J.B. Bullock et al., eds. Oxford University Press. 398–420.
- Wirtz, B.W. and Müller, W.M. 2019. An integrated artificial intelligence framework for public management. *Public Management Review*. 21, 7 (Jul. 2019), 1076–1100. DOI:<https://doi.org/10.1080/14719037.2018.1549268>.