



Traversing Technology Trajectories

Frederick Klaessig 

Received: 9 November 2020 / Accepted: 30 July 2021 / Published online: 28 August 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract Scholars in science and technology studies, as well as economics and innovation studies, utilize the trajectory metaphor in describing a technology's maturation. Impetus and purpose may differ, but the trajectory serves as a shared tool for assessing social change either in society at large or within a market sector, a firm, or a discipline. In reverse, the lens of a technology trajectory can be a basis for assessing technology, estimating economic growth, and selecting among plausible product development pathways. Emerging technologies pose a challenge in that the promissory claims that attracted investment may also challenge social safeguards, leading to necessary, but unforeseen hurdles to later technology acceptance. The experiences with biotechnology and nanotechnology demonstrate that knowledge gained in the intervening years was to have tempered exaggerations and addressed the criteria and the responsibilities of those who would later be deciding on product acceptability. In particular, the promissory claim for nanotechnology, unique phenomena, challenged regulatory practice, while claims for biotechnology products were often the fruition of themes found in the educational experience common to both the biotechnology researcher and the regulator. The social science and humanities literature mirrored these points as did the responses of civil society

organizations. A prospective trajectory for artificial intelligence is developed using the example of driverless cars and then qualitatively compared to the retrospective trajectories for nanotechnology and biotechnology. Experience with automation, algorithm transparency, and computational modeling of biological mechanisms are identified as pertinent to responsible AI development.

Keywords Technology trajectory · Promissory claim · Hype · Regulatory receptivity · Nanotechnology · Biotechnology · Artificial intelligence

Introduction

Biotechnology, nanotechnology, and artificial intelligence (AI) are examples of emerging technologies, a phrase with a plain reading of there being a capability that is new and its full extent unknown. In this paper, the concept of technology trajectory that has figured prominently in innovation studies, economics, and science and technology studies (STS) serves as a framework for identifying critical points that lie along each technology's developmental pathway. The more recent experience with nanotechnology will be examined, validated against earlier events with biotechnology, and then applied to AI. Elapsed time becomes a significant factor in that the knowledge base, practice, and responsibilities of those who decide on the

F. Klaessig (✉)
Pennsylvania Bio Nano, LLC, Doylestown, PA, USA
e-mail: fred.klaessig@verizon.net

technology-as-product differ from those who initiated the technology-as-a-research-program.

The concept of technology trajectory finds expression through a number of phrases: technology dynamics [1], research trajectory [2], technological trajectory [3, 4], technological evolution and technological paradigm [5], trajectory of improvements [6], trajectory [7], and product/technological life cycle [8]. Each field uses the concept for its respective narrative; i.e., business strategy for innovation studies; Schumpeterian creative destruction for economics; and the social construction of technology for STS, but do so without offering a formal definition for trajectory or identifying its intermediate stages. As such, the term is malleable and can be viewed as a useful tool for visualizing innovation as a generic process, as a model that is separate from product-level details (see Fig. 1). There are other terms such as dominant design [8], technological regime [1, 9], regime [1], and technological paradigm [5]. When considered by the same author [1, 2, 5], paradigm and regimes have evolutionary implications as one technology succeeds another, while a trajectory is a “pattern of ‘normal’ problem solving” occurring within the paradigm [5]. A further refinement would be to view the technology trajectory as the aggregate of individual product actions, recognizing that regulatory actions generally involve products and their uses.

These fields treat innovation as being separate from invention (i.e., creativity), which clearly establishes that a trajectory has a commercial intent [10] that may range from introducing technology improvements to displacing current technology and, finally, to establishing new capabilities. Nelson and

Winter (9, p. 258) capture the dynamic very well when describing a technology as a frontier of available capabilities that evolve as the physical and biological constraints are tested either in existing markets or in creating new market opportunities. The several fields differ in their emphasis with economists stressing the allocation of resources, primarily financial, while STS scholars focus on societal influences and expectations. For Geels, the STS variant of technology trajectory is a middle-range theory [1], which though at odds with Merton’s advice that a theory not be based on “observed uniformities” [11], nevertheless does express the STS sense of an overarching pathway shared by many technologies [12]. In innovation studies, colleagues follow the technology in the form of a commercial firm with a new product or as a collection of firms establishing a new market [13]. In sum, the technology trajectory is an organizing principle, a scaffold, and a shared tool that these fields use to characterize an innovation, to influence its direction, and to map its implications for social change and economic development. With an emerging technology, the several fields have a shared focal point but will differ regarding the actors, their roles, responsibilities, and values.

In order to compare the three technologies that are our subject, it is important to recognize the investment decisions that occur when transitioning from exploring a technology’s promise (invention, basic research) to pursuing tangible product forms (innovation, applied research). These decisions are justified by promissory claims intended to assure those present at the time that success is plausible. The phrase promissory claim is used in this paper as a more neutral version of hype. Others in STS have used it [14, 15], and there are similar terms such as a promissory note in philosophy [16–18] and also promissory research, promissory statements, promissory commitments, and promissory economy [19–22]. Each is a claim of future success if given resources; the promissory note is a financial instrument, the promise of future payment, while in the mining of minerals a claim is a physical location that has demonstrated sufficient promise to warrant further digging [23]. The same span of concepts is true for those seeking investment in a technology. Some exaggeration at the outset is allowable as long as there is a commitment to obtaining the evidence and a recognition

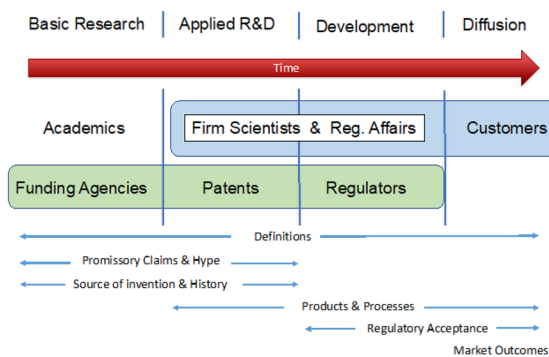


Fig. 1 Actors and roles in traversing a technology trajectory

that there are discrete time points when that evidence is required, e.g., patent examiners during applied research, regulators during development, and customers during product diffusion (see Fig. 1). There is a challenge throughout of disentangling the scientific hypothesis from the promissory claim, or as expressed by Nickles [16, p. 14], “the method of hypothesis, rather like the old method of analysis in Greek mathematics, lets scientists use the assumed hypothesis in logical reasoning as if it were already established.... For a hypothesis is really only a loan or promissory note that must be repaid through confirmation, else the enterprise fails. Whereas a conclusion derived from established knowledge can be detached and asserted, inferences involving hypothetical premises (or, indeed, any fallible claim) must be remembered and tracked. In this respect, hypotheses are like lies!” The difference between a scientific hypothesis and a promissory claim is a measure of the knowledge gained in the time period between the investment decision and its market realization.

The experiences in biotechnology and nanotechnology indicate that there can be occasions along the trajectory, where the breadth of an early promissory claim causes those at a later stage to demand an unforeseen level of evidence. To explore these issues, a trajectory can also be viewed as tracking the interplay of the following: the technology’s accepted definition, its impetus expressed as sources of invention and history, the promissory claims, the resulting palette of products or processes, regulatory receptivity, and final market outcomes. These elements are illustrated in Fig. 1 using the linear model’s standard categories to capture the overlap between stages and terms. The actors, their affiliations (light blue for commercial; light green for governance), their roles, and responsibilities differ over time, which is especially significant when considering governmental actors who were technology proponents at the outset, but are regulators with statutory responsibilities at later stages. Initially, however, there is not just one path; rather, there is a set of potential paths reflecting technology specifics, financial resources, firm strategy, and market acceptance that, when aggregated, describe the technology trajectory. The path actually taken is a response to the landscape and can be tracked

Table 1 A comparison of AI, Biotechnology and Nanotechnology Trajectories in Relation to Fig. 1

AI trajectory	Biotech	Nanotech
Definition	Like	Like
Promissory claims	Unlike	Like
Sources of invention and history	Like	Unlike
Evaluated as a process/technique	Like	Unlike
Evaluated as products	Like	Like
Regulatory receptivity	Unlike	Unlike
Creating new markets	Unlike	Like

as products and processes becoming more tangible as they encounter regulators’ and customers’ judgments.

One aim of this paper is to compare three technologies using the trajectory concept and a second aim is to draw conclusions from the comparisons, especially to bridge the experiences gained with nanotechnology to those that might be anticipated for AI. At the suggestion of one anonymous reviewer, Table 1 is offered as a navigation tool. Originally in the concluding remarks, it summarizes in a qualitative manner how these technologies might compare using the elements in Fig. 1. There is a projected AI trajectory that is developed in the paper’s later sections, which will refer back to this table. For each element, the projected AI trajectory is expressed as either “like” or “unlike” the experience with biotechnology or nanotechnology. The actors and roles in Fig. 1, however, change as the technology takes on a tangible product form (when traversing the trajectory) such that those making promissory claims are not those deciding on regulatory acceptance. The commercial firm has the greatest influence as its actors are present for much of the trajectory. Governmental actors change as one moves from funding to patenting to regulating. Definitions are the one element spanning all actors and stages, which is why it is emphasized when examining each technology’s narrative.

This work reflects participation in US activities including public nanotechnology meetings and personal discussions at NSF-funded centers and standards committees. These are occasionally discussed in footnotes. Though biotechnology was chronologically first; the experiences surrounding nanotechnology are more recent and provide a fuller overview

of the interactions outlined in Fig. 1. In addition, the prominence that ethical, legal, and social implications (ELSI) research gained in the debates surrounding biotechnology led to a larger, more structured program in the US nanotechnology program. For these reasons, nanotechnology is discussed in more detail leading to analysis that is validated using biotechnology and then extended to artificial intelligence.

Return on Investment and Hype

The word trajectory implies that there is an entity having an impetus, following a path and arriving at a target. The landscape between impetus and outcome can be used to differentiate among the fields that use the trajectory metaphor. In a market economy, profit is the primary impetus underlying both innovation studies and economics, which tend to differ when drawing conclusions or taking actions to influence the trajectory's eventual outcome, i.e., competitive advantage and firm growth for innovation studies and productivity and GDP growth for economics. Where innovation studies and economics investigators examine all forms of products, services, and technologies that might enter the market, social science, and humanities scholars (SSH) tend to focus on facets that embody societal aspirations such as grand challenges, policy objectives, and transnational governance, e.g., sustainability.

Those in SSH, therefore, have a partial view of the many paths a technology and its products might take when traversing a trajectory. For example, the STS field is generally disconnected from the innovation studies and economics literature [13, 24], which is surprising considering that innovation studies, in particular, had its origins in the SSH literature on cultural change [10, Chapters 1 and 2]. Perhaps this is to be expected as the SSH commitment is to understanding innovation for its effects on those societal aspirations. SSH practice is, therefore, less interventionist at the firm level than it is for innovation studies and economics. SSH colleagues, particularly in STS, have therefore focused on influencing the trajectory through innovation governance, e.g., responsible research and innovation (RRI) that finds particular expression for nanotechnology in the concept of safe(r)-by-design [25–27]. A gap nevertheless appears when RRI is limited to commercial firms and

not applied to the funding agencies or to the academic scientists who may have been the first to decide that an emerging technology will meet a societal goal.

This gap widens when the hype cycle starts with a government initiative [12] and then passes through a progression of firm-level and market-level promissory claims. The promissory claims are initially weighted towards scientific hypotheses that are then combined with return on investment (ROI) calculations, thereby contributing to that perception of commercial inevitability that lies at the core of the Collingridge dilemma (which basically says that when change is easy, the need for it cannot be foreseen, and when the need for change is apparent, change has become expensive, difficult, and time-consuming [28]). The more significant distinction, therefore, lies between those promissory claims translatable into marketplace supply and demand analyses (technology push, market pull, risk–benefit ratio, cost avoidance) and those that are aspirational (justice, equity, public health, climate, and sustainability). Each has a separate purpose, one societal investment, and one a financial investment, as well as audience, one political in nature, and the other economic. The efforts of SSH scholars, Civil Society Organizations (CSO), and others enriching public engagement are impaired in their pursuit of aspirational goals if they become mired in exaggerated promissory claims that were designed to attract commercial investment. Their efforts will be misdirected if they expend their limited resources on issues occurring early in the trajectory (known as upstream engagement) should other intervention points be more favorable to their goals. Anticipating those circumstances, the technology trajectory can be a roadmap for translating past experiences with one technology into future actions with a second (Table 1). These aspects are captured in the SSH discussions that follow each historical recounting of promissory claims in nanotechnology and biotechnology.

Nanotechnology: Definitions as History and in Modulating Promissory Claims

The promissory claim can be found in definitions or in the description of expected outcomes. In the case of nanotechnology, the claim was initially found in the definition and underwent later adjustments that

align well with the STS concept of the social construction of technology [1]. Definitions for nanotechnology and the closely related terms of nanoscale and nanoparticle were proposed during the formation of the National Nanotechnology Initiative (NNI) circa 2000 and underwent change when enacted into a law, the 21st Century Nanotechnology Research and Development Act [29]. There are variants. The organizations and agencies generating those variants have presumably decided that earlier versions were incomplete, perhaps even incompatible or misleading when viewed from that organization's or that agency's perspective. Collectively, the resulting adjustments act upon the technology, especially if offered by an organization or agency situated along the trajectory in Fig. 1. Recounting the history of nanotechnology definitions, therefore, identifies both the trajectory's actors and their technical contributions.

The core concepts underlying the NNI's definitions are "matter" with a specific size ("roughly" 1 to 100 nm in 2006, becoming approximately 1 to 100 nm afterward) leading to "unique" phenomena that enable "novel" applications. Whether "matter" is atomic, molecular, or macromolecular or whether the state of "matter" is a gas, liquid, or solid is not explained, nor is guidance offered for identifying those properties that might emerge or might change or if those changes are gradual or abrupt. The 21st Century Nanotechnology Research and Development Act [29] does not mention the terms unique, novel, size, nor define the nanoscale. The act creates a coordinating structure for the NNI to which is added some economic and aspirational concerns that are incorporated into funding guidelines. The NNI's activities,

therefore, express a funding strategy that relies on later research outcomes for justification.

Efforts at two standards developing organizations,¹ ASTM International's E56 and the International Organization for Standards' (ISO) TC-229 committees, faltered over "unique phenomena" and resulting "novel properties," yielding definitions that emphasize objects in the size range of "approximately 1 to 100 nm." The E56 and TC-229 memberships included academics, regulators, lawyers, and industry scientists, the latter interested not only in the commercialization of new materials, but also defensive about the status of existing commercial products, and the policies in force when "ultrafine" materials had been measured in millimicrons, not nanometers. Unsuspected and unspecified "unique phenomena" and "novel properties" might have biological implications and might cast doubt on the toxicological test methods used in past regulatory actions. In parallel and with input from the Joint Research Centre, the European Commission (EC) offered an interim definition for materials in a size range of "1 to 100 nm," and replacing "approximately" with a criterion regarding particle size distribution. Overall, the important nanoparticle characteristics for regulation and toxicological testing became composition, size, shape, and surface chemistry [30].

The environmental health and safety (EHS) concerns and increased funding found in the 21st Century Nanotechnology Research and Development Act arose from a "wave of concern" that "arrived only later in 2002–2003 when industrial participation has increased" [31]. The National Science Foundation's first EHS-oriented center at Rice University was followed by two Centers for the Environmental Implications of Nanotechnology, one headquartered at Duke University (CEINT) and one at the University of California, Los Angeles (UC-CEIN).² Clearly, this late awakening to EHS issues also meant that the NNI leadership had not been fully aware of the difficulties their promissory claims and definitions posed to the Environmental Protection Agency's (EPA) procedures for evaluating a new chemical substance. In order to respond within the 90-day period mandated by the Toxic Substances and Control Act (TSCA),

¹ These are organizations with the purpose of developing consensus-based documentary standards regarding terminology, test methods, and administrative processes that may be used on a voluntary basis. Each standard developing organization has membership rules, voting procedures, and formalities designed to achieve consensus. ASTM Int. committee members are individuals voting as firms, non-governmental agencies, governmental agencies, and trade associations who are categorized as producers, users, consumers, or general interest; all negative votes must be addressed, and steps are taken that voting is not dominated by producers or users. ISO committee members are individuals representing a similar range of institutions, but voting as a national body (one country, one vote) with the required majority depending on the type of standard (Technical Report, Technical Standard, and International Standard).

² It should be noted that I was a UC-CEIN consultant on stakeholder involvement.

the EPA's practice is to augment any submitted (eco) toxicity data with results from analogous compounds already in commerce. The claim of "fundamentally new molecular organization, properties, and functions" [29] created a dilemma in that the regulators could not know if the "new" chemical substance was fully characterized using standard tests or if an analogous chemical substance even existed. The unidentified "unique" phenomenon could be toxicity or might significantly alter the interpretation of toxicity testing for both the new substance and the analog. Further, the term "approximately" was not legally defensible, as had been noted in the EC's interim definition.³

The EPA eventually decided that it did not have the authority under TSCA to view size as determinative, and therefore, had no basis for re-examining the TSCA listing of carbon black, metals (e.g., silver), and metal oxides (e.g., synthetic amorphous silica). In view of the EHS concerns, the EPA resorted to an array of administrative tools that can be illustrated with carbon nanotubes (CNT): CNTs were determined to be a new allotrope of carbon; each CNT was itself viewed as a unique chemical substance thus requiring each firm to provide EHS data before CNTs could be viewed as a recognized class; and consent orders stipulating personal protection equipment, environmental releases, and production limits were used when permitting limited marketing. Alternatively, the firms might obtain a low release and low exposure exemption by demonstrating that their manufacturing process did not lead to worker or environmental exposure. In effect, the EPA substituted exposure criteria for incomplete definitions.

Validating promissory claims is not uncommon for the Food and Drug Administration (FDA). The first drug to target a disease or a different enzyme associated with a disease and its mode of action (the mechanism) represents a promissory claim requiring FDA clearance in order to pursue clinical trials. The

FDA imposes a trajectory onto the drug's evaluation by insisting on staged clinical trials (phase I, phase II, and phase III) that incorporate FDA findings into the trial's design. One can argue that the FDA's product is the set of instructions that accompany a drug and that inform the patient about proper dosage and possible side effects. In the case of nanotechnology, the FDA reports [32] that it received 359 submissions for drug products between 1970 and 2015: 234 as an investigational new drug (IND), 62 as a new drug application (NDA), and 63 as an abbreviated new drug application (ANDA). Of the 234 INDs, 15% received approvals at the NDA stage, which is a slightly higher rate than for drug formulations without nanomaterials. As an ANDA represents the transition from a proprietary (and usually patented) drug to a generic, the FDA's activity demonstrates that there has been a commercial success without significant adverse effects being reported. It is noteworthy that the FDA's experience with nanomaterial-containing drug formulations led them to hold a different view from the NNI regarding particle size and to take steps to clarify the type of material (the "matter" in the NNI definition). The FDA guidance document incorporates a size range up to 1000 nm and exempts proteins, cells, viruses, nucleic acids, or other biological materials [33].

Caution should be taken when describing EPA or FDA actions as the two agencies respond to statutory language from different Congressional acts. For example, the FDA can limit changes in the drug manufacturing processes, but not the EPA. The agencies also differ in terms of accepted test methods or interpretations of test results. However, when viewed from a governance perspective, the regulatory gaps [34, 35] for nanomaterials are comparable in degree to gaps for other chemical substances. As will be discussed with biotechnology, the nature of FDA's methodology has been more accommodating to the NNI's promissory claims with the *proviso* that the FDA's experience leads it to differ with the NNI's definition of size.

Events in Europe paralleled those in the USA. During its implementation of the Registration, Evaluation, Authorisation of Chemicals (REACH) regulation, the European Chemicals Agency (ECHA) incorporated the Commission's definition in its proposal for the term nanoform, which distinguished particles by composition, size, shape, and surface coating. Two industry associations objected, leading to a judicial

³ It should be noted that I represented my employer on the American Chemistry Council's Nanotechnology Panel that industry raised EHS concerns with the NNI and the EPA. As a member of ISO TC 229's terminology working group, I attended an ISO meeting with senior EPA and NNI representatives present at which the EPA colleagues specifically asked: why is there an "approximately" in the ISO definition of nanomaterial? And why is there no mention of 'unique' phenomena?

review deciding in industry's favor on procedures. Eventually, ECHA used a more involved procedure in revisiting the topic, and nanofirm is now one basis for grouping nanomaterials. ECHA and the European Medicines Agency (EMA) have not documented their actions at the product level to the same extent as the EPA and FDA.

In summary, the nanotechnology example demonstrates that firms, organizations, and agencies located at the development and diffusion stages in Fig. 1 acted as counterweights to the NNI's promissory claims. The focus on size and novel properties was ultimately supplemented by additional criteria involving manufacturing and product use (EPA), was modified by dropping "unique" and "novel" (ASTM and ISO and EC), and was revised by raising the size boundaries (FDA). The plausible outcomes that the technology trajectory might exert on definitions had been overlooked, leading to mid-course corrections at the local level (country, agency, firm, product).

Nanotechnology and the Origin of Hype

Time is an essential factor in Fig. 1, but one that is difficult to portray fully. For example, the promissory claims that attracted investment will over time have become tangible products that a firm's R&D and regulatory affairs staffs present to the regulator. The potentially "unique" phenomena will have become particles with known compositions and specific functional properties. For the tangible product, technology time has stopped, and time-to-market has begun. Beyond the firm, though, the promissory claims may persist. Essentially, the promissory claims once thought reasonable when justifying a large, government initiative are either validated by the funded research or retrospectively considered mistaken, exaggerated, speculative, and even hype. In the case of nanotechnology in the USA, the NNI's initial promissory claims changed shortly after it started, a change that heightened later regulatory and societal skepticism about safety and effectively becoming a form of hype.

The NNI's origin can be traced to the National Science Foundation's (NSF) commitment to materials science and engineering research as a means for promoting the USA's economy [31, 36, 37]. Turning "[a]n orchestrated effort to assemble fragmented

disciplinary contributions" into a more cohesive one was communicated through workshops, interagency committees, strategic proposals, and White House presentations [31]. Dr. Mihail Roco was prominent throughout this time period, eventually becoming Senior Advisor for Nanotechnology at the NSF and Chair of the US National Science and Technology Council's subcommittee on Nanoscale Science, Engineering, and Technology. In recounting the NNI's origin at an August 2019 NNI stakeholder meeting [37], Dr. Roco spoke of being given an administrative challenge in 1999 of having five or six government agencies designate nanotechnology as a "top" priority and of having to do so within five months. He and his associates used a "method" of approaching the "right level" at each agency (Under Secretaries or Chief Science Officers) offering a "surprise," something to "spark the imagination," noting that "[i]f it's an improvement people are not interested." The claim of 30% lighter rockets to the National Aeronautics and Space Administration or of changed metabolism to the National Institute of Health was based on "preliminary material" that was a "seed activity in each agency." In effect, he and his colleagues recognized that their claims were predictions that required the level of testing that only funding would allow and also accepted that their counterparts at other agencies were exercising a form of peer review. The overall process led to the initial promissory claims found in the NNI's introductory documents [38] that were primarily functional and assumed that "major industrial markets are not yet established" [38, p. 27].

After the NNI had begun, there were NSF-funded workshops [39] that pursued concepts such as nanobio-info-cogno (NIBC) convergence and the more speculative claims of "enhancing human intelligence" and "developing artificial intelligence which exceed human capacity" that were to be enshrined in the 21st Century Nanotechnology Research and Development Act [29, Sect. 5(2)]. Additionally, Dr. Roco's public presentations increasingly included marketplace estimates (\$1 trillion for 2015 with 2 million nanotech workers was projected in 2003). The sources of these claims were vague and appear to stem from what one STS scholar describes as the NSF's institutional attachment to "frontier rhetoric" [36] and another viewed as responding to the recurring crises for the NSF's "foundational premise" of basic research in material science and engineering [22]. However,

these claims were well beyond the NNI's capabilities in terms of its administrative authority, budgets, and operational time scales and, therefore, fostered a misleading set of expectations that contributed to regulatory skepticism. In essence, having a "method" to spark imagination with technically oriented colleagues who understood the purpose (budgets) was a basis for acceptable exaggeration in that these colleagues also understood the trajectories that would follow. Those colleagues were also well aware of the non-economic, societal aspirations associated with their agencies (health, defense, safety). Colleagues at an NSF workshop, on the other hand, are potentially the recipients of the funding needed to validate the promissory claims being discussed. Extending the initial promissory claims through workshops and Congressional action contributed to a hyperbole injurious to later adoption.

Coursing throughout the 21st Century Nanotechnology Research and Development Act is the NNI's role in promoting the transfer of this technology to industry [29, Sect. 2(6)]. As noted above, the NNI Implementation Plan assumed that "major industrial markets are not yet established [38, p. 27], even when acknowledging an existing \$34 billion catalyst market and a further \$34 billion market for giant magnetoresistance memory in the computer industry. Overlooked were commercial products such as carbon black, a legacy nanomaterial from the 1930s with a yearly global production volume of 10,000,000 mt primarily for the tire industry. This can be contrasted to the rather low volumes for products frequently mentioned in NNI literature: nano-ceria with < 1000 mt per annum, CNTs at > 250 mt, and nano-silver at > 70 mt [40, Supplementary Information]. Also in its marketplace arguments, NNI literature anticipates that several decades may pass before a new material becomes significant, an argument that does not apply to CNTs. The first CNT patent was granted in 1987 (referred to then as a carbon fibril [41]), and the first patent using the term carbon nanotube was granted in 1994 [42]. These points became significant when scientists from those overlooked industries participated in ASTM International or ISO working groups or when their product's TSCA registrations were questioned at the EPA. Presumably, the NNI's exaggerated marketing claims can be ascribed to the NSF's untempered "frontier mentality."

In summary, the initial promissory claims of the NNI's implementation plan were appropriate to a budgetary initiative involving a collaboration among colleagues whose budgets might be affected. The later claims combined an incomplete definition with the NSF's institutional attachment to economics whenever justifying materials science research. Societal aspirations became secondary, and the perception created that nanotechnology is inevitable, leading to the urgency found in the SSH literature.

Nanotechnology: Social Science and Humanities Response

The foundation of effective public engagement is a public informed by the advocacy from civil society organizations and by the insights of a vibrant SSH literature. Engagement is difficult to achieve without some means of interaction along the technology trajectory. As noted before, the fields of innovation studies and economics have an a priori purpose of promoting a viable innovation environment and have access as consultants to the governmental development agencies and commercial firms found in Fig. 1. The SSH community's experience, on the other hand, is not as interventionist, striving primarily to understand those aspects of innovation that might affect societal aspirations. Elements of the trajectory become opaque, especially at the product level when interactions between firms and agencies are confidential, validating the commentary earlier that the SSH literature has a partial view of the trajectory.

The NNI's strategy of funding ethics and legal and social implications (ELSI) research illustrates this assessment while also highlighting the implications arising from the NNI's stated purpose: ELSI funding was intended to "help us identify potential problems and teach us how to intervene efficiently in the future on measures that may need to be taken" [38]. There were three NSF-funded ELSI centers, one headquartered at the University of South Carolina, one at Arizona State University, and one at the University of California, Santa Barbara (the latter associated with the EHS-focused UC-CEIN). There were also many individual ELSI grants. The 2003 Act specifically mentions ELSI issues [29; Sect. 2(10)] surrounding "enhancing human intelligence" and "developing artificial intelligence which exceeds human capacity."

These objectives are repeated in the Act's Sect. 5(c) on responsible development with an additional one of "self-replicating nanoscale machines and devices." Essentially, the ELSI program represented a promissory claim about guiding the outcomes of the physical science research.

Fisher recently published a thematic overview on the changing role that the NNI and Congressional policymakers were expecting of SSH scholars [43]. In his view, the 21st Century Nanotechnology Research and Development Act of 2003 marked a transition from an ELSI grounded in the Human Genome initiative over to one with a more active stance for "shaping trajectories and, by extension, societal outcomes." The Act's language is "integrating research on societal, ethical, and environmental concerns with nanotechnology research and development, and ensuring that advances in nanotechnology bring about improvements in quality of life for all Americans" [43]. Clearly, the boundaries of SSH involvement were under discussion, and the debate reflected some dissatisfaction with the conventional ELSI interventions aimed at influencing funding and regulatory policies (upstream in Fisher's parlance) and risk communication and public education programs (downstream). For Fisher, "they [sociotechnical integration] explicitly target routine R&D activities, which have traditionally been shielded from both external influences and internal value reflections, in relation to the governance of science and technology in society." The article is also an example of SSH scholars organizing their analysis around the term governance,

In arguing that sociotechnical integration itself should become policy, there is an acknowledgment that it is an underutilized means of SSH inquiry, perhaps due to the infrequent funding or due to the hurdles that only a policy mandate could remove. Placing the SSH discussion in the context of a governance mechanism indicates that there is some unease regarding the ability of conventional ELSI practice to recognize the societal implications of ongoing research, and, further, that this unease warrants more oversight of the scientists in upstream laboratories. It is less clear if the purpose is simply to gain insight or to guide the laboratory research program towards preferred outcomes. Not explicitly considered are the other locations along the trajectory where sociotechnical integration might clarify more fully the

social and economic forces acting upon an emerging technology.

Contrasting examples from the SSH literature are offered to highlight the value additional sites might bring, especially those closer to market introduction. Regulatory unpreparedness for evaluating nanotechnology was frequently noted by ELSI colleagues primarily in terms of identifying gaps in the regulatory framework (an issue of law, not science) or of questioning the ability of regulators to comprehend nanotechnology [34, 35]. Returning to CNTs as an example, one set of SSH authors [34] viewed the EPA's approach as an "unsustainable path," not realizing that the CNT-by-CNT review would place the administrative burden and costs onto the manufacturer not the EPA and not being aware of the TSCA exemptions. In a second study [35], the ambiguity in defining nanomaterial is recognized as undercutting regulatory actions, and the authors recommend that the FDA defines the term properly, but without fully anticipating that this Agency's interests might differ from those of the NNI. In contrast, STS investigators from the Center for Nanotechnology in Society at Arizona State University utilized the work of Kline and Rosenberg [4] when preparing for 19 "downstream" construction industry interviews through which they uncovered the importance of building codes to eventual nanotechnology adoption [44]. A second "downstream" ethnographic study at an AI firm pursuing big-data innovation found that an apparent regulatory void was in reality "governed by contextual legislation... and industry guidelines" [45]. In these contrasting examples, conventional ELSI methodology had a greater level of success if augmented by the specifics of firm-level or market-place-level interactions.

Clearly, the argument in this paper is that paying more attention to trajectories, the set of plausible innovation pathways, might address the policymakers' unease or at least allow SSH colleagues to translate their generalized concerns into ones more specific to the emerging technology. Yet, there are two dynamics that complicate this argument: one involves the STS community's commitment to technology assessments that combine the social construction of technology (SCOT) with the Collingridge dilemma [28] to become a form of governance, and one involves misinterpreting hype as reasonable promissory claims. The hype drives the dilemma leading to a

perceived inevitability if a technology is not engaged with during its early laboratory stages such as basic research (see Fig. 1).

Two prominent STS schools of technology assessment are real-time technology assessment (RTTA) associated with the Arizona State University and Constructive Technology Assessment (CTA) pursued extensively at the University of Twente. Both explicitly emphasize government-initiated technology development over firm-level activities [46, p. iii; 47, p. 3], potentially leading to an analysis gap as it is the firm that engages with regulators and marketplaces. The gap is compounded when the “interactional expertise” gained from pre-market, and “upstream” sites becomes the basis for generalized statements about post-market outcomes [48]. It becomes difficult to combine the results of “upstream studies with the considerable literature on SCOT, which asserts that technology is co-produced when social groups respond to technological offerings, a “downstream” site. A similar discordance occurs with the innovation studies literature that finds that there is considerable technological change occurring after market introduction, during the downstream diffusion process where products and technology undergo incremental changes initiated by their users [7]. Restated, the “interactional expertise” is sub-optimal if gained predominantly from “upstream” interventions. These two factors magnify the Collingridge dilemma by viewing nanotechnology introduction as inexorable and by focusing on early, upstream engagement. Set aside is Collingridge’s “logic of monitoring” where expert decisions are framed so as to be falsifiable [28, Chapter 10] and his guidance to act in reversible steps [28 pp. 193–194], e.g., the FDA’s evaluation stages. An STS narrative develops centering on the causal ambiguity underlying the NNI’s promissory claims and this narrative benefits little from the “interactional expertise” to be gained from examining the production experience, regulatory assessments, and customer acceptance of historical products [49].

The second STS dynamic, hype considered as promissory claims, revolves about claims that express commercial intent such as return on investment. ROI clearly communicates that profit is a motive and is one of several metrics used in administering a firm’s research project portfolio, i.e., management techniques such as the stage-gate model for project selection and monitoring.

(Stage-gate functions as the firm’s form of governance and is proposed by the NANoREG for safe-by-design [27].) While the ROI may be exaggerated, it expresses the underlying market assumptions and is not misleading within the firm nor among informed investors. One can explain the NNI’s initial promissory claims, for example, as an exaggeration intended to give the initiative credibility when setting Federal budgets. The exaggeration was inconsequential to nanomaterial research involving synthesis and characterization. However, the transition to hype occurs with the topics found in the 2003 Act, e.g., bioaugmentation, that lie well beyond the likely outcomes of a ten-year budget initiative encoded into law. These can be misleading relative to future NIH and FDA responsibilities. In that respect, nanotechnology hype merges with concepts like disciplinary capture [50] or even disciplinary imperialism [51], when taking the view that one field, materials science, and engineering [37], is dominating methodological decisions far beyond its recognized domain. These distinctions are challenging if the STS literature is disconnected from the innovation studies’ practice due to an “upstream” focus.

In general, the SSH literature views the NNI’s original hype as descriptive of nanotechnology’s commercial outcomes, which of course heightens concerns regarding the Collingridge dilemma. It would be difficult not to, but there is a circularity in that NNI-funded ELSI considered the NNI-hype to be a given rather than as a construct with parts warranting independent analysis. Alternatively, hype was viewed as having a positive effect in alerting the SHS community to the possible normative impacts and of doing so at an early stage of scientific development [52]. Less attention, however, was paid to the roles that the remaining actors in Fig. 1 might exercise in modulating the hype. In effect, the STS governance proposals focused on “upstream” actions so as to ensure that the “downstream” agencies only considered acceptable product forms. Nordmann captured this dynamic in terms of hype’s seductive nature in overwhelming the critical analysis expected in responsible research and innovation [53–55]. Here, I would argue that the technology trajectory might serve to situate the gradations in promissory claims for a fuller understanding regarding sources, motivations, and intended audiences.

Biotechnology Overview for Validation

In the case of nanotechnology, the history of promissory claims and hype could be aligned with the adjustments made to its definitions when products pass along their respective trajectories. A similar history occurred with biotechnology, an umbrella term in the life sciences utilized when the techniques for manipulating biological organisms anticipate a commercial purpose. In 1991, the FDA defined biotechnology as “the application of biological systems and organisms to technical and industrial processes,” aligning it with bioengineering, but in the EPA’s glossary, the definition is “the science of modifying the genetic composition of plants, animals, and microorganisms,” having the narrower sense of gene engineering. The Cartagena Protocol on Biosafety differentiates historical and current methods by defining “modern biotechnology” as the use of *in vitro* nucleic acid techniques involving genetic materials or the fusion of cells “beyond the taxonomic family” in order to overcome the “natural physiological reproductive or recombination barriers and that are not techniques used in traditional breeding and selection.” Clearly, a distinction is being made between historical techniques that steer heritable traits and the current methods that manipulate the germ plasm directly. There may be confusion in using the broader term, biotechnology, when mainly alluding to the narrower “modern” biotechnology. The concept of “modern” biotechnology has a clear foundation in gene engineering, itself an outgrowth of past achievements in establishing a chemical explanation for genetics, namely DNA, though without explicitly considering epigenetics [56]. While the history of their respective definitions is similar, there is a distinction in promissory claims. Nanotechnology is a means (nanoscale materials) with surprising outcomes (unique phenomena); biotechnology is a biological means (genes) with a preferred biological outcome (disease control). In Table 1, the two are considered “like” for definitions and “unlike” for promissory claims.

Jasanoff’s book, *Designs on Nature: Science and Democracy in Europe and the United States*, provides an overview of the legal, cultural, and regulatory issues encountered when introducing biotechnology products in Germany, the UK, and the USA [57]. The book provides a touchstone for both the benefits and the limitations of a technology trajectory approach, especially

regarding the importance of definitions, history, and marketplace actors. In general, Jasanoff’s examples reflect the broader definition of biotechnology when being applied to human health and the more focused “modern” biotechnology when examining plant products. (The issues of cloning, cell fusion techniques, and creation of laboratory research animals through *in vitro* means may not have been not prominent when the book was written.) The book demonstrates that there is a considerable history in terms of regulatory concepts as well as legal precedents that demand the attention of experts, committees, academic researchers, and industry scientists. Throughout that history, existing products, markets, and policies influenced the technology trajectory. Taken together, biotechnology was not a decidedly government-initiated effort, unlike nanotechnology (Table 1), for the sources of invention were dispersed among many people, firms, institutions, and the potential for policy implications were well communicated beforehand.

Realizing that regulatory approval would be required for foods, plants, and medicines does not require a trajectory analysis. It was a well-known fact that points to a limitation for this type of analysis. A trajectory may prepare the user for the sequence of interactions, agencies, events, and participants, but it does not anticipate the nature of these influences without the type of analyses that Jasanoff and others in the SSH community provide. For example, it was shown through bibliometric techniques that the initial “technological trajectory” for broad-spectrum cancer therapy agents underwent adjustments to become more targeted after “typical problems, opportunities, and targets” acted as “focusing devices” [58]. In a study involving industrial biotechnology, the “innovation trajectory” for manufacturing three products considered promising “sustainable solutions based on natural resources” were disappointing due to a reframing of naturalness [59]. Bioengineering vanillin production displaces the petrochemical plant recipe but is less natural than the vanilla orchid. Bioengineering of artemisinic acid, a precursor for an antimalaria drug, was less natural than the process based on the Chinese sweet wormwood plant. While Asveld et al. concluded that following RRI precepts would have alerted the companies sooner to these issues, these case studies and SSH analyses arose from there being tangible products before the results could be reflected back upon the overall biotechnology trajectory.

Viewed from the perspective of innovation studies, the firm and the market sector assume that the early

stages in the trajectory will have a global scientific reach that may at first need to be tailored to the specifics of a local jurisdiction, but that later are likely to be harmonized through such agencies as the Codex Alimentarius and the Organisation for Economic Cooperation and Development (OECD). This explanation was offered for the Australian acceptance of Bt-cotton [60] and applies also to the initial openness of French regulators to cultivating Bt 176 GM maize, a decision later overturned by the Prime Minister [61]. Following the initial acceptance, objections from farmers and consumers surfaced across several EU nations, often in the form of citizens participating in technology assessments. The SSH literature tracked these events in detail, including how citizen participants were eventually chosen to ensure that “the trajectory of biotechnological innovation has been protected from challenge” [61], and how industry complaints about regulatory policies hindering innovation ignored “the fact that regulations can influence the direction of technological trajectories, toward, for example, providing safer, more useful, and more sustainable products and processes” [62]. It was noticed by political scientists that the USA and Europe had traded places in terms of the environmental and regulations [63]. In the case of plant biotechnology, the global reasoning behind the technology’s promissory claims was at first shared by the local regulator, but the resulting policy entered into a legitimacy crisis due to the public’s reaction, which led to the diversity of national approaches described in Jasanoff’s book. Once again, the unexpected societal reaction arose from there being tangible products to consider and those reactions would have been difficult to measure otherwise.

There was no initial regulatory acceptance with nanotechnology. The NNI definitions and later hype undercut the EPA’s standard practice, while the FDA’s methodology exhibited greater resilience to the point of amending the NNI’s definition [33]. The different responses can be ascribed to the FDA’s experience with the promissory claims found in every IND and NDA, and therefore the steps taken to dampen the influence of hype. However, such resilience can also be viewed as flexibility bordering on accommodation. One line of SSH thought utilizes the concept of “regulatory objectivity” when examining the FDA’s use of consensus conferences to revise protocols, establish new conventions, or adjust test

requirements [64]. The FDA remains current but also arrives at intermediate decisions that may drift with the state-of-the-art and thereby lose sight of broader societal issues, e.g., the initial French approval of Bt maize. Another line of reasoning describes the influence that “promissory technological visions” have as forms of regulatory capture or “pharmaceuticalization.” [65, 66]. The regulator’s receptivity to promissory claims that regarding gene expression, either in drug therapy or in altering the germplasm in plants, can be connected to the biotechnology scientist’s and the regulator’s shared university education in biology and a likely commitment to exploiting genetic knowledge. The regulator’s role later becomes one affirming the progress made by “consenting” to the next stage of clinical trials and thereby providing the “informed” component of informed consent. In the case of personalized medicine, the boundary separating the clinic and the pharmaceutical research lab has changed in that the patient’s response to new drug probes their genetic predilections to cancer expressed as a tumor [67] rather than the clinician deciding if the drug demonstrates safety and efficacy. The regulator and clinician may even become advocates through their roles in devising the trial’s design as the drug itself has become an investigative tool. The parallels in “modern” biotechnology of plants or specialized microbes are an openness to field trials and a defensiveness regarding the test methods used in the approval process [68].

In summary, biotechnology has its roots in the historical utilization of biological processes for domestic and industrial uses. “Modern” biotechnology is a well-defined subset of techniques that pose significant issues regarding heritable traits. Past difficulties in regulating therapeutic products and processes have led to a regulatory framework consisting of defined, reversible stages (Collingridge’s logic of monitoring [28]), which means that laboratory testing and its accompanying promissory claims are known from the outset to require validation before being presented to a regulator. However, the test methodologies for “modern” biotechnology products face a considerable challenge, which has led to a greater openness for clinical trials and field tests as the means for validating these same test methodologies in the form of scientific hypotheses rather than demonstrated promissory claims. Both the biotechnology proponents and the regulators have a misplaced sense of security in

assuming that the regulatory framework is sufficient when extended to a topic where exposure is more widespread, i.e., beyond the clinic or the farmer's field. As documented in the extensive SSH literature on biotechnology, the general public's reaction can be unexpected and lead to political action, new statutory language, and changes in the technology's trajectory.

Applying the Lessons to Artificial Intelligence

Definitions were an instructive tool for understanding the trajectory taken by the products of nanotechnology or biotechnology and are a reasonable starting point for situating this analysis of AI. According to ISO/IEC 2382:2015(en), Information technology—Vocabulary, artificial intelligence is the “capability of a functional unit to perform functions that are generally associated with human intelligence such as reasoning and learning” [definition 2123770], where a “functional unit” is an “entity of hardware or software, or both, capable of accomplishing a specified purpose” [definition 2122865]. There is considerable overlap with the same source's description of automation as the “conversion of processes or equipment to automatic operation, or the results of the conversion” [definition 2121284], should the “processes or equipment” have once involved human intervention, e.g., reading, calibrating, interpreting, and maintaining gauges. The overlap between AI and automation is demonstrated in a bibliometric study of the AI literature where “automation & control systems” along with “instruments and instrumentation” formed the third of five subject categories [69, Figure 8]. As with biotechnology and “modern” biotechnology, a distinction is being made regarding historical techniques, which for automation is displacing mechanical devices, sensors, and motors by computer technology. In biotechnology, there is a mechanistic underpinning to the distinction, the use of *in vitro* nucleic acid techniques involving genetic materials or the fusion of cells, while with AI, there is an additional distinction to be made between the “functional unit's” operational mechanism and that of “the process or equipment” being acted upon. As noted earlier, AI and nanotechnology are similar (“like” in Table 1) in that the means and outcomes traverse different subject matter domains, whereas, in biotechnology, a biological means leads to a biological outcome (“unlike” in Table 1).

The potential for conflating the two mechanisms can be illustrated using recent EU activities. The European Commission facilitated the formation of the High-Level Expert Group on Artificial Intelligence (AI HLEG) as a step for developing an AI definition [70] and the accompanying set of ethical guidelines [71]. The suggested definition centers on software (and possibly hardware) needed to achieve a complex goal exemplified by a listing of AI techniques as if it were a scientific discipline. In the same time period, the Joint Research Centre conducted two workshops on AI. One was “to identify opportunities for meeting the EC demands on AI” [72], and the second pursued opportunities for AI in risk assessment [73]. The first workshop's report stated that AI systems will “take intelligent actions or propose decisions” and notes that “many of the methodological developments in AI date back more than 50 years.” The second workshop's report has no definition of AI. In it, machine-reading and machine-learning are expected to enhance the scientific-technical process as well as the social aspects surrounding decision-making to the point that the AI system becomes “an additional expert around the table,” one that is “neutral.” Not addressed are the criteria that will connect machine-learning and machine-reading, the functional units, to the mechanisms undergirding the EHS data to be gathered and interpreted, i.e., the modes of toxicity leading to adverse outcomes. Neither sets a “complex goal” for “intelligent actions or purposive decisions,” neither alludes to issues found in the AI HLEG ethical guidelines, and both use computational techniques that are not unique to AI. This dynamic of workshops and promissory claims tracks the history of nanotechnology in the USA. AI is an umbrella term open to the seductive pressures identified by Nordmann [53–55].

There have been similar episodes in the USA. Multiple workshops sponsored by the US Defense Advanced Research Projects Agency (DARPA) were the basis for a 1988 report on the promise of neural networks during an earlier period of AI research [74]. As with the NNI workshops, a future-shaping community developed reasonable definitions intended to distinguish the proposed effort from the then-existing hype. Neural networks were a computer architecture “modeled on biological processes.” More recently, colleagues from the US Army [75] wrote about “cybertrust” stating that “[a]ssuming that the AI

is correct more often than a human under the same time and resource constraints (underscoring the utility of AI applications), on average, the human who disagrees with the AI should still follow the AI's recommendations. The challenge of that moment of discordance, then, is to convince humans to trust the AI output despite their own opposing judgment." Starkly stated, as is appropriate to the battlefield, the issues of discordance, trust, and favoring AI decisions over personal judgments are pertinent to the civilian applications of this technology, such as the tangible example of driverless cars.

According to AI's promissory claims [76, 77], driverless cars will eventually have a lower accident rate than those driven by humans, and a plausible argument can be made that lives could be saved. End-user testing under controlled conditions, such as trucks on limited-access highways, will be pursued to measure lives saved as well as measure the productivity gains that will enter into ROI estimates. Historically, the automobile industry has faced comparable technology shifts (e.g., electric cars, unleaded gasoline) [8] and already has experience with driver assistance through cruise control and self-parking cars. Further, automobiles and trucks are part of a network of institutions that will influence the final outcome. For example, the EU's Joint Research Centre (JRC) has been active in this arena of connected and automated vehicles [78] as well as the implications for road utilization and traffic safety [79]. One unresolved aspect for both the driverless car and traffic safety is the insurance industry's role [77, 79, Sect. 3.6]. In the USA, personal liability is borne by the vehicle owner using the insurance company as an intermediary. Where does liability lie with a driverless car? With the owner? With the manufacturer? With the software company? With the local maintenance garage?

The essence of AI is the nature of the decision, one that once required a human, and one that requires a distinction between control and process. James Watt for example introduced the centrifugal governor to control the steam engine's flywheel. All of the linkages were mechanical and the logic was one of negative feedback. In the past, a driver controlled a car's speed by combining a knowledge of the speed limit with the speedometer reading to adjust the carburetor's fuel valve through its mechanical linkages to the foot pedal. Currently, cruise control devices sense the topography, interpret the data relative to a selected

target speed, and adjust fuel flow to the injectors. The linkages are electronic, the decision resides in a computer, and human intervention is minimal. In these cases, the engine's process, its causal mechanism, is separate from the means used for adjusting its control parameters, the "functional entity's" causal mechanism. The driverless car extends the range of parameters further, where the underlying machine's causal mechanism is tangible, and the issues regarding AI center on the "functional entity's" software and hardware. There are other topics, however, such as toxicity, migration, economic markets, and cognition [72] where accepted understanding of the underlying processes falls short of describing a causal mechanism. The issues of AI are now twofold: an understanding of the underlying process and an understanding of its control by AI's "functional unit." These elements align with the "principle of explicability" found in the AI HLEG's ethical guidelines for trustworthy AI where explicability comprises traceability, auditability, and transparent communication regarding system capabilities [72, p. 13].

Addressing the control aspect first, computational models and big data analytics have already been used to generate recommendations that a human might act upon, but with AI, the recommendation is itself the decision to act. Hence, the question becomes how does the "functional unit" possess the capability "of accomplishing a specified purpose"? How was the AI algorithm selected, updated maintained? For example, rule-based algorithms, a deductive approach, dominated the early development of driverless cars. Progress was accelerated, however, when the decision-making algorithm moved to an inductive, probabilistic model where neural networks (1) analyze the images coming from the car's sensors, (2) compare them to a library of actual street scenes with any resulting incidents, (3) determine the probable outcomes, and (4) take an action [76]. In this example, the learning process for control is the regular recalculation of probable outcomes using an ongoing centrally administered collection of sensor images, street scenes and incidents that must be uploaded regularly to the car's computer. Whether rule-based or probabilistic, there will be an algorithm (a computational method) based on a model that represents the target scenario (automobile approaching an intersection), and these will be housed in a computer, perhaps a device, to be sold with the automobile as property.

In viewing driverless cars from the perspective of automation, the initial AI justification of lower accident rates becomes an examination of humankind's understanding of the subject matter that is being automated and the ability to express that knowledge in an algorithmic form. There is a progression in the controlling "functional unit" in terms of its operation, mechanical linkages to electronic signals, and its identity, from a human reading a speedometer to an algorithm doing the same. The analogy to automation also raises the issue of governance, especially during the product development stage (Fig. 1). Automation is often internal to the firm or market segment [77, 80, pp. 61–64, 91–95], and there is rarely pre-market approval through a formal regulator [80, 81]. This of course poses a challenge to public engagement. CSOs, for example, often have a regulatory counterpart, which is yet-to-be-determined with driverless cars [82]. There may be surprises when a more obscure agency, advisory panel [83], or industry standards group [44] have responsibility [83] without the wherewithal to address discordances [75] or the potential for algorithms to discriminate when used beyond their applicability domain [84].

Unlike nanotechnology and biotechnology, the issues of control and causal mechanisms are not totally new to the SSH and natural science communities centered on biology and by extension to those working with models. Explanation in physics and chemistry is derived from general laws in a deductive fashion. Explanation in biology on the other hand does not rely on laws and instead takes the form of mechanisms (the claim) with a supporting narrative or mathematical model (the warrant) [85, 86]. A narrative and mathematics do not prove that the proposed mechanism is causal. Hodgkin and Huxley, for example, recognized that their mathematical model of the action potential provided a predictive capability that was based on a conjectural mechanism and was not a causal explanation [87]. SSH colleagues have further demonstrated that the knowledge base and its history influence the specific type of model that is considered acceptable [88] and have even suggested a taxonomy for mechanism-oriented models purporting to explain biological phenomena [89]. The modeling experience in biology even carries over to nanotechnology EHS, where there are similar considerations regarding model types and the validity of the descriptors used in those models [56] for the purpose of chemical risk analysis [73, 90].

The promissory claims for AI will cover many applications, and each may have a different set of criteria for selecting an acceptable model. Validating that the code in the algorithm expresses the model's functions properly will become increasingly difficult as targets become more challenging, the code becomes voluminous, the number of iterations increases, and the database expands. Without assistance from those who did the coding, the concepts and assumptions underlying the algorithm may be lost from view [91]. This loss of transparency in the coding will compound uncertainties about the model and has led to the concept of epistemic opacity [92]. Additionally, models utilizing mathematical formulae introduce a separate source of error when digitizing analytical functions or when relying on numerical methods for solving differential equations [91]. Essentially, each model is a middle-range theory [1, 11] requiring verification and validation to be considered trustworthy [71].

It is noteworthy that the experience from biology is mirrored in the concerns expressed by regulatory agencies. Computational modeling to supplement laboratory testing is already used in EPA and FDA regulatory submissions, and this experience anticipates future questioning of AI. In establishing credibility, the regulator's overriding concern is that the model might only be a correlation with a limited applicability domain. Deliberations for judging the credibility of quantitative structure–activity models [93] led the OECD to set the qualitative criteria of: (1) defined endpoint (phenomenon); (2) unambiguous algorithm; (3) defined domain of applicability; (4) appropriate measures of goodness-of-fit, robustness, and predictivity; and (5) "a mechanistic interpretation, if possible." More recently, the European Food Safety Authority [94] accepted a toxicokinetic/toxicodynamic model that evidently satisfied these criteria by specifically providing (a) two sets of "ring data" to be used in verifying that proprietary programs were functioning correctly (described as an implementation), (b) connecting choice of model format directly to the experimental design, and (c) providing an internet-accessible implementation of the model so that the regulator has a means for an independent validation. It is clear from these actions that the epistemic opacity will be questioned with a degree of rigor appropriate to the risk associated with the application (end-use). However, the forum and the criteria

to be used in that forum are not presently known for AI applications. This returns us to the earlier discussion regarding the technology trajectory's choice of upstream and downstream locations where the downstream regulatory step may involve a recognized agency like the FDA or may be the organizations administering building codes [44].

As with biotechnology and nanotechnology, awareness will be driven by the availability of AI products. Adoption will be based on the change criteria established in past innovation episodes [8] and may include conducting trials, setting industry standards, and, of course, calculating financial justifications (ROI). This was noted earlier with the FDA's methodology for new drugs [33] and is consistent with Collingridge's "logic of monitoring" based on reversible stages [28]. Hence, AI will have many sources of invention, and its products will encounter widely dispersed adoption criteria. Less clear is the combination of formal governance (regulators) and end-user evaluations needed when balancing AI's inherent epistemic opacity with the potential for injury should the computational techniques be incomplete. Likely, there will be a gradation of applications where some require more rigorous criteria for acceptability, e.g., predictive toxicity and others will rely on industry standards for managing change.

In summary, artificial intelligence has historical roots in automation, which previously relied on mechanical devices and electronic systems (sensors, signal transmission, electric motors) to achieve improved productivity. Definitions in AI are descriptive, emphasizing the use of computer algorithms, but leaving open issues regarding the deductive (deterministic, rule-based, mechanistic) and inductive (correlations, Bayesian logic) methods to be used in the modeling. While there will be an AI infrastructure market, a significant one for devices and data management, the primary use as a productivity tool and the accompanying promissory claims will be widely spread across existing markets. Inventiveness will be broad as new computer capabilities, software, and modeling approaches are tailored to plausible market scenarios. Where new markets are created, the issues of hype and monitoring will intersect with societal aspirations. Relative to the experience with nanotechnology and biotechnology, there is an open question on governance.

Concluding Remarks

The invitation to the special section in NanoEthics: Studies of New and Emerging Technologies that this article is part of asked if "the social learning processes intertwined with technology hype" that were uncovered with nanotechnology and biotechnology are sustainable "new patterns of interpretation" or are "tools" that can be applied to artificial intelligence. In response, the concept of technology trajectory has been proposed as an organizing principle for connecting these tools (upstream and downstream engagement, RRI, RTTA CTA, SCOT) with the sequence of events, actors, and responsibilities that occurred during the "social learning" experienced with nanotechnology and biotechnology. For each technology, the innovation studies and SSH perspectives were viewed in parallel with the resulting overall trajectory becoming the aggregate of the individual product experiences. As illustrated in Fig. 1, the large commercial firm with research, marketing, and sales functions has the greatest knowledge of the trajectory and has multiple opportunities to influence events. This is not the case for public engagement.

A prospective trajectory for AI is presented in Table 1 based on the paper's examination of nanotechnology and biotechnology. AI definitions will be adjusted "like" with nanotechnology and biotechnology; however, the nature of AI's promissory claims is closer to those of nanotechnology ("like") than to those of biotechnology ("unlike"). The sources of invention and history resemble biotechnology in not being as government-directed as with nanotechnology. These statements ("like" & "unlike") reflect that the trajectory is a cyclic process where products evaluated at the diffusion stage are intertwined with the research investment decisions on products under development (Fig. 1). It is in this dynamic that promissory claims can be judged as reasonable exaggeration or hype. While the product-level experience is "like" for all three technologies, there are two themes in Table 1 where neither biotechnology nor nanotechnology can provide guidance, i.e., evaluated as a process and regulatory receptivity.

A repeated theme in this article has been the nature and extent of regulatory influence on the trajectory, as the regulatory agency represents the first encounter between technological promise and a legal

understanding of societal acceptability. The regulators are primarily the FDA in the USA for biotechnology and nanotechnology products as medicines and the EPA for nanotechnology as industrial chemicals. The regulatory framework is yet to be determined for artificial intelligence (AI) and may be quite diffuse as the promissory claim extends to any “specified purpose” that a “functional unit” might control. Liability insurance was emphasized in the discussion of driverless cars, but there are open issues as well surrounding licensing, privacy, cybersecurity, and infrastructure [77] that nominally involve a spectrum of agencies, organizations, laws, and standards [44, 45, 77, 83]. Their respective capabilities and receptivity to AI are uncertain. This is “unlike” either nanotechnology or biotechnology (see Table 1) and poses a challenge for policy advocacy by civil society organizations [82].

For those reasons, the AI discussion was framed as an extension of automation with one component the “specific purpose” and the other the “functional unit’s” algorithm. The result is a gradation along two axes: knowledge of the phenomena underlying the specific purpose and the current ability for expressing that knowledge in algorithmic form. For AI, judging promissory claims as hype will occur when knowledge of the causal mechanisms explaining the phenomena are questioned. This approach aligns with the six levels of automation found in the Society of Automotive Engineers International standards, which tie the “dynamic driving task” with the “hardware and software” capabilities [77]. AI “like” biotechnology will be viewed as a process (“unlike” nanotechnology), but “unlike” both relative to regulatory receptivity.

It is noteworthy that the SSH community centered on biology is well positioned to contribute significantly to the public discourse on AI. Stated simplistically, those already familiar with databases, modeling of biological mechanisms, neuronal networks, and cognition have credibility for the central issues of a technology that has the objective of automating decisions once involving human cognition. Paramount to their contribution to public engagement will be drawing distinctions regarding causal mechanisms (knowledge of the “specific purpose”) and the extent of epistemic opacity (the articulation of that knowledge in algorithmic form). These are also the distinctions necessary to counteract the influence of

hype when deciding on the reversible stages found in Collingridge’s logic of monitoring.

Acknowledgements I wish to thank Drs. Fautz and Seifert for their efforts in pursuing this topic. They and the anonymous reviewers provided valuable insights that improved this work greatly. I have benefited from attending S.NET meetings and the occasional interactions with colleagues there, especially Professors Nordmann and Fisher. The manuscript includes personal observations made at meetings, including the multi-disciplinary stakeholder workshops conducted at UC-CEIN organized by Professors André Nel, Yoram Cohen, Barbara Herr-Harthorn, and Patricia A. Holden.

Declarations

Conflict of Interest I represented my employer, Degussa Corporation (now Evonik) at the American Chemistry Council’s Nanotechnology Panel, ASTM E56 and ISO TC-229. After leaving the industry in 2008, I continued as an independent consultant at ASTM E56 & ISO TC-229 and advised on stakeholder involvement at the UC-CEIN. These affiliations led to participation in meetings closed to the public.

References

1. Geels FW (2007) Feelings of discontent and the promise of middle range theory for STS: Examples from technology dynamics. *Sci Tech Human Values* 32:627–651
2. Shrum W (1984) Scientific specialties and technical systems. *Social Stud Sci* 14:63–90
3. Teece DJ (1987) Capturing value from technological innovation: Integration, strategic partnering, and licensing decisions. In: Guile BR, Brooks H (eds) *Technology and global industry: Companies and nations in the world economy*. National Academies Press, Washington, pp 65–95
4. Kline SJ, Rosenberg N (1986) An overview of innovation. In: Landau R, Rosenberg N (eds) *The positive sum strategy: Harnessing technology for economic growth*. National Academies Press, Washington, pp 275–306
5. Dosi G (1982) Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Res Policy* 11:147–162
6. Landau R, Rosenberg N (1986) Editors’ overview. In: Landau R, Rosenberg N (eds) *The positive sum strategy: Harnessing technology for economic growth*. National Academies Press, Washington, pp 1–16
7. Freeman C (1994) *The economics of technical change*. Cambridge J Econ 18:463–514
8. Utterback JM (1987) Innovation and industrial evolution in manufacturing industries. In: Guile BR, Brooks H (eds) *Technology and global industry: Companies and nations in the world economy*. National Academies Press, Washington, pp 65–95
9. Nelson RR, Winter SG (1982) *An evolutionary theory of economic change*. Belknap Press, Cambridge, MA

10. Godin B (2017) *Models of innovation: The history of an idea*. The MIT Press, Cambridge, MA
11. Merton RK (1968) On sociological theories of the middle range. *Social Theory Social Struct* 2:39–72
12. van Lente H, Rip A (1998) The rise of membrane technology: From rhetorics to social reality. *Social Stud Sci* 28:221–254
13. Klaessig FC (2016) The nexus of innovation management, safer-by-design and responsible research and innovation. In Bowman, DM Dijkstra A, Fautz C, Guivant, Konrad K, van Lente H, Woll (ed) *Responsibility and emerging technologies: Experiences, education and beyond*. IOS Press, Amsterdam pp 11–23
14. Miller G, Wickson W (2015) Risk analysis of nanomaterials: Exposing nanotechnology's naked emperor. *Rev Policy Res* 32(4):485–512
15. Brown N, Beynon-Jones SM (2012) Reflex regulation: An anatomy of promissory science governance. *Health Risk Soc* 14:223–240
16. Nickles T (1996) Deflationary methodology and rationality of science. *Philosophica* 58:9–50
17. Ankeny RA (2000) Descriptive models in biology: Of worms and wiring diagrams. *Philos Sci* 67(Proc.): S260–S272
18. Hacking I (2012) Language, truth and reason 30 years later. *Stud Hist Philos Sci Part A* 43:599–609
19. Gläser J, Laudel G (2016) Governing science: How science policy shapes research content. *Eur J Sociology* 57:117–168
20. Stilgoe J, Owen R, Macnaghten P (2013) Developing a framework for responsible innovation. *Res Policy* 42:1568–1580
21. Sovacool BK, Ramana MV (2015) Back to the future: Small modular reactors, nuclear fantasies, and symbolic convergence. *Sci Tech Human Values* 40:96–125
22. Eisler MN (2013) The ennobling unity of science and technology: Materials sciences and engineering, the department of energy, and the nanotechnology enigma. *Minerva* 51:225–251
23. Klaessig FC (2013) Setting useful nanotechnology definitions examined from the standpoint of central and peripheral claim language used in intellectual property law. In: Konrad K, Coenen C, Dijkstra A, Milburn C, van Lente H (eds) *Shaping emerging technologies: Governance, innovation, discourse*. IOS Press, Amsterdam, pp 127–136
24. Bhupatiraju S, Nomaler Ö, Triulzi G, Verspagen B (2012) Knowledge flows—analyzing the core literature of innovation, entrepreneurship and science and technology studies. *Res Policy* 41:1205–1218
25. von Schomberg R (2013) 'A vision of responsible research and innovation. In: Owen R, Heintz M, Bessant J (eds) *Responsible innovation: Managing the responsible emergence of science and innovation in society*. Wiley, New York, pp 51–74
26. van Wezel AP, van Lente H, van de Sandt JJM, Bouwmeester J, Vandenberg RLL, Sips AJAM (2018) Risk analysis and technology assessment in support of technology development: Putting responsible innovation in practice in a case study for nanotechnology. *Integr Environ Assess Manage* 14:9–16
27. Gottardo et al. (2017) NANoREG framework for the safety assessment of nanomaterials, EUR 28550 EN. <https://doi.org/10.2760/245972>
28. Collingridge D (1982) *The social control of technology*. St. Martin's Press, London
29. US Congress (2003) 21st Century Nanotechnology Research and Development Act of 2003. Public Law No 108–153, 117 STAT. 1923
30. Klaessig FC (2017) Nanotechnology definitions at ISO and ASTM International: Origin, usage, and relationship to nomenclature and regulatory and metrology activities In: Mansfield E, Kaiser D, Fujita D, van de Voorde M (ed) *Metrology and Standardization of nanomaterials. Protocols and industrial innovations*. Wiley-VCH, Weinheim, pp. 49–73
31. Roco MC (2007) National nanotechnology initiative—past, present, future. In: Brenner D, Lyshevski SE, Iafate GJ (ed) *Handbook on nanoscience, engineering and Technology*. Taylor and Francis, London, pp. 19–44
32. D'Mello SR, Cruz CN, Chen M-L, Kapoor M, Lee SL, Tyner KM (2017) The evolving landscape of drug products containing nanomaterials in the United States. *Nat Nanotechnol* 12:523
33. Food and Drug Administration (2017) Drug products, including biological products, that contain nanomaterials, Guidance for Industry, Draft Guidance, December 2017. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/drug-products-including-biological-products-contain-nanomaterials-guidance-industry> (accessed 5 November 2020)
34. Beaudrie CEH, Kandlikar M, Satterfield T (2013) From cradle-to-grave at the nanoscale: Gaps in US regulatory oversight along the nanomaterial life cycle. *Env Sci Tech* 47:5524–5534
35. Ramachandran G, Wolf SM, Paradise J, Kuzma J, Hall R, Kokkoli E, Fatehi F (2011) Recommendations for oversight of nanobiotechnology: Dynamic oversight for complex and convergent technology. *J Nanopart Res* 13:1345–1371
36. Gallo J (2009) The discursive and operational foundations of the national nanotechnology initiative in the history of the national science foundation. *Perspect Sci* 17:174–211
37. Roco M (2019) Comments to the synthesis of workshops discussions (48:37 minutes), the future of the NNI: A stakeholder workshop. (<http://www.tvworldwide.com/events/nanotech/190801/>) accessed 5 November 2020
38. National Science and Technology Council, Committee on Technology, Subcommittee on Nanoscale Science, Engineering and Technology (2000) National nanotechnology initiative, the initiative and its implementation plan, July 2000. <https://www.nano.gov/node/243>. Accessed 5 November 2020
39. Converging technologies for improving human performance: Nanotechnology, biotechnology and cognitive science, an NSF/DOC sponsored event, Mihail C. Roco and William Sims Bainbridge (eds.) 2002 http://www.wtec.org/ConvergingTechnologies/1/NBIC_report.pdf. (accessed 5 November 2020)
40. Holden PA, Klaessig FC, Turco RF, Priester JH, Rico CM, Avila-Arias H, Mortimer M, Pacpaco K, Gardea-Torresdey JL (2014) Evaluation of exposure concentrations used in assessing manufactured nanomaterial environmental hazards: Are they relevant? *Env Sci Tech* 48:10541–10551
41. Tennent HG (1987) Carbon fibrils, method for producing same and compositions containing same, US 4,663,230, assigned to Hyperion Catalysis International, Inc

42. Green MLH, Tsang, SC (1994) Uncapped and thinned nanotubes and process, US 5,346,683, assigned to the Gas Research Institute
43. Fisher E (2019) Governing with ambivalence: The tentative origins of socio-technical integration. *Res Policy* 48:1138–1149
44. Arora SK, Foley RW, Youtie J, Shapira P, Wiek A (2014) Drivers of technology adoption — the case of nanomaterials in building construction. *Tech Forecasting Social Change* 87:232–244
45. Arthur KNA, Owen R (2019) A Micro-ethnographic study of big data-based innovation in the financial services sector: Governance, ethics and organisational practices. *J Bus Ethics*. <https://doi.org/10.1007/s10551-019-04203-x>
46. Larédo P, Jolivet E, Shove E, Raman S, Rip A, Moors E (2002) Management tools and a management framework for assessing the potential of long-term S&T options to become embedded in society. TSER Programme of the European Commission, Paris
47. Smits R, van Merkerk R, Guston DH, Sarewitz D (2008) The role of TA in systemic innovation policy. Utrecht University, Department of Innovation Studies, No. pp 08–11
48. Collins HM, Evans R (2002) The third wave of science studies: Studies of expertise and experience. *Social Stud Sci* 32:235
49. Klaessig FC (2012) Undone science and science un-done at nanotechnology’s periphery. In: Van Lente H, Coenen C, Fleischer T, Konrad K, Krabbenborg L, Milburn C, Thoreau F, Zülsdorf T (ed), *Little by little: Expansions of nanoscience and emerging technologies*, IOS Press Amsterdam, pp 43–53
50. Brister E (2016) Disciplinary capture and epistemological obstacles to interdisciplinary research: Lessons from central African conservation disputes. *Stud Hist Philos Biol Biomed Sci* 56:82–91
51. Mäki U (2013) Scientific imperialism: Difficulties in definition, identification, and assessment. *Int Stud Philos Sci* 27:325–339
52. Van De Werff T, Slatman J, Swierstra T (2016) Can we “remedy” neurohype, and should we? Using neurohype for ethical deliberation. *AJOB Neurosci* 7:97–99
53. Nordmann A, Schwartz A (2010) Lure of the “yes”: The seductive power of technoscience. In: Kaiser M, Kurath M, Maasen S, Rehmann-Sutter C (ed) *Governing future technologies: Nanotechnology and the rise of an assessment regime*, *Sociology of Sciences Yearbook* 27:255–277
54. Nordmann A (2007) If and then: A critique of speculative NanoEthics. *NanoEthics* 1:31–46
55. Nordmann A (2014) Responsible innovation, the art and craft of anticipation. *J Responsible Innovation* 1:87–98
56. Griesemer J (2002) What is “Epi” about epigenetics? *Ann NY Acad Sci* 981:1–14
57. Jasanoff S (2005) *Designs on nature: Science and democracy in Europe and the United States*. Princeton University Press, Princeton
58. Coccia M, Wang L (2015) Path-breaking directions of nanotechnology-based chemotherapy and molecular cancer therapy. *Technol Forecast Soc* 94:155–169
59. Asveld L, Osseweijer P, Posada JL (2019) Societal and ethical issues in industrial biotechnology. In: Hiete M (ed) Fröhling M. Sustainability and life cycle assessment in industrial biotechnology, Springer Cham, pp 121–141
60. Wickson F (2009) Reliability rating and reflective questioning: A case study of extended review on Australia’s risk assessment of Bt cotton. *J Risk Res* 12:749–770
61. Levidow L (2009) Democratizing agri-biotechnology? European public participation in AgBiotech assessment. *Comp Sociol* 8:541–564
62. Stirling A, Glover D, Millstone E (2015) Regulating genetic engineering: The limits and politics of knowledge. *Issues Sci Technol* 31:23–26
63. Kelemen RD, Vogel D (2010) Trading places: The role of the United States and the European Union in international environmental politics. *Comp Political Stud* 43:427–456
64. Cambrosio A, Keating P, Schlich T, Weisz G (2006) Regulatory objectivity and the generation and management of evidence in medicine. *Social Sci Med* 63:189–199
65. Abraham J, Davis C (2007) Deficits, expectations and paradigms in British and American drug safety assessments. *Sci Tech Hum Values* 32:399–431
66. Abraham J (2011) Evolving sociological analyses of ‘pharmaceuticalisation’: A response to Williams. *Martin and Gabe Sociol Health Illness* 33(5):726–728
67. Vignola-Gagné E, Keating P, Cambrosio A (2017) Informing materials: Drugs as tools for exploring cancer mechanisms and pathways. *Hist Phil Life Sci* 39:10
68. Wickson F, Wynne B (2012) Ethics of science for policy in the environmental governance of biotechnology: MON810 maize in Europe. *Ethics Policy Environ* 15:321–340
69. Liu N, Shapira P, Yue X (2021) Tracking developments in artificial intelligence research: Constructing and applying a new search strategy. *Scientometrics* 126:3153–3192
70. AI HLEG (2019) A definition of AI: Main capabilities and disciplines—EU high-level expert group on artificial intelligence. [on-line] European Commission. Retrieved on 5 November 2020
71. AI HLEG (2019) Ethics guidelines for trustworthy AI. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>. Accessed 5 November 2020
72. Nativi S, Gómez Losada A (ed) (2019) *Artificial intelligence at the JRC*, EUR 29798 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 9789276088424, <https://doi.org/10.2760/705074>, JRC117232
73. Wittwehr C, Blomstedt P, Gosling JP, Peltola T, Raffael B, Richarz A-N, Sienkiewicz M, Whaley P, Worth A, Whelan M (2020) Artificial intelligence for chemical risk assessment. *Comput Toxicol* 13:100114
74. Guice J (1999) Designing the future: The culture of new trends in science and technology. *Res Policy* 28:81–98
75. Linkov I, Galaitsi S, Trump BD, Keisler JM (2020) Kott A (2020) Cybertrust: From explainable to actionable and interpretable artificial intelligence. *Computer* 53(9):91–96
76. Lipson H, Kurman M (2016) *Driverless: Intelligent cars and the road ahead*. The MIT Press, Cambridge
77. Lee D, Hess DJ (2020) Regulations for on-road testing of connected and automated vehicles: Assessing the potential for global safety harmonization. *Transp Res Part A: Policy Pract* 136:85–98

78. Gómez LA (2017) Data science applications to connected vehicles: Key barriers to overcome, EUR 28923 EN (JRC108572). Publications Office of the European Union, Luxembourg. <https://doi.org/10.2760/822136,JRC108572>
79. Raposo AM, Ciuffo B, Makridis M, Thiel C (2019) The r-evolution of driving: From connected vehicles to coordinated automated road transport (C-ART), part I: Framework for a safe & efficient coordinated automated road transport (C-ART) system. <https://doi.org/10.2760/225671>
80. Stahl BC (2021) Artificial intelligence for a better future an ecosystem perspective on the ethics of AI and emerging digital technologies. Springer, Cham
81. Ulicane I, Knight W, Leach T, Stahl BC, Wanjiku W-G (2020) Framing governance for a contested emerging technology: Insights from AI policy. *Policy Soc.* <https://doi.org/10.1080/14494035.2020.1855800>
82. Hess, DJ (2020) Incumbent-led transitions and civil society: Autonomous vehicle policy and consumer organizations in the United States. *Technol Forecast Soc: 119825*
83. Verhoest K (2018) Agencification in Europe. In: Ongaro E, van Thiel S (eds) *The Palgrave handbook of public administration and management in Europe*. Palgrave Macmillan, London, pp 327–346
84. Williams BA, Brooks CF, Shmargad Y (2018) How algorithms discriminate based on data they lack: Challenges, solutions, and policy implications. *J Inform Policy* 8:78–115
85. Winning J, Bechtel W (2018) Rethinking causality in biological and neural mechanisms: Constraints and control. *Minds Mach* 28:287–310
86. Braillard P-A, Malaterre C (2015) Explanation in biology: An introduction. In: Braillard P-A, Malaterre C (ed) *Explanation in biology*. Springer, Dordrecht, pp. 1–28
87. Craver CF (2006) When mechanistic models explain. *Synthese* 153:355–376
88. Winther RG (2014) Parts and theories in compositional biology. *Biol Philos* 21:471–499
89. Hunt CA, Erdemir A, Lytton W, Mac Gabhann F, Sander E, Transtrum M, Mulugeta L (2018) The spectrum of mechanism-oriented models and methods for explanations of biological phenomena. *Processes* 6:56
90. Haase A, Klaessig F (2018) EU US Roadmap Nanoinformatics 2030. <https://doi.org/10.5281/zenodo.1486012>; <https://www.nanosafetycluster.eu/Nanoinformatics2030.html>. Accessed 5 November 2020
91. Hocquet A, Wieber F (2017) Only the initiates will have the secrets revealed: Computational chemists and the openness of scientific software. *IEEE Ann Hist of Comp* 39:40–58
92. Humphreys P (2009) The philosophical novelty of computer simulation methods. *Synthese* 169:615–626
93. OECD (2007) Series on Testing and Assessment No. 69, Guidance document on the validation of (quantitative) 'structureactivity relationship [(Q)SAR] models. ENV/JM/MONO(2007)2
94. EFSA Panel on Plant Protection Products and their Residues (2018) Scientific opinion on the state of the art of toxicokinetic/toxicodynamic (TKTD) effect models for regulatory risk assessment of pesticides for aquatic organisms. *EFSA J* 16:5377. <https://doi.org/10.2903/j.efsa.2018.5377>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.