

Travelling ML Models Across Different Socio-Technical Contexts and Purposes

Diletta Huyskes, University of Milan
Maria Sapignoli, University of Milan
Giuseppe Primiero, University of Milan

Abstract

The design and deployment of artificial intelligence (AI) often rest on specific cultural assumptions, perpetuating one-size-fits-all approaches and systems of oppression when applied uncritically across contexts (Birhane, 2020; Gabriel, 2020; Milan & Treré, 2024; Gammelgaard et al., 2023). While AI aspires to universal generalization, its outputs depend on data sources and design choices rooted in specific socio-cultural, economic, and technical settings (Chun, 2021). These values and assumptions shape how models function and make decisions. When transferred across geographies or repurposed for new tasks without translation, AI risks reinforcing inequities, erasing local specificities, and embedding hegemonic norms.

This paper examines how cultural values, positionality, and problem framing influence technological systems, stressing the dangers of models that travel without translation (Lu & Qiu, 2023). Building on science and technology studies (STS) and the philosophy of computer science, we extend the concept of travelling technologies (Rottemburg, 2002), proposing a framework for evaluating models through the dual lenses of context, the dynamic sociotechnical configuration of norms, values, institutions, regulation, and infrastructures (Suchman, 2007) and purpose, or the aims of a system and the project behind it. Through these axes, we develop four scenarios of technological travel, linked to the technical elements implied in redesign and reconfiguration (Facchini & Termine, 2022; Buda, Manganini, & Primiero, 2025). The framework helps trace epistemic and material

consequences of technological displacement, contributing to debates on coloniality and technoscience, and advocating for more just, situated approaches to AI.

Introduction

The design and development of artificial intelligence (AI) are inextricably tied to the socio-cultural, political, and technical contexts in which they originate. These contexts profoundly shape how systems are later deployed, conditioning both their functionality and their social consequences. Predominantly shaped by dominant epistemologies, AI systems embed implicit cultural assumptions and worldviews that, when uncritically transferred across domains, perpetuate universalistic and solutionist approaches that risk reinforcing existing systems of oppression (Birhane, 2020; Gabriel, 2020; Milan & Treré, 2024; Gammelgaard et al., 2023).

In recent years, several cases have highlighted these tensions, across Europe (Heikkilä, 2022), India (Rajalakshmi, et al. 2018), South Africa (Hao & Swart, 2022) and others. Generalization in AI does not ensure that fairness and accuracy are preserved. As Pham, Zhang, and Zhang (2023) demonstrate, models that perform accurately and fairly in their training environments can fail when applied elsewhere, precisely because their generalization relies on statistical correlations and data-specific patterns rather than transferable causal understanding (Chun, 2021). These datasets, and the design choices surrounding them, are shaped by particular social and institutional conditions, meaning that models remain contingent on the priorities and assumptions of their creators (boyd & Crawford, 2012). For example, a model trained in one institutional framework may not function effectively in another, challenging the idea that AI can seamlessly “travel” across contexts.

While the arguments developed in this paper could apply broadly to data-driven technologies, our focus is specifically on machine learning (ML), whose distinctive learning feature makes it particularly sensitive to the social, cultural, and epistemic conditions under which it is developed. The application of ML in real-world cases is often framed as a linear process of diffusion, yet literature shows a more complex picture (Jasanoff, 2004; Suchman, 2011). Models explicitly designed for export or

informally repurposed both carry embedded values that may not align with the ethical or practical needs of new environments.

Building on the notion of travelling technologies (Rottenburg, 2002; Lu & Qiu, 2023), we apply it to ML models to show that they are never simply transferred but become entangled with the sociocultural and institutional environments they enter. Transferring models across geographies without meaningful translation risks erasing local specificities and reinforcing hegemonic norms (Behrends et al. 2014). This raises a central question: are ML models translated when they travel, and what happens when they are not?

In this paper, our analysis deliberately begins at the point of origin of ML models. This is not to overlook the importance of deployment sites, but to demonstrate how powerfully initial assumptions and purposes condition what models can become once they travel. We understand design not merely as a technical phase, but as a socio-technical space in which context and purpose are inscribed, negotiated, and stabilized – a process through which particular worldviews, values, and problem framings become materially encoded into the system. If a machine learning (ML) model is trained on data collected within a particular financial system, healthcare institution, or legal framework, its ability to function effectively in a different context will depend on how closely that new environment aligns with the original data's underlying assumptions. Against this backdrop, this paper examines how cultural values, positionality, and socio-political contexts shape ML systems, and proposes a framework that conceptualizes technological travel along two axes: context and purpose. To address this, the paper proceeds as follows. First, we situate our approach within scholarship on technological diffusion and travelling models, highlighting how AI complicates classical accounts of innovation and transfer. We then introduce the concepts of context and purpose as key axes for understanding the movements of ML, and propose a framework that distinguishes four scenarios of technological travel. Each scenario is illustrated with concrete cases that show

how values and assumptions embedded during design shape deployment across geographies and domains. Finally, we discuss the broader implications of this framework for ongoing debates on epistemic injustice, coloniality, and the governance of AI, advocating for more just and situated approaches (Haraway, 1988) to technological design and translation.

Travelling Technologies

The study of how technologies travel across different contexts has been a central concern in STS, anthropology of technology, and postcolonial studies. Epistemic practices, extractive processes, and exploitative materials that potentially limit the agency of communities in shaping their own technological future has been challenged. Early scholarship on technological diffusion tended to emphasize the spread of innovations as a process of knowledge transfer from “advanced” to “less developed” regions (Hassan, 2023; Rogers, 1983). When writing about the generation of innovations, Rogers argues that past research typically began with the first adopter, ignoring the crucial sequence of decisions and processes that occur before diffusion begins. The first step that he described is that of identifying a need or issue – social, technical, or environmental. This recognition can be top-down (e.g., by scientists or governments) or bottom-up (from users or communities). Clearly, much changes depending on this distinction, as the recognition of needs and purposes can already encode different values and power relations. Innovations then take shape through research and iterative development, a stage that includes prototyping, testing, feedback loops, and successive refinements. These are essential steps, as purpose and location of innovation development greatly shape the trajectory of diffusion. While this phase is often underreported in traditional studies – which tend to focus on adoption outcomes – it has been a central concern in science and technology studies (STS), particularly within social constructivist

approaches. These perspectives sought to “open the black box” of technological development, drawing attention to the social, political, and material negotiations that underpin design processes (Bijker, Pinch and Hughes 1987). Technologies are the contingent outcomes of controversies, alignments of interests, values and negotiations. Different groups may propose or prioritize different features, uses, or design logics, leading to a range of prototypes or configurations. This is why Bijker, Pinch and Hughes (1987) introduced the concept of interpretative flexibility to describe how, initially, technological artifacts do not have a single, fixed meaning or use.

STS research has long shown that technologies are shaped well before diffusion: values, power dynamics, and assumptions are inscribed into artifacts and continuously reconfigured as they circulate (Latour 1987; Akrich 1992). Akrich’s notion of inscription (1992) similarly emphasizes that designers encode assumptions about users into artifacts, but these scripts are often reinterpreted or subverted in practice. Other scholars added a systemic dimension, emphasizing that technologies are shaped not only by isolated actors but by entire socio-technical networks including infrastructures, organizational practices, and political forces (Hughes 1994; Bowker & Star 2000).

Building on this scholarship, we take up the notion of travelling models (Behrends, Park and Rottenburg 2014) to highlight how the assumptions inscribed during the design and development of ML systems condition their trajectories across sociotechnical and geopolitical settings. Models do not circulate intact; they are translated, reinterpreted, and negotiated through interactions with local actors, infrastructures, and epistemologies. As Tsing thoroughly emphasized, global forms travel as assemblages, not as complete, intact packages (2005). The same ML model may mean radically different things – and behave differently – when “plugged into” different infrastructures, values, or institutional logics. Claims of universality in domains such as credit scoring, policing, or disease detection mask the fact that

models depend on local actors for interpretation and operationalization, and each step – design, adaptation, deployment – is a site of political and creative action.

Emphasizing the design phase means highlighting the moment when epistemic, normative, and technical assumptions are codified for the first time. These early inscriptions profoundly shape later outcomes and delimit what kinds of translation are possible. The question, then, is: what happens when models travel without being translated? Such situations often generate what we call “epistemic frictions”: tensions that arise when models built on assumptions of universality, rationality, and objectivity confront sociotechnical contexts governed by different knowledge systems, values, or ontologies (Haraway 1988). In these cases, misalignment does not simply reduce performance; it produces deeper forms of epistemic violence by imposing partial, often Western-informed assumptions – which reflect the most common design contexts – onto settings with divergent logics (Birhane 2021; Ricaurte 2019).

Rottenburg (2009) conceptualizes translation as a chain of transformations in which every act involves both omission and addition. His later work (2024) further explores how science and technology circulate and are reconfigured in African contexts, reproducing imperialist and colonial logics and showing that translation is not merely technical or linguistic, but also epistemological and ontological. Yet, even when translation occurs, it does not necessarily neutralize the historical and structural asymmetries embedded in these processes. Efforts may remain superficial, focusing on technical or linguistic adaptation while leaving intact deeper epistemological and ontological assumptions. As such, translation can sometimes serve to legitimize and obscure underlying power asymmetries, rather than resolve them. Research on AI development in the Minority and Majority Worlds (Mohamed et al., 2020; Madianou, 2025) shows how models and frameworks designed in dominant geopolitical centers are often imposed elsewhere without meaningful translation, reinforcing colonial logics.

These dynamics are evident in current debates also on AI ethics, which should address these issues. Global frameworks, often written in Europe and North America, are presented as neutral but reflect liberal, Western traditions (Roche et al. 2023). Their universal principles – such as fairness, accountability, or transparency – take on different meanings in diverse cultural settings (Mager et al. 2025). Finally, empirical studies show how the failure of models to generalize across diverse populations (e.g., Eubanks 2018) makes visible the cultural particularities of data and design that usually remain hidden until breakdown occurs. This raises critical questions about the adaptability and accountability of AI systems when they move beyond their original design contexts. Understanding the travelling nature of AI models is thus particularly urgent considering increasing concerns about AI ethics and the global governance of technology.

Values Embedded in Machine Learning

Scholars have shown that AI is deeply entangled with colonial histories and epistemologies. Adams (2021), for example, argues that AI’s emphasis on abstraction and objectivity emerges from Enlightenment and eugenicist traditions, excluding non-Western forms of knowledge, and reproducing colonial “dividing practices” like classification, enumeration, and surveillance (Chollet, 2019; Adams, 2021; Appadurai, 1993). Adams also highlighted that the “knowledges upon which AI is built – statistical enumeration of people and land – were advanced by imperial powers to control and contain colonial populations” (2021, 177) and that “AI as a field depends on, and was made possible by, these logics” (2021, 179).

Critiques of central ML values – such as accuracy, efficiency, and generalization – reveal that performance metrics often serve technical community priorities, overshadowing ethical and contextual dimensions and ignoring the fact that accuracy is always socially constructed (MacKenzie, 1990). As Birhane et al. (2022) showed,

ML research actually encodes and advances specific performance values but rarely justifies how the project connects with societal needs or even discusses negative consequences. Moreover, technical practices like feature engineering, hyperparameter tuning, labelling, selection of proxies, are all sites where values and normative assumptions are encoded, often invisibly. Suresh & Guttag's (2019) framework is particularly instructive, as it shows how sources of downstream harm arise across multiple layers of the ML pipeline. In the same direction, philosophy of computer science adds that computational artifacts can be described at multiple levels of abstraction (Floridi & Sanders, 2004; Primiero, 2019). In ML, values appear at least at three interrelated layers: the training data, encoding social and historical biases; the training engine, which amplifies these through model architecture and optimization; and the learned model, which operationalizes correlations in new inputs (Buda, Manganini & Primiero 2025). Each level models specific entities or activities relevant to the functioning of the artifact (Floridi & Sanders, 2004; Primiero 2016, 2019), and each level is implemented in the level below, and abstracted in the one above (Primiero, 2019). The top level is called the Functional Specification Level (FSL): at this level the artifact's behavior is given in abstract mathematical terms describing the function it computes. This function is implemented at the Design Specification Level (DSL): at this level, the function is translated in semi-formal processes corresponding to the function. The DSL is operationalised at the Algorithm Design Level (ADL) algorithmically as rule-based systems, typically in pseudo-code. At the Algorithm Implementation Level (AIL) algorithms from the ADL are encoded into high- and low-level programming languages. Finally, at the Algorithm Execution Level (AEL) one is concerned with the physical execution of those programs on actual hardware. The FSL has normative force over the behaviour of all levels below: this means that the function described at FSL establishes the correct behaviour of the artifact, and that the values embedded at this level should trickle down to all levels below; any additional norm implemented at a level below may become a source of error (Fresco & Primiero, 2013).

Accordingly, each of these components, with its own set of cultural values, establishes a context of use for the entire ML system. In particular, in this structure, the behavior of the model is not predetermined (in LoA jargon: there is no well-defined FSL): instead, it emerges dynamically when the learning model is trained and then when it is executed on a given test set. This means that the norm according to which the artifact will behave cannot be defined a priori without knowing the training and test data. This illustrates why the functionality of ML systems cannot be fully specified in advance, and why contextual awareness is critical.

Accordingly, when ML models travel, they do not simply export neutral technical functions. Recognizing that values are embedded throughout the entire ML pipeline – not just at the level of application, but in technical choices and assumptions – complicates the notion of translating models across contexts, as it reveals how deeply entangled these artifacts are with their original sociotechnical environments. Translation therefore requires not just technical transfer but active reinterpretation of values, objectives, evaluation criteria. Without this, what circulates is not only a model but a worldview – one that risks being imposed as universal while erasing alternatives.

A Framework for Travelling Models

Machine learning models – whether predictive policing algorithms or health risk scoring systems – can be understood as travelling models. They originate in one context (e.g., a tech company, research lab, or public administration) and are adopted in others (e.g., cities in Africa, schools in Europe, hospitals in India). In this sense, context plays a fundamental role as the sociotechnical dynamic configuration of development sites including social norms, cultural values, institutional infrastructures, regulatory environments, and material technologies (Suchman 2007). Alongside context, we foreground purpose: what a model is for, both its explicit

goals (e.g., predicting risk, optimizing services) and its implicit project (e.g., enforcing governance logics, extracting data, enacting forms of social control), as well as those of the broader socio-technical project in which it is embedded. The purpose of AI tools is rarely limited to the technical problem they are said to solve; it is always shaped by broader institutional logics and political imaginaries. These elements together co-produce the meaning, function, and legitimacy of technology as it is developed, deployed, and possibly translated. We use the term *travel* to refer to the movement or use of a technological artifact across different contexts and/or purposes. By contrast, *translation* denotes the process through which such an artifact is adapted or redesigned to fit the epistemic, cultural, or practical specificities of a new context or purpose. While travel may occur without translation, our argument is that every technological travel should entail some form of translation.

Context and purpose allow us to trace how models travel, are translated or fail to do so. Moreover, the two axes are mutually constitutive: the sociotechnical context informs how a model's purpose is defined, while the model's purpose, in turn, shapes which aspects of the context are foregrounded, neglected, or transformed. This is why, as discussed in the previous paragraph, every AI project should be examined in relation to the broader global structures that shape it: from infrastructure and research funding to standardization and imaginaries of what it should do and for whom.

We develop a framework that delineates four scenarios of technological travel. As shown in Figure 1, these scenarios are positioned along two intersecting axes representing the degrees of transformation in context and purpose, illustrating how different forms of technological travel emerge from their interaction: (I) Native Deployment, (II) Model Translation, (III) Function Creep, and (IV) Double Drift. This typology is not a rigid classification but a diagnostic lens for researchers and practitioners to interrogate the epistemic and political consequences of ML implementations.

I. Native Deployment

This scenario represents the most straightforward and foundational form of model deployment, as neither the context nor the purpose change. In this case, the model is both designed and deployed within the same socio-technical environment, and the system is built with a specific set of infrastructural, institutional, cultural, and regulatory assumptions in mind, and it is implemented in precisely that setting, following (at least initially) the function it was intended to serve. The objectives and operational logics of the model are developed in response to specific needs or priorities, and the deployment does not require major adaptation or reinterpretation. This continuity between design and implementation is often taken for granted in ML narratives, where the original context is treated as the baseline or reference case for evaluating technical performance. However, it is important to recognize that even within this original scenario, the values, norms, and problem frames embedded in the model are never neutral. Often, even in this starting context, they are far from explicit and emergent from AI-based projects, and unlike other scenarios, they are simply more easily naturalized because they reflect the dominant logics in continuity with their environment of origin.

A well-known example for this scenario is the use of the COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm in the U.S. criminal justice system. Developed by the private company Northpointe, COMPAS was explicitly designed for use within that legal framework to assist judges, parole officers, and other legal actors in assessing the likelihood of reoffending, or recidivism, among defendants. Its purpose – predicting future risk for the sake of sentencing, bail, or parole decisions – was tightly coupled with its intended deployment context: the North-American criminal justice system.

The tool was trained on U.S. criminal records and behavioral data (Lagioia et al. 2023), and implemented in judicial settings across various states without major alterations in its objective or institutional environment. Both the context (U.S. courts and penal logic) and the purpose (risk-based decision-making) remained consistent with the assumptions embedded during the tool’s design. However, this alignment did not prevent controversy. Investigations (notably, ProPublica 2016) revealed that the algorithm systematically reflected racial biases present in the training data and in the broader criminal justice system. The choice of training data selected and pruned from a historical data lake resulted in an embedding of societal values in the trained model, according to which the Black community is at higher risk of reoffending than members of the White community. This case illustrates how even when a model operates in its “native” setting, with apparent continuity of context and purpose, it can still reproduce and intensify structural inequalities if the values and assumptions baked into its design go unchallenged.

II. Model Translation

In this scenario, a model is designed in one context but implemented in another. Two situations can occur: first, the system is transferred largely unchanged, reflecting techno-solutionist assumptions of universality and portability. In the second, developers adapt the model to local conditions – by translating interfaces, localizing data, or aligning with regulatory frameworks – while keeping its original goal intact. In this case, the goal of the model remains the same, but it is transferred to other conditions, for example translated into another language, integrated into different bureaucratic processes, or calibrated to reflect local norms or infrastructure. In this scenario, developers or implementers engage in processes of *adaptation*, such as localising data inputs or aligning with local regulatory frameworks, values and needs, to ensure that the model can function effectively in a new context. However, even

in this best-case scenario, what remains unchanged is the underlying scope of the system: the optimization goal, predictive goal, or administrative logic that justified its original development.

The value of this scenario may lie in its potential recognition of context without yet interrogating purpose. That is, while efforts – more or less explicitly – are made to translate it, the model’s goal is treated as universally desirable, or at least stable. From the point of view of the ontology of the ML system, this means to assume that given a fixed learning engine, an update of the training and test data will imply a trained model as efficient and appropriate to the desired task as the original. Keeping the goal or task of the system fixed means not to be aware of the impact that a change of context may imply on the trained model. A local alignment without goal re-evaluation may reproduce extractive or unjust logics under the guise of technical adaptation, and normally it is “often less about the artificial intelligence system itself and more about the activities surrounding it” (Berk 2020, 224). Most crucially, as noted by Heaven (2020, 5), these tools tend to exhibit reduced accuracy when deployed in contexts that differ from those in which they were originally trained. In this sense, Model Translation helps reveal the limits of technical adaptation when the foundational goals of a system may themselves require contestation or rethinking.

Healthcare provides a telling example (Asiedu et al. 2024). In 2024, Google announced licensing agreements in India and Thailand to deploy its diabetic retinopathy model, aiming for six million free AI-assisted screenings over the next decade (Philpott 2024). While the project involves local partners and ministries, it remains unclear whether adaptation extends beyond technical integration to interrogate the epistemological and ontological assumptions embedded in the model’s design.

III. Function Creep

This scenario describes the situation where a ML model is deployed within its original cultural or national context but is repurposed to serve a different function than initially intended. While the environmental and institutional settings remain familiar, the model's objectives are altered, often without adequate consideration of the new domain's specific requirements or ethical implications. This repurposing can lead to the model enforcing logics or practices that are misaligned with its new application, potentially resulting in systemic biases or injustices.

A prominent example is the Dutch childcare benefits scandal. In this case, the Dutch Tax and Customs Administration employed a machine learning algorithm to detect fraudulent claims for childcare benefits (Heikkilä 2022). The algorithm, originally developed by a private company for assessing credit risk in the banking sector, was repurposed to evaluate the likelihood of fraud among benefit applicants (Huyskes 2025). This shift from financial risk assessment to welfare fraud detection occurred without sufficient adaptation to the public sector's ethical standards and accountability requirements. This factor contributed to the severity of the system's social impact, which disproportionately flagged people with dual citizenship or a migrant background as high risk, resulting in unfair accusations, financial hardship, and social stigmatization for thousands of families. From the ontological point of view, this means to ignore that the function of the trained model emerges from its training, and from the specificity of its training data and tuning. In this sense, the application to a different context presupposes an alignment of the original and new application context, of their specificities and sensibilities to variables and correlations across test data. This assumption is often left implicit in the translation process. This case underscores the dangers of repurposing machine learning models without thorough evaluation of their suitability for new applications.

IV. Double Drift

In this final scenario, a model is deployed in a setting that diverges both in purpose and context from its original design. This dual misalignment can lead to significant challenges, as the model’s foundational assumptions may not hold true in the new environment, potentially resulting in ineffective or even harmful outcomes.

This scenario is particularly important as it highlights cases where models are deployed without meaningful translation, leading to epistemic friction, social misfit, or harm. Ontologically, double drift entirely ignores the nature of ML systems: the fact that their functionalities are generated from training; that tuning may strengthen properties emerging from this process; that testing is sensible to the local data, their selection and structure; that the output may amplify any bias along the process (Buda et al. 2025). In the double drift scenario, the general purpose nature of ML systems is not only assumed, but unjustifiably put to work.

For example, iSentry’s video analytics software was designed to detect anomalies in high-risk operational environments, originally developed for the Australian military. Today, that same technology has been repurposed for a very different context: urban civilian surveillance in South Africa (Hao and Swart 2022). The system now powers Vumacam’s Proof 360 platform, which monitors real-time footage from cameras across the city. Instead of assisting military personnel in combat zones, the software is used to flag “unusual” civilian behavior, such as running, lingering, or crossing virtual boundaries. What began as a military tool for situational awareness has thus transitioned into a mechanism of everyday urban control.

Translation Risks

Beyond these four scenarios, it is important to stress that even when AI systems are translated and adapted to new contexts, significant risks remain. Adaptation often prioritizes superficial alignment – such as language localization or regulatory compliance – while deeper cultural, epistemological, and ontological assumptions remain unexamined. This can reproduce dominant worldviews, marginalize alternative knowledge systems, and reinforce existing power hierarchies under the guise of inclusion (Birhane 2021; Okolo 2020). What we call translation is also frequently shaped by external actors and global market forces, leading to a form of “soft imposition”, where technologies appear context-sensitive but continue to serve the interests and logics of their origin rather than those of local communities (Couldry & Mejias 2019). Technologies are not simply adopted but actively reconfigured by users, who may appropriate services in unexpected ways (Maurer 2012). This highlights the need for local actors to go beyond technical contextualization: they must be able to question whether a model is needed at all, redefine its goals, and establish evaluative frameworks aligned with their own social and epistemic realities. Without such agency, AI risks becoming a new form of technological imperialism. Moreover, when infrastructures like cloud services or training platforms remain under external control, local communities’ ability to govern or reject technologies is further undermined.

Importantly, we tried to illustrate how each scenario can carry risks. Here, alignment between purpose and context may conceal the reproduction of entrenched assumptions and systemic biases. Often, since ML systems “fit”, their harms are harder to detect and critique, as their logics align with established norms. Continuous monitoring, auditing, and critical reflection are therefore essential, even when systems appear well-adapted.

Thus, the key issue is not merely achieving alignment between model, purpose, and context, but rather developing the capacity to critically interrogate and examine these elements itself – its exclusions, its underlying normative assumptions, its criteria for determining what counts as worth collecting, labeling, and classifying, and its definitions of what constitutes a problem worth solving. Without this deeper form of engagement, all travelling ML systems risk becoming mechanisms for naturalizing inequality, even when deployed with the best of intentions and under conditions of apparent translation.

Conclusions

This paper has illustrated how the social, epistemic, and normative conditions of ML development shape the ways models travel and take effect. Building on a Science and Technology Studies (STS) and the philosophy of computing, it proposed a framework for analysing these trajectories through two analytical dimensions – context and purpose. Crossing these axes, we identified four ideal-type scenarios of technological travel to show how values, assumptions, and problem framings embedded in models interact in different environments.

While this typology provides a useful lens for examining how machine learning models travel across sites, it is important to acknowledge its limitations. These four scenarios are intended as ideal types, not as exhaustive or mutually exclusive categories. By emphasizing the design phase and the originating conditions of models, the framework also risks underplaying the interpretative flexibility (Bijker & Pinch, 1984) and agency of receiving contexts, where models may be resisted, reinterpreted, or subverted. Finally, the scenarios should not be read as static or linear: models can move between them over time, and their classification depends on situated analysis rather than fixed definitions. Recognizing these limits is crucial to avoid reifying the very universalist assumptions we aim to critique.

Yet, the value of this framework lies in its diagnostic function: it centres the opaque dynamics through which assumptions, values, and problem framings embedded in ML models travel across settings. By foregrounding design-stage inscriptions and their interaction with subsequent deployments, the framework helps disentangle when adaptation constitutes meaningful translation and when it risks perpetuating epistemic violence under the guise of neutrality. Its value lies not in offering a final taxonomy, but in providing a structured way to interrogate how models move, change, or fail to change across sites. In a global landscape marked by deep asymmetries in power, infrastructure, and representation, this means reorienting AI not toward universal replication but toward situated, plural, and sustainable forms of technoscience.

Bibliography

- Adams, R. (2021). Can artificial intelligence be decolonized? *Interdisciplinary Science Reviews*, 46(1–2), 176–197. <https://doi.org/10.1080/03080188.2020.1840225>
- Akrich, M. (1992). The description of technical objects. In W. E. Bijker & J. Law (Eds.), *Shaping technology/ Building society: Studies in sociotechnical change* (pp. 205–224). MIT Press.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias. ProPublica. Available at: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Accessed on: 07/10/2025.
- Appadurai, A. (1993). Number in the colonial imagination. In C. A. Breckenridge & P. van der Veer (Eds.), *Orientalism and the postcolonial predicament* (pp. 314–339). University of Pennsylvania Press.
- Asiedu, M. N., Dieng, A., Haykel, I., Rostamzadeh, N., Pfohl, S., Nagpal, C., Nagawa, M., Oppong, A., Koyejo, S., & Heller, K. (2024). The case for globalizing fairness: A mixed methods study on colonialism, AI, and health in Africa. In Proceedings of the 4th ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '24) (Article 10, pp. 1–24). Association for Computing Machinery. <https://doi.org/10.1145/3689904.3694708>
- Behrends, A., Park, S.-J., & Rottenburg, R. (Eds.). (2014). *Travelling models in African conflict management: Translating technologies of social ordering*. Brill. <https://doi.org/10.1163/9789004274099>
- Berk, R. A. (2019). *Machine learning risk assessment in criminal justice settings*. Springer. <https://doi.org/10.1007/978-3-030-02272-3>
- Bijker, W. E., Hughes, T. P., & Pinch, T. (Eds.). (1987). *The social construction of technological systems: New directions in the sociology and history of technology*. MIT Press.
- Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPT-Ed*, 17(2), 389–409. <https://doi.org/10.2966/scrip.170220.389>
- Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. MIT Press.

boyd, d., & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Society*, 15(5), 662–679.

Buda, A. G., Manganini, C., & Primiero, G. (2025). A Philosophical Framework for Data-Driven Miscomputations. *Philosophies*, 10(4), 88.

Callon, M. (1984). Some elements of a sociology of translation: Domestication of the scallops and the fishermen of St Brieuc Bay. *The Sociological Review*, 32(S1), 196–233.

<https://doi.org/10.1111/j.1467-954X.1984.tb00113.x>

Chollet, F. (2019). On the measure of intelligence. arXiv preprint arXiv:1911.01547.

Chun, W. H. K. (2021). *Discriminating data: Correlation, neighborhoods, and the new politics of recognition*. MIT Press.

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

de Laet, M., & Mol, A. (2000). The Zimbabwe bush pump: mechanics of a fluid technology. *Social Studies of Science*, 30(2), 225–263. <https://doi.org/10.1177/030631200030002002>

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Facchini, A., Termine, A. (2022). Towards a taxonomy for the opacity of AI Systems. In: Müller, V.C. (eds) *Philosophy and Theory of Artificial Intelligence 2021. PTAI 2021. Studies in Applied Philosophy, Epistemology and Rational Ethics*, vol 63. Springer, Cham.

https://doi.org/10.1007/978-3-031-09153-7_7

Floridi, L., & Sanders, J. W. (2004). The method of abstraction. *Yearbook of the Artificial. Nature, Culture and Technology. Models in Contemporary Sciences* 177–220.

Fresco, N., & Primiero, G. (2013). Miscomputation. *Philosophy and Technology*, 26, 253–272. <https://doi.org/10.1007/s13347-013-0112-0>

Gabriel, I. (2020). Artificial Intelligence, Values, and Alignment. *Minds and Machines*, 30(3), 411–437. <https://doi.org/10.1007/s11023-020-09539-2>

Gammelgaard, A. Ø., Pedersen, C. G., Kaspersen, E. S., Thomsen, M. R., Samson, J. K., & Fabricius, A. H. (2023). Decolonizing Language Resources in the Human-Machine Era: A Critical Reflection on Universalist Ideologies and Language Technologies. *Interventions*, 26(8), 1189–1210. <https://doi.org/10.1080/1369801X.2023.225280>

Hao, K., & Swart, H. (2022). How facial recognition is transforming surveillance in South Africa. *MIT Technology Review*.

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575–599.

Hassan, Y. (2023). Governing algorithms from the South: A case study of AI development in Africa. *AI & SOCIETY*, 38(4), 1429–1442. <https://doi.org/10.1007/s00146-022-01527-7>

Heaven, W.D. (2020). Predictive policing algorithms are racist. They need to be dismantled. July. *MIT Technology Review*, 17: 1-16. Available at: <https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machinelearning-bias-criminal-justice/>. Accessed on: 12 May 2025.

Heikkilä, M. (2022). What Europe can learn from the Dutch childcare benefits scandal. *MIT Technology Review*. Available at: <https://www.technologyreview.com/2022/04/04/1048622/what-europe-can-learn-from-the-dutch-childcare-benefits-scandal/>. Accessed on: 12 May 2025.

Hughes, T. P. (1994). Technological momentum. In M. R. Smith & L. Marx (Eds.), *Does technology drive history? The dilemma of technological determinism* (pp. 101–113). MIT Press.

Jasanoff, S. (Ed.). (2004). *States of knowledge: The co-production of science and social order*. Routledge.

Lagioia, F., Rovatti, R. & Sartor, G. Algorithmic fairness through group parities? The case of COMPAS-SAPMOC. *AI & Society* 38, 459–478 (2023). <https://doi.org/10.1007/s00146-022-01441-y>

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.

Lu, M., & Qiu, J. L. (2023). Transfer or translation? Rethinking traveling technologies from the Global South. *Science, Technology, & Human Values*, 48(2), 272-294.

<https://doi.org/10.1177/01622439211072205>

MacKenzie, D. (1990). *Inventing accuracy: A historical sociology of nuclear missile guidance*. MIT Press.

Madianou, M. (2025). *Technocolonialism: When technology for good is harmful*. Cambridge, UK: Polity Press.

Mager, A., Eitenberger, M., Winter, J., Prainsack, B., Wendehorst, C., & Arora, P. (2025). Situated ethics: Ethical accountability of local perspectives in global AI ethics. *Media, Culture & Society*. SAGE Publications. <https://doi.org/10.1177/01634437251328200>

Maurer, B. (2012). Mobile Money: Communication, Consumption and Change in the Payments Space. *The Journal of Development Studies*, 48(5), 589–604.

<https://doi.org/10.1080/00220388.2011.621944>

Milan, S., & Treré, E. (2024). Against decolonial reductionism: The impact of Latin American thinking on the data decolonization project. *Big Data & Society*, 11(4), 20539517241270694. <https://doi.org/10.1177/20539517241270694>

Mohamed, S., Png, M. T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4), 659–684. <https://doi.org/10.1007/s13347-020-00405-8>

Okolo, C. T. (2020). AI in the “Real World”: Examining the Impact of AI Deployment in Low-Resource Contexts. arXiv preprint arXiv:2012.01165.

Pham, H., Zhang, M., & Zhang, C. (2023). Fairness and accuracy under domain generalization. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '23), 585–596.

Primiero, G. (2019). *On the foundations of computing*. Oxford University Press.

Primiero, G. (2016). Information in the philosophy of computer science. In L. Floridi (Ed.), *The Routledge handbook of philosophy of information* (pp. 558–573). Routledge.

<https://doi.org/10.4324/9781315757544.ch10>

Rajalakshmi, R., Subashini, R., Anjana, R. M., Mohan, V., & Deep learning team. (2018). Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *JAMA Ophthalmology*, 136(10), 1080–1086.

Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 20(4), 350–365.

Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics*, 3(3), 1095–1115.
<https://doi.org/10.1007/s43681-022-00218-9>

Rogers, E. M. (1983). *Diffusion of innovations*. Free Press.

Rottenburg, R. (2024). Translating technology in Africa. In R. Rottenburg, F. Ballim, & B. Kotzen (Eds.), *Translating Technology in Africa. Volume 1: Metrics* (pp. xii–xxi). Brill.
<https://doi.org/10.1163/9789004678354>

Suchman, L. (2011). Anthropological relocations and the limits of design. *Annual Review of Anthropology*, 40, 1–18.

Suresh, H., & Gutttag, J. V. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. Proceedings of the ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21), 1–9.

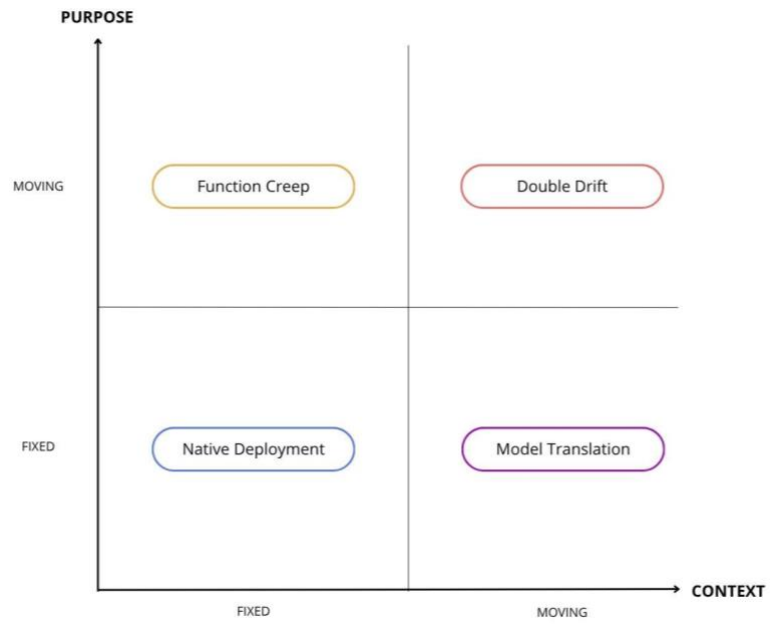


Figure 1: The four scenarios of technological travel mapped along the axes of context and purpose