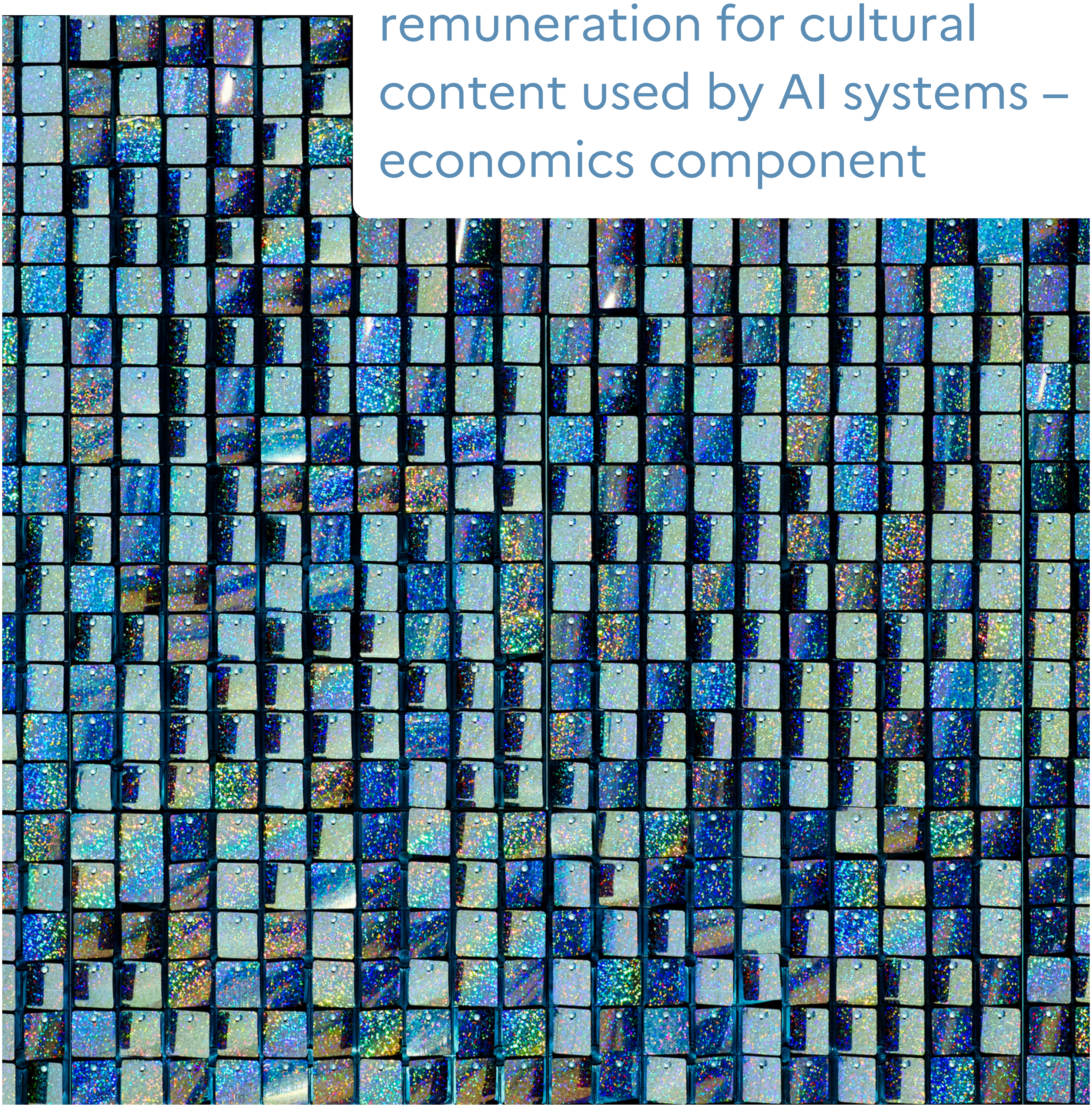
MINISTÈRE
DE LA CULTURE
Liberté
Égalité
Fraternité

Task force report on remuneration for cultural content used by AI systems – economics component



PRESENTED TO THE CONSEIL SUPÉRIEUR DE LA PROPRIÉTÉ LITTÉRAIRE ET ARTISTIQUE [SUPERIOR COUNCIL OF LITERARY AND ARTISTIC PROPERTY]

Task force co-chair : Joëlle Farchy
Co-rapporteur : Bastien Blain

Task force co-chair
Joëlle Farchy
Professor at Paris 1 Panthéon - Sorbonne
Honorary member of the CSPLA

Co-rapporteur
Bastien Blain
Assistant Professor at Paris 1 Panthéon - Sorbonne

Report presented to the plenary meeting of the CSLPA, June 23, 2025
Its content is the sole responsibility of its authors

A document focusing on the legal aspects of this issue has been developed by Professor Alexandra Bensamoun, task force co-chair and expert member of the CSPLA, and Julie Groffe-Charrier, associate professor and co-rapporteur.

Remuneration for cultural content used by AI systems

Economics report

In France, the Minister of Culture asked the Conseil Supérieur de la Propriété Littéraire et Artistique [Superior Council of Literary and Artistic Property] (CSPLA) to create a task force "relating to the remuneration for cultural content used by artificial intelligence systems".

In a letter dated April 12, 2024, the chairmanship of this task force was assigned, for its legal component, to Professor Alexandra Bensamoun and, for its economics component, to Professor Joëlle Farchy. Subsequently, the rapporteurs for the task force, respectively Julie Groffe-Charrier and Bastien Blain, sent out a questionnaire. Responses to this questionnaire served as the basis for initial reflection.

The economics report presented here does not aim to provide definitive solutions to all the issues raised. Its objective is to lay the groundwork which will inform the phase of in-depth analysis that is now beginning, with a view to future work.

The content of this report is the sole responsibility of its authors.

Joëlle Farchy

With Bastien Blain

June 2025

Executive summary

The development of AI can be a source of exciting innovations for the future of creation. However, in the cultural sector, beyond philosophical or environmental questions, economic concerns are emerging, as certain uses of works are carried out without the consent and remuneration of copyright holders. To perform the various tasks necessary for their operation, AI systems require vast datasets, some of which are protected by intellectual property. Faced with these concerns, pragmatic solutions for valuing this data—adapted to very large parameters while respecting the general principles associated with intellectual property rights—must now be found. The objective is to value data-works within an ecosystem that guarantees both the circulation of European works in AI systems and the sustainability of their funding.

To clarify these issues, this report is structured around three main questions. Why are value transfers necessary? How should value transfers be organised? And what could be the rules for determining the sums of these envisaged value transfers?

1 - Why value transfers?

The first part of the report aims to highlight the economic reasons likely to justify value transfers between AI operators and intellectual property rightholders.

AI operates in a circular movement: AI systems need to "feed" on multiple data-works protected by intellectual property to obtain results while, at the same time, producing a certain number of "outputs" that we call "synthetic quasi-works" which, in turn, feed the AI models (on this movement, see diagram Part 1).

The novelty here, compared to other forms of non-consensual use of protected works, lies in the fact that numerous results generated by AI directly compete with human creations used in their development. It is no longer just individual authors who see their intellectual property rights infringed; we are witnessing a macroeconomic destabilisation of a set of professional fields linked to the risk of the replacement of human works. A dystopian world in which only synthetic works would be available is, in theory, conceivable. The act of self-consumption—corresponding to the circularity of the movement in which AI feeds on its own synthetic productions—would lead to the creation and existence of only "mad," synthetic quasi-works, which would all end up resembling one another. The consequences of such a world would be twofold.

First, this lack of diversity would be detrimental to the disruption processes that have marked the entire history of artistic activity. Further, the replacement of human creation by AI would lead to an internal contradiction in the AI models themselves, and to their possible collapse—that is, to the degeneration of models, if they are no longer (or poorly) fed by new human creations. The short-sightedness of economic actors and a short-term vision of the markets could lead to a failure to grasp this dual issue. In the absence of funding, the incentive to create and produce new, high quality, and diverse human works could one day dry up.

Value transfers between AI operators, who are at the source of true disruption, **and cultural actors** guaranteeing the future of human creation, **therefore aim to ensure incentives for investment in creation**; these investments are **necessary**

for the future of that which ensures the nobility of culture, as well as for supporting innovation in AI.

Recommendations

No. 1 – Explain and publicise the joint interest of cultural actors and AI operators, to invest in a sustainable ecosystem that guarantees both the presence of European works in AI systems and the sustainability of their funding.

No. 2 – Implement and/or consolidate support and training policies adapted for the professions most directly impacted by the rise of AI.

2 - How should value transfers be organised?

The second part of the report indicates the possible framework for implementing value transfers.

In terms of copyright, we usually distinguish schematically between:

- On one hand, *market-based solutions* concerning the contractualization of rights, which correspond to the "full" exclusive right with a monopoly on authorisation and the right to remuneration (regardless of whether the management of rights is exercised individually or collectively);
- On the other hand, *mandatory transfers* in which the ability to authorise or prohibit is removed, while the capacity for market contractualization gives way to solutions organised under the supervision of public authorities.

With AI, under current conditions, in France as in many other countries, direct negotiations between stakeholders—in other words, the first solution—are proving rare. This is why mandatory transfers are occasionally considered.

Given the limits of negotiations within a context of scattered initiatives, and taking into account the lack of responsiveness to market developments inherent in mandatory transfers, the report proposes **a complementary and optional path, that of collective support for the structural conditions facilitating the formation of a marketplace**—that is, a structured space for exchange, allowing contractualization in respect of sectoral specificities and without the need for legislative changes. The objective is to bring together, in the same digital technical space and in the case of defined catalogues or parts of catalogues, a triple activity of access/authorisation/remuneration. The aim is not to create ex nihilo a completely new system but rather, to rely on recognised expertise and skills and the varied missions of the different stakeholders in place (collective management organisations, publishers, producers) to **offer AI operators an integrated and clear supply, and to play a role in accelerating negotiations between culture and AI. Unlike a single market or a mandatory management system, participation in the marketplace would be solely on a voluntary basis.** While the expected benefits are numerous, rendering this solution operational requires that a certain number of sensitive issues be debated and decided among the stakeholders, particularly concerning the financing conditions.

Such market-based solutions do not preclude the opportunity for, in specific situations, other, more binding legal mechanisms. Aside from developments within the framework of intellectual property, other mechanisms are indeed conceivable. Fiscal mechanisms first, such as earmarked taxes on turnover inspired by existing models; funding obligations following the economic logic requiring that which is downstream to finance that which is upstream; or even the establishment of a support fund for the benefit of human creation. These avenues remain to be explored.

Recommendations

No. 3 – In the context of a consultation between cultural actors and AI operators, consider the opportunity and feasibility of constructing a marketplace—a structured exchange space—allowing for contractualization while respecting sector-specific characteristics.

No. 4 – In the context of a consultation between cultural actors and AI operators, explore the opportunity and feasibility of compensation mechanisms and value transfers in addition to those provided by intellectual property.

3 – How should the envisaged value transfers be quantified?

Whatever framework in which the transfers will eventually be organised, the question of the value of the data-works will arise. These questions are addressed in the last two sections. The report highlights the importance of considering value transfers, in the case of AI, with different methodologies depending on what one seeks to determine. First, the basis of these transfers—that is, the sites of *value creation* by AI operators; second, the proportion allocated to culture during the *distribution* between AI operators and cultural actors; and finally, sharing among the works and the rightholders within cultural sectors.

3.1 Value creation and the basis of transfers

At this stage, the objective is to analyse the value chain, and to precisely identify the companies and activities connected with the use of protected works. The value chain of an AI system goes through a development stage and a deployment stage.

The development of an AI model first relies on a pre-training phase followed by a fine-tuning period. This last consists of specialising the foundation model by retraining it on specific data or tasks. Inference corresponds to the operation of putting the model into production—in other words, the process by which a previously trained model will produce a result: predictions on new data. The inference can be supplemented by the contribution of fresh data, so that the model provides information considering current events, or very specific data that it will search for in an external source. This is called retrieval augmented generation (RAG) or the *grounding* of the model.

Thus, three main categories of data are used in the development of models:

1) Training data, which is plentiful, to the order of several million or several billion;
2) Fine-tuning data, which is specialised and can be available on the internet, or carefully selected by a company or organisation;
3) “Fresh” data, which consists of grounding the model in the latest events or new data without the need for training.

The influence of a dataset is different at each stage. Removing a dataset in the training phase only weakly influences the model's performance, as it is trained on an immense quantity of data. For fine-tuning, a dataset relevant to the model's use is crucial; if the dataset is not relevant, its value is null. The same goes for grounding.

The development phase is not a source of value creation in itself; it is the deployment phase that creates value when the model is integrated into a system that can be commercialised. Once the models are developed, they are published, meaning they are available for deployment. The activities resulting from the development can then be monetised towards services and end-users (businesses and individuals). Based on the models published at the end of the development phase, the financial returns (see Figure 5) mainly come from two major categories of activities: first, the creation of software and applications to interact with the models; and second, fine-tuning to specialise the models.

Two main types of companies perform these two activities. On one hand, there are operators who act in the market for developing foundation models. In this market, which is dominated by an oligopoly of American companies, the monetisation of services to which the models provide access is now almost systematic, whether for businesses or individual clients. Further, there are third-party operators, business users, or intermediaries acting on behalf of a client. By paying the developers of foundation models, companies thus create their own applications which, in turn, they bill to their clients.

This observation of an ecosystem under formation, the deployment phases of which allow a user, upon request, to produce a result, are sources of various valuations, leading to the idea that **the basis for remuneration should be both broadened and refocused**: refocused on deployment activities where value is created, rather than on development activities; broadened to companies that are not only the few giants at the origin of foundation models, but also to those that create other activities during deployment. Indeed, to carry out their commercial activities, some companies have an imperative need for quality cultural data to perform various actions during deployment, notably for specialisation and freshness to the proposed results.

Recommendations

No. 5 – Carry out, with the services of the Ministry of Economics and Finance, a precise mapping of the sites of *value creation* and relevant markets; and follow the value chain in the deployment phase to provide the basis for value sharing.

3.2 Value sharing between AI providers, cultural actors and within cultural sectors

The objective is, first, to help determine the levels of value sharing between AI operators and cultural actors. This task force recommends that the rates practised be proposed by rightholders to AI actors based on categories of uses and users. **We therefore propose prioritising the intended purpose of an AI system downstream to appraise the value of the data upstream.** The application of the intended purpose criterion—already commonly practised in intellectual property matters to adjust levels of remuneration—should not pose any problems of principle.

However, in AI matters, the idea of adjusting prices according to the intended purpose of the input data (pre-training, fine-tuning, RAG, etc.) is sometimes raised, since the value of certain data depends on a given case (see above). In our opinion, such mechanical solutions should not be retained for operational reasons; they would undoubtedly lead to spillover effects and opportunistic behaviours. For example, in the absence of clarity on the intended purpose of the input data, it would not be possible to prevent an actor from having access at a low price to "training" data only to then exploit these same data for fine-tuning purposes, at a stage in which the data should be much more expensive.

This is why valuation according to intended purpose is not centred on the input data but rather on the visible results produced by the models, systems, or applications. It involves distinguishing, within the value chain of AI systems and their applications, between those whose activity is directly linked to the use of works, and those who have access to the works but whose AI system has an intended purpose that is not necessarily related.

With the degree of distribution corresponding to the exploitation of data, works would thus be based on an economic presumption of use according to the intended purpose of the model. The intended purpose criterion provides a guideline for distinguishing, very schematically, three major categories of remuneration levels to carry out the distribution between cultural actors and AI operators.

- Intended purpose 1: generalist models - basic remuneration levels;
- Intended purpose 2: specialised cultural and media models, without competition on outputs - intermediate remuneration levels;
- Intended purpose 3: specialised cultural and media models, with competition on outputs - high remuneration levels.

Based on these categories, a continuum of pricing levels could be established by the cultural actors themselves according to annual licences, renegotiable each year with AI operators, for example, within the proposed marketplace. **The visible intended purpose criterion of activities that feed on protected data thus provides a first basis for establishing a scale of distribution.**

More precisely, different quantification methods are presented in the report to help approximate the valuation of a dataset of works. A wealth of new literature dedicated to the notion of data attribution consists of calculating the marginal contribution of each dataset to the performance of the model in general, and to the generation of a particular result (an output) following a user's request. Three approaches co-exist. The first consists of changes in the parameters of the models (whether by training the models on subsets of data or by altering the parameters of the already trained model) to establish causal links. The second, which is correlational, seeks to measure the similarity between the result generated by the model and the elements constituting the training data set. The third, which is causal and proactive (and which cannot be applied to already trained models), corresponds to watermarking the ingested data.

In terms of value sharing, quantification techniques are operational in limited cases: inoperable on foundation models but potentially more so on specialised models, provided they move from proof-of-concept to operationality on use cases or, for some of these techniques, to prove the use of works.

Finally, the objective of a final step, that of redistribution, is to share out the amounts from the previous step within the cultural sector itself, among the multitude of different works and rightholders. **It is at the level of redistribution that quantification techniques, with their respective advantages and disadvantages, seem the most promising**; they would thus help ensure that all cultural actors benefit from value sharing commensurate with their contribution.

Recommendations

No. 6 – Refine the operationality of the criterion of economic presumption of use according to the intended purpose of the results produced by the models, systems, or applications that utilise protected data, to produce a *scale of value sharing*.

No. 7 – In collaboration with the Centre of Expertise for Digital Platform Regulation (PEReN), delve deeper into case studies on the operationality of scientific quantification methods to prove and/or evaluate the contribution of certain data-works to the results produced and/or to the overall performance of specialised models. Promote among cultural actors and AI operators the solutions deemed most relevant according to specific cases.

To perform the various tasks necessary for their operation, AI systems require multiple datasets. Among these, there may be data-works protected by intellectual property. However, the voracity of these models should not be an argument for the generalised, unbridled consumption of all protected works. Cultural stakeholders express the concern that non-compliance with intellectual property rules could lead to a situation where funding for creation is no longer guaranteed.

Intellectual property law has been designed around the notion of the work, considered in its unity and singularity. Traditional instruments of authorisation and remuneration often seem inappropriate in the face of the volumetric approaches of AI, which disrupt authorisation mechanisms designed for the individual exploitation of specific objects, seeing as the notions of data and content obey inverse movements of permanent flow and large masses within the digital economy (Benabou, 2018).

This is why **pragmatic solutions must be found, adapted to vast parameters and respecting the general principles associated with intellectual property. The objective is to valorise European data-works within an ecosystem that guarantees both their circulation and the sustainability of their funding**. In a context of intense international competition, innovation in AI has become a major issue of competitiveness in terms of industrial sovereignty but also cultural sovereignty. Only a broad marshalling of European ideas and content will effectively limit the major cultural risk that AI systems may ultimately propose content from which European cultural references would be entirely excluded.

To clarify the issues, this report is structured in four parts. In the first part, we show why the incentive to invest in human creation, through value transfers, is necessary both for cultural stakeholders and to support innovation in AI. The second part indicates the possible framework for implementing value transfers. After presenting in a third section the stages of the value chain and the sites of value creation for AI system stakeholders, we address, in the final part, the valuation of data-works for AI systems, as well as guidelines around value sharing to the benefit of cultural actors.

1 - Maintaining the cultural heritage of humanity to preserve it from the disease of the “mad work”

In a globalised and competitive market, the notion that rightholders can share in the value of AI systems is not always clearly grasped. It is undeniable that AI offers a multitude of exciting opportunities for the creators of the future. **The situation we are interested in this report is not that of the relations between AI and culture in general, but that in which operators need, for various purposes, to use certain data protected by intellectual property rights.**

As representatives of OpenAI pointed out during an inquiry by the Communications and Digital Committee of the House of Lords in the United Kingdom (House of Lords Communications and Digital Select Committee inquiry, 2023): “We believe that AI tools are at their best when they incorporate and represent the full diversity and breadth of human intelligence and experience. To do this, today's AI technologies require a large amount of training data and computation, as models review, analyse, and learn patterns and concepts that emerge from trillions of words and images. OpenAI's large language models, including the models that power ChatGPT, are developed using three primary sources of training data: (1) information that is publicly available on the internet, (2) information that we license from third parties, and (3) information that our users or our human trainers provide. Because copyright today covers virtually every sort of human expression—including blog posts, photographs, forum posts, scraps of software code, and government documents—**it would be impossible to train today's leading AI models without using copyrighted materials.** Limiting training data to public domain books and drawings created more than a century ago might yield an interesting experiment, but would not provide AI systems that meet the needs of today's citizens.”

This first section therefore aims to highlight the economic reasons likely, in this specific case, to justify value transfers between AI model operators and intellectual property rightholders.

1.1 - From inspiration to substitution: the fear of the great replacement

It is possible to consider that, given the mass of input data used by the models, beyond classic questions of exploitation of a work, we are in the presence of an act of “inspiration,” like a human who reads hundreds of books or listens to thousands of pieces of music to create or compose themselves.

However, for most cultural stakeholders, the use of protected works by an AI is not comparable to the reminiscences of other works that fuel human creation. While an author is often inspired by the creations of others (without this always giving rise to copyright), the systematisation and the scale, both quantitative and qualitative, constituted by artificial intelligence completely changes the scope of the phenomenon. The legal questions of respect for the monopoly of exploitation and licensed access to works are addressed elsewhere in the report. The economic question on which we concentrate here is whether certain acts of utilisation[1] of protected works to train or improve the performance of an AI model must be remunerated based on any potential loss of value incurred.

In the case of actions carried out by AI systems, the real disruption lies not only in the change in scale of the actions performed, or in the mass or technical access to the works, but in **the fact that many AI-generated results resemble what might be called "quasi-works"** (Benabou, 2023) **and directly compete with the human creations that were used in their development (see Figure 1, point 1) which, in the long run, could undermine the very conditions for human creation.**

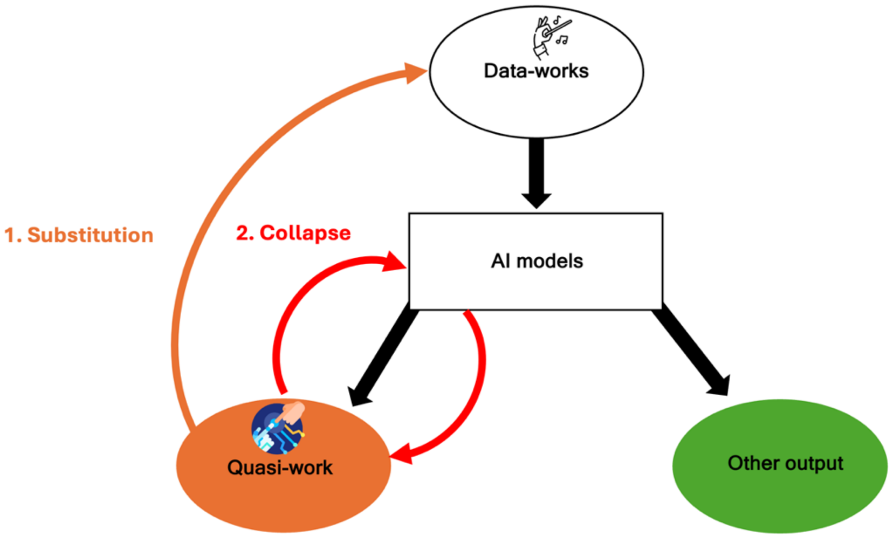


Figure 1. Circuit for fuelling AI models with data-works

The potential for a "great replacement" of humans by machines thus makes the issue of value transfer especially acute. It is important to note that the loss of value incurred when works are used for AI is not just the classic and direct one affecting an individual author who is "copied" when the output closely resembles the ingested input, or that of the loss suffered by an economic actor who has invested in the creative process. The loss also stems from the degree of substitutability between human and synthetic works at a macroeconomic level. In tangible terms, it corresponds to the effects of competition on entire professional sectors and on the future of those sectors.

The crowding-out effects on human creation are felt first through price competition, as AI allows for the creation of outputs at a faster rate and at a lower cost than humans. This is the case, for example, with translations, which have dropped from around 21 euros per 1,500-character page to 17 or 18 euros (Vulser, 2024). Crowding-out effects also occur through volume. The overabundance of AI-generated content risks saturating the market and, as a result, reducing the visibility and accessibility of human-created works. This is the case, for instance, with the proliferation of unauthorised vocal clones,

[1] The term "utilisation" used in this report is intended to be legally "neutral," without referring to a specific qualification. Not every use of a work corresponds to the notion of "exploitation" within the meaning of the Intellectual Property Code, justifying remuneration under copyright law.

particularly on YouTube and TikTok, where they cause an economic issue by diverting attention away from official recordings.

The effects of AI on employment have been the subject of various studies. These effects appear limited in the short term, but 33% of jobs in advanced economies could eventually be replaced by AI, according to an IMF report (Cazzaniga et al., 2024). In March 2023, researchers from OpenAI, OpenResearch, and the University of Pennsylvania estimated that large language models (LLMs) could affect the professional tasks of 80% of the workforce in the United States (Eloundou et al., 2023). Furthermore, researchers from the Massachusetts Institute of Technology, the London School of Economics, and Boston University detected a negative correlation between AI adoption and job recruitment between 2010 and 2018: for each 1% increase in AI use, firms exposed to AI reduced their hiring by about 1% (Acemoglu et al., 2022). According to other estimates by the French Treasury, focusing more specifically on the emergence of foundation models, while 80% of American workers could see at least 10% of their tasks replaced, only 19% of them could see that share reach 50% or more, thus facing a significant risk of substitution (Besson et al., 2024).

AI adoption would disproportionately threaten the most highly qualified professions (higher education graduates with high salaries), as it can substitute for certain highly skilled workers in tasks that require advanced competencies. AI is capable of handling abstract, non-routine cognitive tasks, thereby expanding the range of tasks that can be substituted (for instance, translation, diagnostics). Among skilled professions, not all would be affected equally. Companies might be more likely to reduce staff in professions focused on writing and programming, which are more exposed to the risk of replacement by generative models.

In cultural sectors, some professions are already feeling particularly threatened. The profession of **translator**, in its current form, is being called into question. Translators are receiving fewer requests for complete translations and more for services that involve correcting translations produced by AI systems such as DeepL. This post-editing work is sometimes viewed as more time-consuming and results in lower pay, according to a December 2022 survey conducted by Association of Literary Translators of France (ATLF) among 400 people (*L'ATLF a Interrogé Ses Adhérents Sur La Post-Édition*, 2022). Indeed, remuneration is lower and, for 68% of respondents, it was even below the average translation rates (Vulser, 2024). While literature, which accounts for less than 10% of editorial production in France each year, is relatively spared, works that leave less room for interpretation are more affected. Conversely, in the world of comics and webtoons (comics in smartphone format), AI tools like GeoComix are being used to translate speech bubbles into multiple languages. The same applies to audiobooks, with AI systems being used by HarperCollins, while Audible offers many books whose voices are generated by AI (Cohen, 2024). Additionally, many websites are now translated using AI (Thomson et al., 2024).

Graphic designers are also seriously threatened by systems like Midjourney, which can generate, for instance, science fiction illustrations. The Society of Authors in the United Kingdom estimated in a survey that about a quarter of illustrators have already lost work due to generative AI, and more than a third (37%) reported a decline in income attributable to generative AI (*SoA Survey Reveals a Third of Translators and Quarter of Illustrators Losing Work to AI*, *The Society of Authors,* 2024).

As for **writing professions** such as **journalism**, a recent study estimated that reporter and journalist jobs are among the most exposed to AI systems (Eloundou et al., 2023). Several media outlets, including BuzzFeed, News Corp Australia, and G/O Media, have already integrated generative AI into their content production. In early 2023, BuzzFeed launched quizzes powered by ChatGPT, travel articles, and a recipe recommendation chatbot called Botatouille (Chin-Rothmann, 2023). At the same time, Google offered national and local media an AI tool called *Genesis*—a generative chatbot capable of drafting headlines, social media posts, and articles—which was presented as a productivity-enhancing tool.

As a result, more than one hundred news and information websites that are fully or almost entirely AI-generated have been identified since 2023 (Sadeghi et al., 2024). Moreover, many books are now written using AI systems, as evidenced by the high number of self-published e-books on platforms like Kindle, where publication is now limited to three titles per day per author (*Update on KDP Title Creation Limits*, 2023).

For voice actors, a study by the Audiens Data Lab quantified the professions in France threatened by AI-based dubbing tools. In 2023, this industry involved 110 companies and employed 7,397 freelance performers and 3,116 permanent staff. AI systems such as HeyGen, Eleven Dubs, or Deepdub make it possible to clone voices and translate videos into multiple languages while synchronising lip movements. Their use could eliminate the need for in-studio dubbing of films, series, video games, and animated content by professional actors. This automation and localisation process could result in massive loss of activity in the sector (Thomas, 2023). Concern is not limited to French voice actors; video game voice actors have also rallied together, for example, in California. 2,600 artists who perform voiceovers or whose movements are used to animate synthetic characters may see AI systems replicate their voices, or create digital doubles of stunt performers, without their consent or remuneration (*Worried About AI Use, Video Game Actors and Voice Artists to Strike in California*, 2024).

In the case of dubbing, the issue of competition between human and AI voice actors—fuelling strong reactions from the profession—intersects with broader questions of national industrial policy. Should AI dubbing become widespread, it could jeopardise a sector in which France has historically excelled. Since the advent of talking cinema, Hollywood has aimed to dominate the global dubbing market from within the U.S., but has never succeeded, as France developed a high-performing dubbing industry. However, since the main players in AI dubbing (such as OpenAI and Eleven Labs) are American, there is a strong likelihood that AI dubbing will be centralised in American studios on American soil.

Beyond the debate on industrial sovereignty, the broader issue—balancing short-term employment threats with long-term human–machine collaboration—is far from resolved.[2] In a Schumpeterian perspective, it reflects the macroeconomic process of "creative destruction" inherent to any innovation. In the short term, during an essential transition period, **social support policies that take into account existing expertise and know-how will be indispensable**. Moreover, as many professions are not disappearing but rather undergoing profound transformation, **training** is essential, so that tomorrow's professionals can take control of these innovations rather than simply being subjected to them. For this reason, the task force alerts public authorities to the urgency of establishing or reinforcing appropriate systems for the sectors directly affected.

[2] In an open letter published in Le Monde dated September 10, 2024, a quote from American translation scholar Alan Melby indicated that AI will not make the profession of translator disappear "except perhaps for those who already translate like machines" (*'No, artificial intelligence will not replace translators!'*, 2024).

1.2 – Degeneration of models trained on synthetic data

Beyond the short-term impacts on the cultural sector, the **risk of replacing human creation with AI could also lead to an internal contradiction within AI models themselves, and to their potential collapse in the medium or long term.**

Once a model is trained on real, human-produced data, it can be used to generate new data, known as synthetic data. Synthetic data is designed to mimic the statistical and structural properties of real data, while being artificially generated. In the case of current generative AI systems, this refers to data created by a model originally trained on human-generated inputs, such as landscape images generated on demand. In principle, a model can be trained entirely on synthetic data or on a combination of synthetic and human data. In other words, the output data from a model trained on human-created content can be used as input for a new model. However, proceeding in this way may lead to the generation of low-quality data, resulting in the collapse of the model itself (see Figure 1, point 2).

The impact of training AI models on synthetic data (rather than on human-produced data) on the quality of the outputs has been the subject of various studies. These studies use measures of result quality (see box 1 in the appendices) and identify sources of error that may cause this collapse (see box 2 in the appendices).

In a recent study published in the prestigious journal *Nature*, researchers from Oxford and Cambridge (Shumailov et al., 2024) demonstrated the degradation of output quality from a model fine-tuned on synthetic data. The term "collapse" here refers to a generative process in which the quality of model outputs is degraded because the data generated by a first generation of models pollutes the training data for the next generation. In the study, the researchers fine-tuned a language model (an LLM) originally pre-trained on a dataset from Wikipedia (human-created data), which they then used to generate new articles (synthetic data). They trained the next generation of the model on these new articles rather than on real data, and so on. When evaluated, the newly trained models quickly showed significant errors compared to the original model trained on real data. In other words, the quality of the newly generated data collapsed once the models had been trained on synthetic rather than real data.

This collapse occurs because each model relies solely on synthetic data, which leads to an overemphasis on common words and a neglect of rare ones. With each iteration, the model increasingly learns from its own erroneous predictions, amplifying errors until it ends up learning almost entirely from incorrect information. The collapse corresponds to a progressive loss of information about the real distribution of data; the models become less and less capable of producing diverse outputs, and the variance of their output distributions shrinks. Rare or unlikely events disappear from the model's knowledge as it continues to train on its own data. At the same time, over successive generations, highly improbable data is generated, data that the original model would never have produced: aberrant data. Collapse manifests through increased repetition and the introduction of errors. The model ultimately fails to perceive the underlying data distribution correctly, focusing instead on its own increasingly inaccurate projections of the world and generating aberrant data.

A study by researchers from Stanford and Rice University confirms that the same phenomenon occurs in the context of image generation (Alemohammad et al., 2023), which is based on diffusion models (the "equivalent" of large language models for images). In this study, the researchers study the impact of training loops for image generation models, which differ in how they combine real data and synthetic data, and show that quality and diversity collapse when only synthetic data is used.

A similar study conducted by researchers from Stanford and Berkeley (Bohacek & Farid, 2023) confirms this result for another type of model that allows image generation, the popular Stable Diffusion (Rombach et al., 2022). Bohacek & Farid, taking the example of face generation, show that the images produced by the base model are of excellent quality; but the collapse of their quality is observed as soon as the percentage of synthetic data used for training exceeds 10%.

The risk of collapse is to be taken seriously insofar as data is often extracted from the internet, which is increasingly filled with synthetic data (Alemohammad et al., 2023), such as images, reviews (Gault, 2023), websites (Cantor, 2023) or annotated data (Veselovsky et al., 2023). Certain popular image databases contain synthetic data whose use is occasionally by design, sometimes linked to the lack of accessible data as in medicine (Pinaya et al., 2022) or in geophysics (C. Deng et al., 2022) or due to the protection of private medical data (Klemp et al., 2023; Luzi et al., 2024).

Furthermore, the recourse to synthetic data could become a consequence of the scarcity of data created by humans (Villalobos, 2022). Assuming that current rates of data consumption and production are maintained, **real data will be lacking**. Research conducted by Epoch AI predicts that "we will have exhausted the stock of low-quality textual data by 2030 to 2050, high-quality textual data before 2026, and visual data between 2030 and 2060." (Villalobos, 2022). It is, moreover, not possible to train models further on existing data because of the risk of "overfit" (i.e., training the model to explain stochastic variations specific to the training dataset and to it alone)—that is to say, the risk of altering the generalisation of the model and therefore its capacity to generate new data.

Faced with these proven risks, how model collapse be prevented? A study shows that, in the context of replacing real data with synthetic data, collapse will not occur if the initial generative models sufficiently approximate the distribution of real data and if the proportion of real data is sufficiently large compared to synthetic data (Bertrand et al., 2024). Similarly, (Dohmatob et al., 2024) suggest that the careful choice concerning the quality of real data mixed with synthetic data can avoid model collapse. Other researchers (Alemohammad et al., 2023) suggest, however, that while selecting good quality images from synthetic data before each training avoids degradation in terms of quality of data generated by the model, this nevertheless leads to a reduction in the diversity of generated data (while not selecting at all leads to collapse on both these aspects).

If investigations on model collapse always include training of the initial generation on real data, they diverge on the approach adopted to train subsequent generations.

The most extreme approach consists of a total "replacement" of real data with synthetic data in training subsequent generations, which leads, in all studies, to collapse.

For other approaches, studies differ on the proportion of real data and synthetic data in training the next generations (a fixed proportion of synthetic data versus an increasing proportion of synthetic data). Numerous studies (Bohacek & Farid, 2023; Martinez et al., 2023; Shumailov et al., 2024) show that the presence of a fixed proportion of synthetic data leads to model collapse, sometimes even when this proportion is very low. In these studies, the first generation of models is trained on real data while subsequent generations are trained on data comprising a proportion of synthetic data associated with real data, which remains constant at each generation. With this approach, the only way to avoid collapse could be to use new, real data on which models have never been trained (Alemohammad et al., 2023).

Another way to combine real data and synthetic data consists of the accumulation of synthetic data from each new generation of models alongside a fixed proportion of real data to train the new generation (the "accumulation" approach). In other words, synthetic data accumulates over time alongside real data to train each next generation. A study shows that collapse is delayed and accompanied by a reduction in data diversity (Alemohammad et al., 2023) (cf. technical appendix 2). Another team of researchers from Stanford University and MIT (Gerstgrasser et al., 2024; Kazdan et al., 2024) suggests that when synthetic data accumulates alongside real data instead of replacing it, catastrophic collapse is unlikely, at least after a few generations. The degradation in quality of what is produced would be much slower and would occur only in case of strong disproportion between (too little) real data and (too much) synthetic data, which would occur in the case of excessively weak creation of new data.

The proportion of real data necessary varies from one study to another, notably regarding the term of their effect. Shumailov et al., (cf. supra) emphasise that in the case of their study, it would be necessary to incorporate 10% of real data so that collapse occurs more slowly, which constitutes an important mass of data when we think that models are trained on trillions of data. Other studies suggest that 10% of synthetic data is sufficient to lead to collapse (Bohacek & Farid, 2023).

The risk of model collapse therefore remains important when the quantity of real data becomes insufficient. Potential collapse does not mean that large language models or other AI systems will cease to function, but rather that this will increase the costs of their development. As synthetic data multiplies online, scaling laws that suggest that models improve with more data may cease to be relevant, because this synthetic data lacks the richness of content generated from "real," human data (Wenger, 2024) (cf. Figure 2 for illustration).

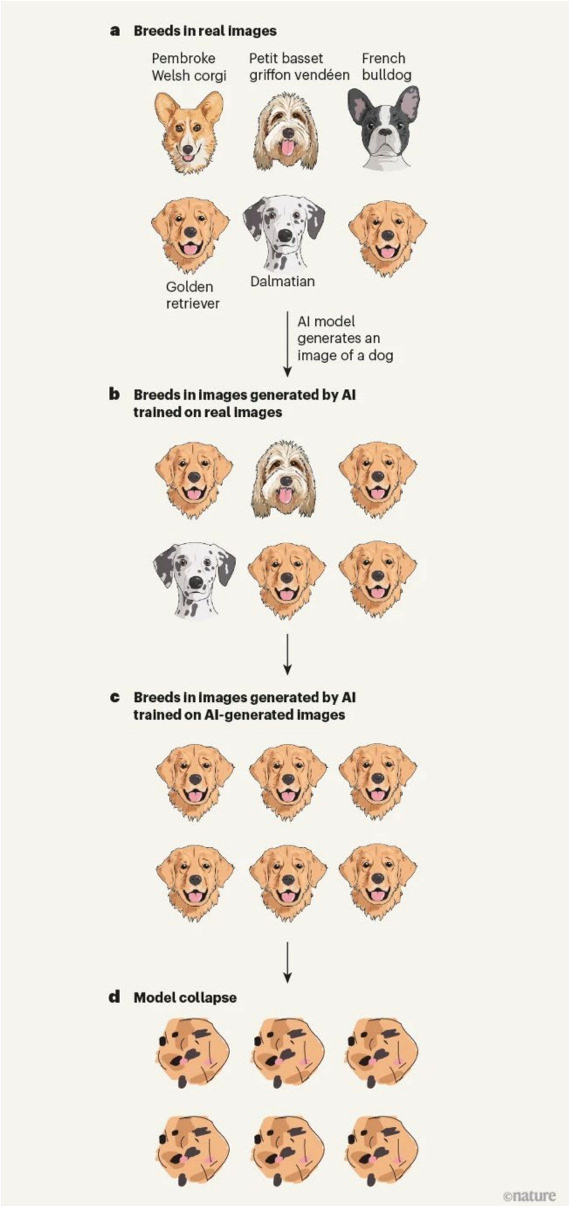


Figure 2. Illustration of model degeneration (Wenger, 2024)

The fact that training models on synthetic data leads to a reduction in the diversity of results (Alemohammad et al., 2023; Shumailov et al., 2024) is all the more worrying in cases where real data is itself biased and lacks diversity (Bolukbasi et al., 2016; Caliskan et al., 2017). Certain groups are more represented than others in data (Glickman & Sharot, 2024). Consequently, training on biased data leads models to generate biased content (Glickman & Sharot, 2024). A known example is the difficulty of AI face detection systems to recognise those of non-white people, because these models are trained on corpora that are not representative of the real world (Buolamwini & Gebru, 2018; Geirhos et al., 2022). This was similarly the case for an Apple algorithm that tended to assign a higher credit limit to men than to women applying for a credit card (Nasiripour & Natarajan, 2019); for a recruitment algorithm used by Amazon that valued male applications more than female ones (Dastin, 2022); and for Google's search engine which proposed more images of men than women in response to a neutral query like “person,” particularly in countries with strong gender inequalities (Vlasceanu & Amodio, 2022).

Similarly, Stable Diffusion, which is used by millions of users, when asked to provide images of high-income professions such as doctor or lawyer, principally generated photos of white men (Bianchi et al., 2023).

The presence of bias in real data, amplified by recourse to synthetic data, can lead to reinforcing users' own biases when they interact with AI models, distorting decision-making (Skjuve, 2023; Troyanskaya et al., 2020). For example, individuals assisted by a biased AI system to make a medical diagnosis reproduce the bias of models in their decisions, even once decisions are made without assistance (Vicente & Matute, 2023). The tendency of humans to be biased by AI systems that are themselves biased as a result of their use extends to other domains such as emotion recognition on faces (Glickman & Sharot, 2024).

New human and diversified data therefore remain indispensable to fight against model degeneration.

Among this human data, some concern more specifically those drawn from cultural sectors and protected by intellectual property. In these sectors, and by analogy to "mad cow disease," we can evoke the disease of "mad quasi-works." Mad cow disease refers to contamination via consumption of animal feed by cattle, which mushroomed due to the recycling of carcasses of sick cattle into animal feed given as food to other cattle. **AI, by replacing human cultural creation, could lead to creating only "mad quasi-works," synthetic ones which all end up resembling one another and which, by nature, are foreign to the disruption processes that mark the entire history of artistic activity**. Moreover, these human works must themselves be diversified if we wish to avoid model degeneration. Access to quality data within an adapted technical infrastructure, and reflecting the diversity of the real world—including the diversity of languages, cultures and regions of the globe—thus appears necessary and in the interest of all parties.

For this heritage to continue to be nourished, investments in human production and creation must be protected; we must take into account investments made by rightholders to produce original content. Beyond the loss of short-term revenue, the longer-term risk is that of an absence of investments allowing cultural industries to exist. Without funding, the incentive to create and produce new human works of quality and diversified could dry up.

In the field of photography, a study on contributors to Unsplash, a popular platform for royalty-free photos and illustrations, which has about 6 million high-quality pieces of content (Peukert et al., 2024) highlights diminishment in this specific case. In summer 2020, Unsplash launched an artificial intelligence research program by publishing a dataset comprising 25,000 images for commercial use. The objective was to analyse contributors' reactions, comparing those whose works were part of this dataset to those whose works were not included. The study results show that contributors whose works were used in this program left the platform at a higher rate than usual and considerably slowed their upload rate. This tendency is more marked among thriving professional photographers than among amateurs. Moreover, affected users decreased the variety and novelty of their contributions to the platform, which could have long-term implications on the stock of works available for the functioning of AI systems.

***

In this way, AI, by replacing human works with "quasi-works," carries the risk, in the short term, of destabilising a set of professional domains. In the longer term, this great replacement could lead to model degeneration if they are no longer (or little) fed by new human creations. The myopia of economic actors and a short-term vision of markets could lead to a failure to fully grasp the stakes at hand. The importance of culture for our societies goes without saying. And beyond any other motivation, investing in human creation among diverse sources is a necessity for AI models themselves.

2 – Value transfers between AI operators and IP rightholders: implementation framework

2.1 Contractual freedom and direct negotiations

2.1.1 Negotiations in the context of the TDM exception

In Europe, the 2019 DSM directive (Directive 2019/790 on Copyright in the Digital Single Market) creates a new exception to copyright for text and data mining (TDM). The response proposed by the directive to reconcile massive data mining and express authorisations from rightholders provided for by intellectual property law, is that of an exception to the monopoly, constructed in two stages.

In a first stage, Article 3 imposes an exception for the benefit of research organisations and cultural heritage institutions to mine, for research purposes, sets of protected works or objects. The exception is provided for without a compensatory remuneration mechanism. Rightholders assert non-compliance with conditions set by the directive, as demonstrated by cases of transfers, by public entities, of research results for commercial purposes.

In a second stage, to encourage uses of data mining, Article 4 provides for another exception, broader this time, for uses including commercial ones, under conditions of lawful access to data, and once more without a compensatory remuneration mechanism. This article results in an unprecedented and highly singular compromise. There is an opt-out possibility from this exception for rightholders, which supposes that they explicitly indicate, if necessary, their refusal of mining. In other words, the directive provides for an exception, but also the possibility of derogating from this exception and returning to the monopoly of exclusive right.

Questions quickly emerged about the practical feasibility of the opt-out mechanism (for a first legal analysis, see (Bensamoun & Farchy, 2020). **Furthermore, this article was inserted at the end of the legislative process, at a time when there were few indications of the lightning-fast development of generative AI to come**; the first application version of OpenAI's chatbot, ChatGPT, was released in April 2022. **The question of whether the TDM exception applies to generative AIs**, and the economic and societal implications that this article would have in such a context, **remains largely controversial**.

Numerous rightholders have chosen to opt-out, for differing motivations. Some do not wish to see their works used by AIs. For others, the faculty to opt-out should constitute a tool for encouraging negotiation. However, the conditions for applying this complex process are the subject of contradictory interpretations; moreover, negotiations leading to remuneration after exercise of the opt-out right provided by the directive often prove fruitless. First, because the TDM exception opens a potential negotiation space only if rightholders have expressed their intention to opt out by "machine-readable means"—that is to say, through automated means—which leads to real difficulties in both the technical application and the interpretation of this provision. What is more, if an operator removes the concerned works from its dataset, there is no longer a need for negotiation (see schema in Figure 3).

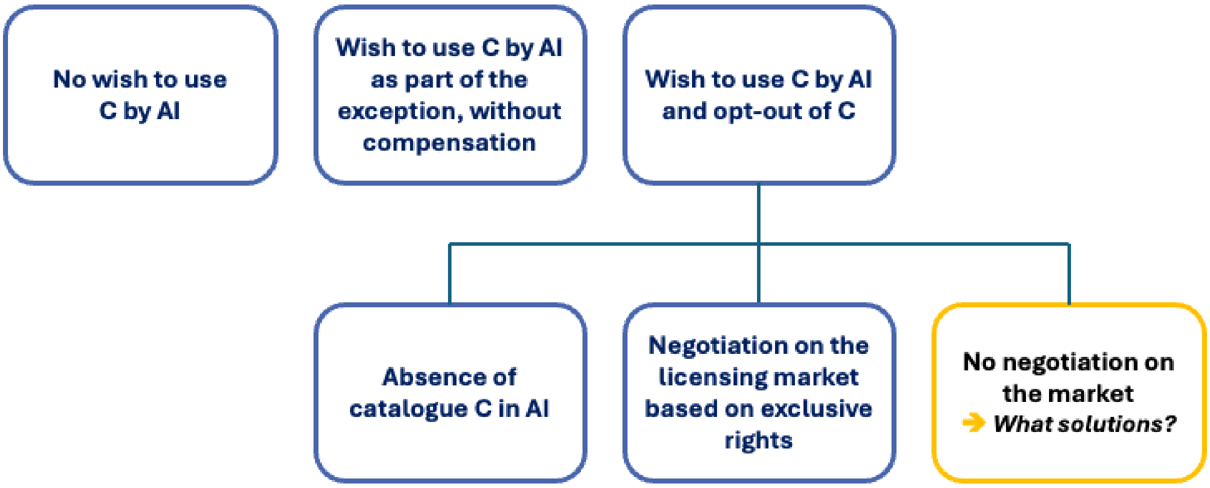


Figure 3. Remuneration of content catalogue C by an AI model within the framework of the TDM exception (in blue).

Currently, in France, when responses to requests are addressed to rightholders, the latter are very generally met with a refusal to negotiate. The arguments invoked are diverse: the infringing works are claimed not to have been used (without any proof for this being provided); the possibility is offered to withdraw works if a list of them is provided, though this withdrawal will not be retroactive on a model already trained; or the AI provider trained its model in a country outside the EU and is thus not bound by the directive. The implementation of the TDM exception is therefore, in practice, a thorny means of establishing licences and remuneration.

2.1.2 Negotiations in the context of independent contractual initiatives

In parallel, AI companies obtain commercial licences from press publishers in order to legally gain access to high-quality data (Cf. OpenAI agreements with NewsCorp, in the United States, Prisa in Spain, Axel Springer in Germany, Le Monde in France, or Vox Media, Associated Press, Financial Times, Google agreements with Reddit, etc.; or, in another area, Adobe and Getty Images). An agreement concluded between Mistral and AFP in January 2025 specifies that "AFP content will not serve to train and advance Mistral's models. This content is a 'module' that plugs into the system and can be unplugged upon contract expiration." Most agreements therefore do not concern pre-training phases but updating data and anchoring (cf. part 3).

These agreements, which take multiple forms (A. Thomas, 2025), have in common the fact that they are established via independent initiatives aimed at providing quality datasets and, beyond that, responding to the varied and specific expectations of each actor.

Producers and publishers, notably in the press or music sectors, put forward, for the future, the advantages of a free and competitive market for individual voluntary licences. These actors emphasise the freedom to grant large-scale licences, as they have demonstrated in the context of digital market development. Free negotiation between parties notably allows music streaming platforms to provide access to almost all global music libraries.

2.1.3 Limits of scattered contractualization

There are certain risks specific to agreements between tech companies and individual actors being concluded in a scattered fashion which must nonetheless be highlighted.

For rightholders, notably in certain sectors, actors are in an economic situation that does not allow them to contract or refuse remuneration conditions proposed during bilateral agreements. Only those whose data is most coveted, or those who will benefit from the "first mover" advantage, will be able to contract. Thus, once OpenAI has negotiated with a renowned press publisher in each European country, it is hard to see what would incite it to negotiate with other press actors who would provide it with news data judged to be similar, even if editorial slants are different.

With exercise of exclusive right conducted in a scattered fashion, only rightholders having technical control to do so and/or having high added value content can hope for remuneration. In this way, other actors would face increasing difficulty accessing this market. Moreover, remittances to authors, natural persons, should be the subject of particular attention.

On the other side, these negotiations could only benefit large **AI actors** who have the financial, human and administrative means to negotiate, once more closing the market to more modest AI providers and deployers. Indeed, the largest operators have competitive advantages due to their vertical integration (*Opinion 24-A-05 of June 28*, 2024).

▪ Upstream, for model production, they benefit from direct access to computing power and a large quantity of data associated with the use of their multiple services.

▪ Downstream, they have pre-existing distribution channels to distribute their models; on one hand, models are sold in complement to the sale of access to their computing infrastructure or cloud services; on the other hand, foundation models are sometimes integrated into existing products or services (search engines, social networks, office suites, smartphones) which already have a large user base; independent companies for distributing AI software can thus be penalised.

Dominant positions or forms of unfair competition could therefore appear due to the aforementioned vertical integration, or because of high fixed costs that push towards natural monopoly situations. The AI market remains however, at this stage, very competitive, as shown by the entry of new actors of more modest dimension such as DeepSeek (cf. part 3). The situation is likely to evolve very quickly, and therefore should be attentively monitored by competition authorities in Europe.

2.2 Mandatory value transfers

To avoid rightholders having to enter this market in a scattered fashion, mandatory transfer responses can be envisaged, either within or outside the intellectual property framework.

Within the copyright framework, the opportunity, legal feasibility, notably in an international framework, of alternative measures and any necessary legislative developments will be examined in the relevant part of this report. We limit ourselves here to economic aspects of the question. Currently, in some situations, the “complete” exclusive right including the monopoly on authorisation and remuneration (cf. part 4) is attenuated. In certain cases, the faculty to authorise or prohibit shifts from individual solutions towards collective ones (collective management, extended collective licence), though they nonetheless remain available. These cases thus form part of the economic framework of **market solutions for rights contractualization**.

In other cases, on the contrary, the faculty to authorise or prohibit disappears. In marginal cases, both prerogatives—authorisation and remuneration—are abandoned, giving rise to an uncompensated exception (for example, the exception for parody). In other cases, the faculty to authorise or prohibit is forgone, replaced by a simple right to remuneration (legal licence giving right to equitable remuneration or compensated exception for private copy). In these instances, **the faculty to authorise or prohibit disappears, as does the capacity for contractualization on a market, giving way to solutions organised under the guidance of public authorities—we speak of this using the term “mandatory transfers”**.

The sums generated by legal licence or private copy are far from negligible. Based on these known examples of right to remuneration, mandatory value transfer mechanisms (whose framework would result from discussions arbitrated by public policy and administration) are therefore proposed in the case of AI. France Digitale (Generative AI and copyright, 2024) thus suggests a compensated exception for content freely accessible on the web, with the exclusive right continuing to apply for “closed” content.

However, mandatory transfer modalities within the copyright framework present limits. Publishers and producers of press and music emphasise that, if recourse to this type of mandatory transfers does provide certain benefits, it nevertheless involves inconveniences relating to the uniform treatment of protected data, which as a general rule leads to a contraction of the value of the most sought-after content as well as the de facto remuneration obtained. Audiovisual producers note that compensatory remuneration—which only comes after broadcast—is not an incentive to produce new works, because it does not allow for such works to be funded, particularly in domains requiring significant financial means like the audiovisual sector.

Next, solutions organised on a purely territorial basis would not respond to the globalisation of the AI market. Moreover, the relation between mandatory measures and free market play remains an essential point of debate (should this operate as an alternative system or as an addition to existing practices, and should it depend on a given sector?). Lastly, public authorities must not contribute to organising a form of unfair competition to the detriment of European companies vis-à-vis their American or Chinese competitors, nor obstruct the development of markets under formation (cf. part 3).

Thus, establishing a mandatory blanket licence—debated then rejected just 20 years ago in France—might appear as a reassuring legal emergency solution to some when faced with the magnitude of digital piracy, but would undoubtedly have made difficult the subsequent formation of legal supply markets. As with blanket licence, in the case of AI, mandatory transfer mechanisms present the same drawbacks, the most significant being that of economic incentives; overall remuneration would no longer be linked to attractiveness and the potentially growing valuation of works with AI, but instead to long negotiations with uncertain outcomes, unlikely to favour rapid developments.

Nevertheless, because a licence market is struggling to form at this stage (see supra), the debate cannot be definitively closed. Prior to implementation, however, such mechanisms should be evaluated on the economic level according to the market failures they resolve compared to other systems (Lutes, 2025).

Beyond developments within the IP framework, other mechanisms are conceivable. Fiscal mechanisms first, like earmarked taxes on turnover inspired by existing models (taxes going straight to the support account of the *Centre national du cinema et de l'image animée*); financing obligations according to an economic logic that obliges those downstream (broadcasters) to finance that which is upstream (creation and production); or establishing a paying public domain or support fund for human creation: these paths remain to be explored.

2.3 The complementary path scenario: supporting the creation of a marketplace

Between the limits of scattered market-based solutions and a lack of reactivity to market evolutions in mandatory transfers, a complementary and optional path, that of collective support of the structural conditions facilitating market formation—that is to say, a structured exchange space allowing contractualization while respecting sectoral specificities—merits exploration.

The 2024 interministerial report on AI proposed establishing a technical infrastructure for holders of heritage data in the public domain (Aghion & Bouveret, 2024). Data from cultural activities used by artificial intelligence systems group together objects with diverse legal statuses. Some of these resources are held by heritage cultural institutions (libraries, museums, archives) and may or may not be protected under IP law. For several years now, the regulatory framework that defines the ways in which public information is made available and reused for public persons—the State, territorial communities, legal persons under public law—as well as legal persons under private law responsible for public service missions, has evolved greatly. Opening this vast heritage to digital use, which can correspond to public policy objectives around the influence of French culture, raises specific questions that are not treated by this task force; we concentrate on data protected by IP and held by private operators.

For this latter data, in opposition to a single market or mandatory management system, and respecting the specificities of different cultural sectors and different actors, participation in the marketplace could be done solely on a voluntary basis. This would not be about establishing a new legal obligation, and in no case could this participation substitute for respect of existing national or regional legislations, nor for the organisation of rights management, which is different according to sectors. Indeed, certain sectors regularly use collective management systems (voluntary or mandatory); others prefer individual negotiations to exercise their IP rights. Certain sectors like those centred around images do not have complete aggregated metadata bases. The goal is not to create ex nihilo a totally new system, but to rely on recognised expertise and competencies and the varied missions of different actors in place (collective management organisations, publishers, producers) to propose an integrated supply to AI operators.

The marketplace would therefore bring together in the same digital space not only the technical infrastructure for making files available, but also the legal authorisations for use as well as the economic conditions for remuneration. The objective of a common marketplace is to **group**, for defined catalogues or catalogue sections, **a triple activity of access/authorisation/remuneration**. As proof of this expectation, negotiations having already led to licensing contracts (cf. supra) are most often those in which operators possessing both rights and data-works (press publishers, scientific publishers, image databases).

Major institutions such as the National Library of France (*Bibliothèque nationale de France*, BNF) or the National Audiovisual Institute (*Institut national de l'audiovisuel*, INA) could play a role in this system. These institutions currently make available to third parties data-works that are still under protection, with the agreement of authors or rightholders (who can then negotiate the legal and financial conditions of use with the third-party as appropriate). This role could be extended to the situations discussed here. The BNF conserves, for example, a substantial corpus of data-works, particularly in the field of writing, which is likely to interest numerous actors developing artificial intelligence systems. Without being able to deliver authorisations or remunerate authors themselves, these institutions, like other actors, could play an aggregator role to deliver, at scale, digitised collections in which they have recognised technical expertise.

On the marketplace, catalogues would be made available within this common space, either by organisations already having established catalogues (collective management organisations, publishers, producers), or by new technical intermediaries, while AI providers would benefit from facilitated access to quality data-works. In all cases, the concerned actors retain their ability to negotiate and set prices. Prices would notably be adjusted according to the intended purpose of AI activities, based on annually negotiated licences (cf. part 4).

In no case can the existence of a marketplace prevent actors from contracting outside this marketplace if they wish, or from not contracting if they prefer to oppose uses by AI, by exercising the legal prerogatives at their disposal. The objective is solely to create incentives encouraging interested actors to take advantage of an economic opportunity.

The expected advantages are numerous.

Ensuring **economic conditions** for return on investment in human creation**.** As noted (cf. part 1), investment in quality human data-works, reflecting the diversity of the real world, including the diversity of languages, cultures and regions of the globe, appears necessary both for culture at large and for innovation in AI.

Avoiding mutual risks for the parties involved, notably of the legal uncertainty associated with potentially fraudulent uses of protected works, by centralising information on rights and persons with authorisation authority. In a globalised AI market, where companies are located in territories with diverse intellectual property legislation, by entrusting the interested parties with the negotiation of granted rights, contracts can, unlike purely national regulations, have global or regional scope. Note that the international scope of the AI Act for training AGI models commercialised in the EU and its potential legal conflicts will be the subject of subsequent work by the CSPLA. Moreover, **legal uncertainty carries major economic risks** for AI companies that could be subject to financially painful judicial decisions, particularly affecting European companies that have critical fundraising needs.

Technically, limiting transaction costs related to data search through simplified, mutualised, and as automated as possible access to organised cultural data, benefiting from a guarantee of reliability, quality and diversity. Some AI operators have highlighted the question of technical accessibility to files; actors who can offer legal authorisations are often, in fact, multiple holders for the same object, and different from those who hold the content files and associated metadata. Hence the place of technical intermediaries when rightholders are not capable of exercising this role.

Finally, in terms of **diversity,** allowing the most modest rightholders or those with specialised catalogues or their representatives, by benefiting from collective infrastructure, to make these catalogues accessible to interested AI companies. In parallel, allowing modest-sized AI actors who do not have the capacity to muster adequate services for the development of diversified offerings, thus opening the market to the innovation capacities of new entrants. Finally, by opening this marketplace to all actors, including extra-European ones who wish to join it, this space would encourage the promotion of minority cultures, in accordance with a French tradition that prides itself on supporting works and creators from around the world.

Rendering this solution operational requires, without doubt, that a certain number of sensitive subjects be debated and decided between the stakeholders, particularly concerning governance and financial rules; building the necessary infrastructure and making quality data technically available entails costs whose coverage must be discussed and distributed in order to associate, for example, private and public funds (AI operators, rightholders, France 2030, etc.). The relation of this marketplace with other initiatives remains to be foreseen.

***

Under current conditions, in France as in many other countries, direct negotiations between actors are rarely taking place. For that reason, this task force proposes a marketplace system that would play the role of an accelerator for these negotiations. Such market solutions do not exclude other, different cases, and more constraining legal mechanisms within or outside the framework of copyright law.

3 – Value chain and AI systems actors

To understand the value chain, we usually distinguish **AI models and AI systems**. AI models are integrated into AI systems of which they are an essential component. AI models require the addition of other components such as a **downstream user interface** or **upstream cloud services** to become AI systems. An AI system is therefore an automated system that is designed to function at different levels of autonomy and can demonstrate adaptability after its deployment. It deduces, from the inputs it receives, how to generate outputs such as predictions, content, recommendations or decisions that can influence physical or virtual environments.

3.1 Typology of systems and models

A general-purpose AI system has the capacity to respond to diverse purposes, both for direct use and for integration into other AI systems; it is based on a general-purpose AI model.

Within models, we distinguish foundation models, for general use, and specialised models. A foundation model is a large-scale model, pre-trained on enormous quantities of unlabelled data. It is designed to be adapted to a wide range of different tasks, notably after additional fine-tuning (see below). Furthermore, models specialised from the outset are developed for specific tasks, without going through the adjustment of foundation models. A specialised model for a domain is not always small-scale; thus BloombergGPT, a financial model from Bloomberg, comprises 50 billion parameters.

Finally, **generative** AI models (generative AI or gen AI) have the capacity to generate text, images and videos from textual instructions. Common applications involve users entering natural language instructions to generate results. Not all generative models are necessarily foundation models. Foundation models are broader in their design and potential application. They can be adapted to non-generative tasks such as classification or analysis.

Part of generative AI is based on large language models (LLM). The model generates the most probable response to a sequence of words produced by the user (a prompt, a query). Another part is based on **diffusion models**, typically used in **image** generation via a prompt. These neural networks are called “large” because of the number of their parameters: GPT-3, for example (used by OpenAI until recently) comprises 175 billion parameters. The quantity of data in a text being enormous, the LLM must understand a very large number of parameters. Concretely, LLMs are trained on large sets of textual data such as Common Crawl, The Pile, MassiveText, Wikipedia or GitHub. These datasets contain up to 10,000 billion words, which is very costly in computing resources and time. **Multimodal models**, for their part, seamlessly associate text, images, and speech. Companies such as Runway (which counts Google and Nvidia among its investors) or Synthesia thus monetise automated video creation solutions by training on text, images and videos, with clients using these videos for, to give an example, marketing purposes.

3.2 Segments of the value chain

In the AI Act (article 83), the "**provider**" is a natural or legal person, public authority, agency or any other body that **develops** or has developed an **AI system** or general-purpose AI model and places it on the market, or puts the AI system into service under its own name or trademark, for a fee or free of charge. The activity includes upstream resource management and development (European Commission, 2021).

The "**deployer**" is a natural or legal person, public authority, agency or other body **using** an **AI system** under its own authority, except when this system is used in the context of a personal activity of a non-professional nature. These are the users. We can cite among them hospitals and health establishments using AI for diagnosis, financial institutions using AI for risk assessment, recruitment companies using AI for CV screening, etc.

The value chain of an AI system can be more precisely segmented into three main blocks (Hoppner & Streatfeild, 2023).

- **Resources.** *Computation.* Upstream, AI developers require, in addition to qualified labour and data at scale, the necessary computing and storage power. To respond to this, a company can directly buy computing capacities, rent them (cloud services) or use existing technical infrastructures. Computations are mainly done on chips, in this case graphics processing units (GPU) whose market is concentrated around the American company Nvidia, which holds 85% of GPU market shares in the world in 2023, or GPUs developed by Google. Cloud services, allowing storage on servers (rather than on a single computer) of models and data, are dominated by American services AWS (Amazon), Azure (Microsoft), GCP (Google). *Data.* Services propose the creation, collection or preparation of data to train AI models.
- **Modelling** corresponds to the **development** and **training** of models, particularly large foundation models trained on gigantic quantities of data. These models are closed or accessible in open source. Certain companies specialise in the storage and sharing of open models.
- Downstream, models are **deployed and commercialised** towards services and end users, via applications developed by model developers or by a third party.

3.2.1 Resources – data

Data is accessed from different sources (Figure 4):

- Publicly available data**.** Data from web scraping (protected or not by IP, obtained in a licit manner or not) and open-source datasets continue to be important for model development. This type of data constitutes the majority of those used for pre-training foundation models, and is accessed once more when a new model supplants the previous one.

- Synthetic data. Several recently published foundation models have used synthetic data. Beyond cost reduction, this data responds to confidentiality/data protection concerns. However, there are limits to their use due to the risk of long-term model collapse (cf. Part 1).
- Third-party proprietary data. This data is collected by external entities such as data brokers.
- Directly proprietary data. This data, held by companies active in foundation model development, may not be accessible to their competitors (cf. Part 2).

Figure 4. Different datasets and their accessibility (AI Foundation Models, 2024)

3.2.2 Development: the stages of modelling

Schematically, the development of an AI model relies on two major successive stages. The use of data varies at each stage.

a) A training phase

Pre-training first refers to the training of a foundation model during which four different tasks are performed:

1) The collection of learning data called training data (in general, the more complex the model, the better its performance and the more it requires a large dataset);
2) The construction of a model (in the form of a neural network, which includes an input layer, intermediate layers and an output layer) potentially capable of linking **input** data (e.g. animal images) and desired **outputs** (the different animal species); initially, parameters take random initial values and the model provides a (random) response called output (e.g. a cat image is labelled as being that of a dog);
3) The definition of a cost function to minimise;
4) The training, i.e. the determination of parameter values that minimise the cost function so as to make model outputs and desired outputs converge (e.g. a cat image is correctly labelled as "cat", etc.).

Other stages can be implemented. **Fine-tuning consists of specialising the foundation model by "retraining" it on specific data or tasks**. For example, processing specific texts (a financial report, a legal text) or a specific task (sentiment analysis, recognition of "professional" terms, detection of defective parts). This can be done by modifying part of the model's parameters, but only marginally (so as not to erase the training), and potentially by adding one or more layers to the (pre-)trained model.

Most often, the model is trained on only part of the data, the **training data** (for example, 80% of the data) that will be pushed as far as possible by minimising the cost function, under the constraint of minimising the generalisation error when attempting to predict the rest of the data (for example, 10%), called **validation data**. However, validation data is not a perfect test of the model's ability to generalise, since they themselves participate in model training. The model could thus fail to accurately predict new data. Finally, the remaining 10% is often used to test the model on its ability to generalise data never seen before, what we call **test data**, and give an indication of model performance.

b) An inference phase, production deployment

After the training phase, inference corresponds to the production deployment operation of the model, i.e. the process by which a previously trained model will produce a result—predictions on new data. Inference can be completed by providing new data, fresh data, so that the model provides information taking into account current events or very specific data that the model will search for in an external source. The model is not trained on this data. This is what we call Retrieval Augmented Generation (RAG) or sometimes model grounding, which does not relate to training.

Grounding provides external information to the model during its use and does not alter its internal parameters. The most obvious external tool that can be called upon is a web browser, allowing the model to stay up to date; but practitioners refine language models so that they can use API calls (Application Programming Interface, corresponding to an interface between software, see below) and thus access a wide variety of tools. When a query is received, the system performs a search in a set of external documents or data. Relevant information is retrieved and used to enrich generation. The language model then generates a response by relying both on this external information and on its internal knowledge. This improves the precision and relevance of generated responses, allowing the use of up-to-date information without requiring model retraining, and reduces the risk of "hallucinations" by grounding responses in verifiable facts. Grounding is particularly useful for companies wishing to use language models on facts such as current events.

To summarise, three major categories of data are used in model development:

1) Training data, which is plentiful, to the order of several million or several billion;
2) Fine-tuning data, which is specialised, and may be available on the internet or carefully selected by a company or organisation;
3) "Fresh" data, consisting of grounding the model in current events without requiring training.

The influence of a dataset is not the same at each stage. Removing a dataset in the training phase only weakly influences model performance, since it is trained on an immense quantity of data. For fine-tuning, a relevant dataset for model use is crucial; if the dataset is not relevant, it has no value. The same goes for grounding. Moreover, it is not always the same actors who conduct these different tasks, relying either on data quantity or on their quality. Finally, for AI start-ups whose products are oriented towards a very specific demand, the main cost is not training but access to pre-trained models.

3.2.3 Model launch and deployment

Once models are trained, they are published, i.e. they are made available for deployment.

The model is then available in open source or in a proprietary format. In open source, it is potentially usable on new infrastructure and can be studied and modified. This is the case, to a certain extent, of LLaMA (Meta) or models from the French Mistral AI. In proprietary models, on the contrary, such as GPT (OpenAI), Gemini (Google), Claude 3 (Anthropic), access is controlled by licences, plug-ins, APIs, etc.

Once published, AI models can be fine-tuned, not by the initial developer (see above) but by a third party, user, company or intermediary (such as Eviden) acting on behalf of a client. These fine-tunings can be integrated into software and applications. Beyond fine-tuning tools available on development platforms for foundation models (FMs), such as those offered by Microsoft, Amazon and Google, certain companies, including OpenAI and Mosaic, offer fine-tuning services. These services can be useful for deployers and clients who do not have the internal technical resources to develop foundation models, but who would like to take advantage of the vast capacities of FMs while benefiting from a personalised solution (which can include a mix of models or access points, such as APIs).

There are several options for accessing and deploying foundation models. Companies can choose how they access models, from API access to open-source models to developing their own model. Depending on their needs and financial constraints, companies can include a mix of the most powerful models (used via APIs, for a fee or not) and less powerful models (for example, in-house), or a mix of proprietary and open-source models.

Models are made available to users via an application (in chat form, such as ChatGPT from OpenAI company or Le Chat from Mistral AI) or via a programming interface for developers (API), which allows a computer program direct interaction with the model. API access consists of authorising users to interact (i.e. by sending queries) with the model stored on a server by the provider. This is the case of OpenAI's GPT service. Access is often paid with usage restrictions, query limits, or differentiated rates depending on query volume or accessed functionalities. Access via Cloud platforms, such as Amazon Web Services, Google Cloud or Microsoft Azure corresponds to hosting AI models accessible to users.

Furthermore, there is an increase in the availability of paid API access directly from developers and via development platforms, such as Amazon Bedrock. For example, via Bedrock, clients can access foundation models, notably Claude 3 from Anthropic, Command from Cohere, Jurassic-2 from AI21, Llama 2 from Meta, Stable Diffusion from Stability AI, 8x7B from Mistral, and Titan from Amazon. Clients can also choose to pay for API access to certain of these models directly from developers, such as Anthropic, Mistral and Stability AI (AI Foundation Models, 2024).

To summarise, a company can choose to access models directly via the developer or through a marketplace that aggregates different models; it can choose to use a large general model, most often a foundation model or a smaller model, often specialised and developed in-house or via a specialised company. These models can be open-source or proprietary.

3.2.4 Users

Recent results from a survey concerning AI use by British companies, from "Business insights and impact on the UK economy" by the ONS, show that 15% of British companies currently use at least one of the AI technologies mentioned in the survey (which included FM services as well as other AI technologies); this figure reaches 46% among the largest companies (those with 250 employees or more). AI adoption tends to increase with company size, and the level of adoption varies considerably by sector (*AI Foundation Models*, 2024; Business Insights and Conditions Survey Team, 2024). The main reasons why companies use AI are: improving their business operations; providing or personalizing a product/service; developing a new product/service; or exploring a new market.

Beyond companies, many individual users use AI models. A 2023 survey measures that 20% of adults use AI to create new text or new images, and that 36% use these models to access information (Online Nation 2023 Report, 2023).

3.2.5 Complex valuation circuits in the deployment phase

Monetisation by foundation model providers of services to which these models give access is now almost systematic. For business clients, Amazon, Anthropic, Google, Microsoft, Mistral AI, OpenAI and Stability AI offer a variety of paid tiers (monthly subscriptions, payments based on the number of input and output tokens used, pay-as-you-go, credit-based systems, etc.). For individual clients, most services, such as those offered by Google, OpenAI, Anthropic, xAI and Microsoft, are either free or around $20 per month for a subscription.

To give orders of magnitude, in 2024, OpenAI's revenue came partly from API revenues (estimated at 27%, $1B), and partly from subscriptions (73%, $2.7B) (Khan, 2024; Sacra, 2024). Anthropic derives its revenues essentially from APIs (60-75% from third parties ($600-750M), 10-25% from direct usage ($100-250M) and about 15% from subscriptions ($150M)) (Khan, 2024; Sacra, 2024).

We also observe a trend towards the monetisation of productivity software with AI integration, presented as a premium complement to other systems; Microsoft's Copilot, for example, can be purchased as a complement for existing users of Windows Home, Pro or Enterprise.

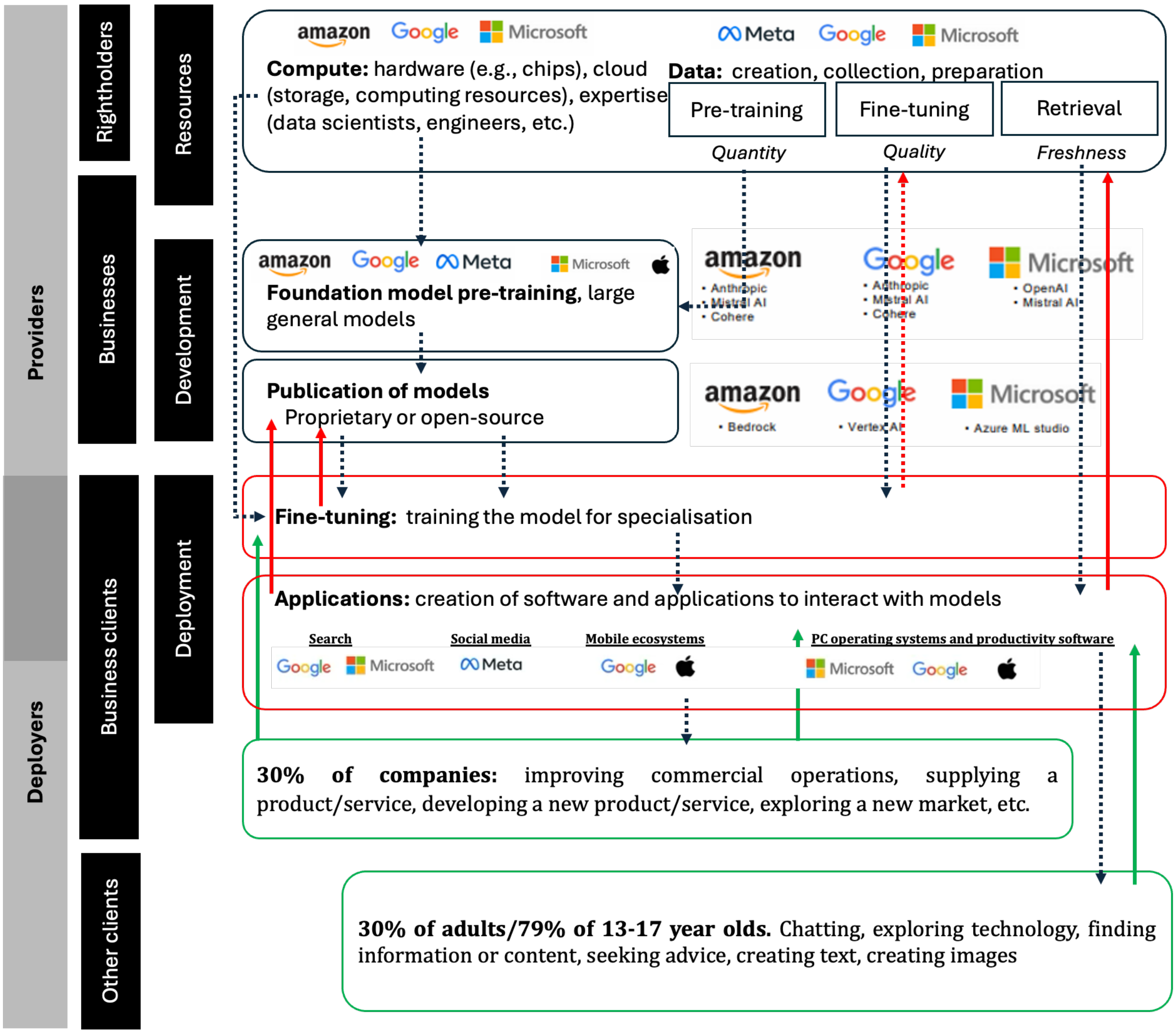


Figure 5. Overview of the foundation model value chain (adapted from AI Foundation Models, 2024; Treasury, 2024). Black arrows indicate the use of different components in foundation model development. Red arrows indicate payment by operators to use models, refine them, or fuel them on data. Green arrows indicate user payment.

Furthermore, **monetisation also occurs, for other operators, through their foundation model fine-tuning activities and the creation of user interfaces**. By paying foundation model developers, companies realize their own applications which, in turn, they bill to their clients. Fine-tuning is sometimes performed on "business" datasets furnished by the client themselves. Provided that certain foundation models remain accessible in open-source, applications from these models may not be overly onerous. Other economic models rely on expensive subscriptions accessible essentially to large companies (*AI Foundation Models*, 2024).

As an example, in music, most content generation applications, trained on a large corpus of existing titles, are intended for amateurs. Companies like Suno AI or Udio AI offer functionalities at prices that are accessible to the general public (from free for limited use to a few dozen dollars per month). Soundful, another music generation company focusing on professionals, charges significantly higher rates for companies.

3.3 Operators and market trends

3.3.1 The rise of partnerships

The British Competition and Markets Authority has identified several categories of value chain organisation between companies (*AI Foundation Models*, 2024):

- Vertical integration. The same entity is present at different levels of the FM value chain, whether in FM inputs (such as computation or hardware), FM development, and their deployment.
- Partnerships in the FM supply chain. Partnerships exist between FM developers and cloud service providers (CSP).
- Dispersion across the value chain. A range of companies operate at different levels of the value chain. For example, one company provides computing resources, another develops FMs, and a third deploys this model in its own products and services.

The authority highlights the continued development of partnerships, investments and strategic agreements for foundation models (*AI Foundation Models*, 2024). Partnerships can offer significant advantages for the parties involved and lead to increased innovation and efficiency gains (accessing rare resources, bringing their models to market more quickly and at larger scale, etc.). Since 2019, more than 90 partnerships between "GAMMAN" companies and "partners" have been identified. GAMMA companies (Google, Amazon, Microsoft, Meta, Apple) and Nvidia (which is the main supplier of AI accelerator chips) are called "GAMMAN". Foundation model developers, deployers of these models, or tool providers for developers of these models are called "partners". Furthermore, AI start-ups are often acquired by larger actors.

The British Competition and Markets Authority also observes a wide variety of partnership structures (*AI Foundation Models*, 2024):

- **Data partnerships**: Partnerships can allow one party to access the other party's data (Meta and Shutterstock, Google and Reddit, etc.);
- **Computing partnerships** allow foundation model developers to access computing resources, including access to specialised supercomputing systems or chips (Microsoft and OpenAI, Amazon and Anthropic, Google and Anthropic, etc.);
- **Distribution partnerships** which can take several forms.
    - **Foundation model distribution**: Certain companies build development platforms offering a library of foundation models. Partnerships can allow a GAMMAN company (1) to add the partner's model(s) to its library or (2) to provide access to the partner's model(s) via the GAMMAN company's foundation model development tools (Amazon and Cohere, Google and Mistral, Microsoft and Meta, Amazon and HuggingFace, etc.).
    - **Tool distribution**: GAMMAN companies can also add the partner's foundation model development tool to their own platform or marketplace (Microsoft and Nvidia, etc.).

- **Foundation model infrastructure distribution**. A GAMMAN company can distribute a partner's AI infrastructure via its own cloud marketplace (Nvidia and Google, Nvidia and AWS, etc.).
  - **Accelerator programme**. GAMMAN companies can create an accelerator programme for AI partner start-ups (the Meta/Hugging Face/Scaleway start-up program). These can offer funding, computing resources, as well as coaching and networking opportunities.
- **Investments**: GAMMAN companies can be one of multiple investors in a partner company, alongside other GAMMAN companies. Venture capital firms also commonly participate in these funding rounds (funding rounds for Runway AI, Cohere, Adept, Inflection and HuggingFace, etc.).

3.3.2 An oligopoly of American companies dominates the foundation model development market

According to ADLC, AI is the first technology to be dominated by industry giants from the outset (*AI Foundation Models*, 2024). Certain companies can play several roles in the value chain. For example, Google can be both a provider of its own AI systems and a user of third-party AI technologies in its products.

The development market is dominated by American foundation models. We count among the providers: Google (developer of Gemini); OpenAI (creator of ChatGPT and GPT); Meta (developer of LLaMA); Microsoft (OpenAI partner); DeepMind (Google subsidiary); Anthropic (developer of Claude); Stability AI (creator of Stable Diffusion), etc. (cf. Figure 6). Google, Microsoft and Meta represent a quarter of recent foundation models out of the 348 models listed (Treasury, 2024).

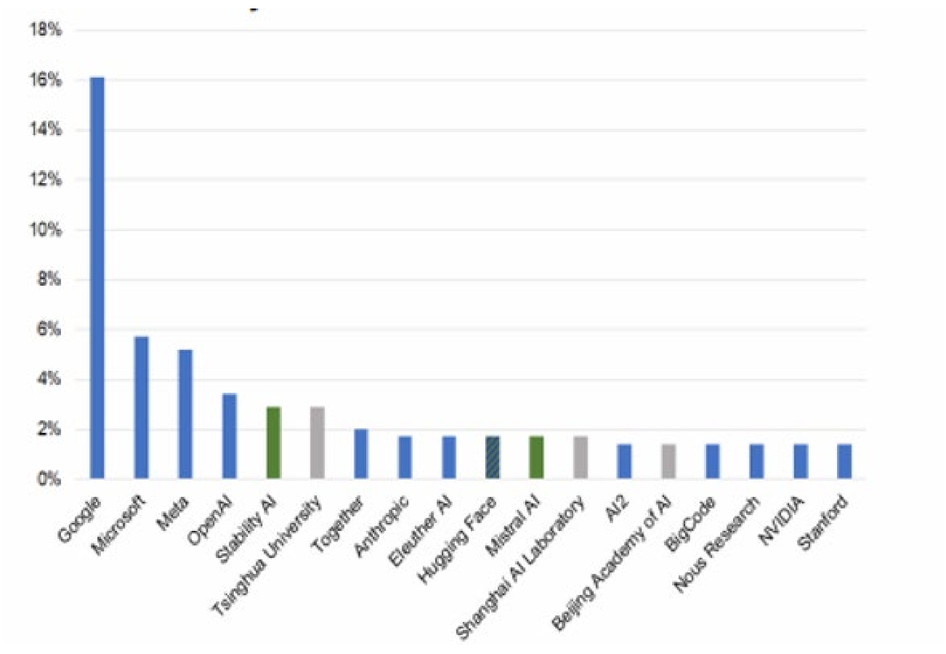


Figure 6. Share of different organisations in the total number of foundation models between January 1, 2022, and October 4, 2024. (Treasury, 2024)

3.3.3 Small is beautiful

Although current foundation models continue to increase in size, it is unlikely that many actors will come to join the restricted club of companies that provide them, given the development cost in both economic and environmental resources. All the more so because, faced with high costs, potential revenues are sometimes aleatory and the economic balance uncertain. This is why, **in terms of development, trends towards economy of means, specialisation and hybridisation** are emerging.

The company Perplexity AI thus offers a search engine based on "hybrid" AI models, i.e. on foundation models offered by other companies but also on its own model and on RAG. Moreover, there is growing interest in developing more compact models (with a reduced number of parameters) which, while offering extended capabilities, require fewer resources for their development or deployment. This trend is partly motivated by computing costs and by the fact that many use cases do not require the full capacity of large, general-purpose models. The newcomers Mistral and DeepSeek, although very different, have made this economy of means a marketing argument. The conversational agent DeepSeek, from a Chinese start-up in early 2025, is a generalist LLM, partially open-source, positioning itself to compete with OpenAI; the system distinguishes itself from others because it operates with reduced-capacity GPUs (Laird, 2025; Nellis & read, 2025). Its development cost, as announced by Chinese authorities, would moreover be much lower than that of American giants. The models of French flagship Mistral are smaller than those of the giants (the "Mixtral 8x7B" model counting 46.7 billion parameters but using only 12.9 billion per token (a token is a unit corresponding to a group of letters). The French company also publishes several small models (between 3 and 8 billion parameters). Among recent examples of small models, we also find Gemma 7B from Google and Zephyr 7B from Hugging Face, which are claimed to match, or even surpass, much larger models on certain criteria. Moreover, "distillation" allows reducing the size of a model while improving its performance for specific tasks; it consists of creating a small, efficient model that learns to imitate a larger model by trying to limit performance loss. The small model is trained on the predictions of the large model rather than on data themselves (*A Three-Step Design Pattern for Specializing LLMs*, n.d.; *AI Foundation Models*, 2024; *Distilling Step-by-Step*, n.d.)

Finally, specialisation is increasing, either through the development of models specialised from the outset, or through the retraining of foundation models via fine-tuning or through grounding models with fresh data. Microsoft's Phi-2 is a "small linguistic model" with only 2.7 billion parameters (Hughes, 2023). Developed with carefully selected training data, Microsoft claims superior performance to Llama 2 70B on specific criteria, such as coding or common-sense reasoning. Microsoft's Orca-Math is another recently published model, in this case dedicated to mathematics and created by fine-tuning the Mistral 7B model (AI Foundation Models, 2024; Hughes, 2024).

***

This brief overview of the value chain aims to better understand value creation in different markets, in order to subsequently raise the question of sharing this value out for the benefit of culture. The question of value creation goes well beyond the pre-training phase of foundation models of the most publicised companies. Mass extraction of data that is directly accessible on the web is a collection method mainly used by foundation models, most of which are American. Other data acquire their value from the specialised uses for which they are intended.

The different activities of pre-training, fine-tuning or grounding are sometimes carried out by the same companies, sometimes by different companies. In an ecosystem under formation, the contours of which remain to be confirmed, foundation model providers monetise the application or fine-tuning services to which these models give access to end users (companies and individuals). Furthermore, other intermediate operators monetise their own fine-tuning activities and diverse application systems (cf. Figure 5). Monetisation activities are likely to diversify and amplify in the coming years.

The value of cultural data in this global ecosystem must therefore be considered in light of this observation. Certain companies, to exercise their commercial activities, have an imperative need for quality cultural data for different actions, which are not limited to scraping, in order to bring specialisation and freshness to the proposed results. The adequacy of available data, with the multiple applications and uses deployed by AI operators, increases the value of this data and, thus, the expected remuneration. It is therefore not only all the actors but also all types of value-creating activities in the relevant markets that must provide the basis for sharing.

4 – Valuation of data-works for AI systems

Whatever the framework in which value transfers are organised (cf. part 2), and given the complexity of the value creation chain in AI systems (cf. part 3), the question of the value of data-works arises, as well as that of the sharing that flows from it.

4.1 Quantifying data value

Recent scientific research on quantifying data value for AI models helps illuminate the debate. An abundant though new literature devoted to the notion of data attribution consists of calculating the marginal contribution of each dataset to the model's performance in general, and to the genesis of a particular result (an output) following a user's query. A central point in this case is that of the substitutability of this dataset, which breaks down into two related questions. Can the model perform in the same way, i.e. correctly respond to the various queries formulated to it, without this dataset? And what is the relative contribution of a dataset to a specific output?

To answer these two questions, in the literature, three approaches co-exist. The first consists of changes in model parameters (whether training models on data subsets, or by altering the parameters of an already trained model) in order to establish causal links. The second, correlational, seeks to measure the similarity between the result generated by the model and the elements constituting the training database. The third, causal and proactive (which cannot be applied to pre-trained models) corresponds to watermarking ingested data.

4.1.1 Establishing causal links by modifying model parameters

The leave-one-out cross-validation method

One way to study the influence of a dataset is to train it without this dataset, and compare the content generated by AI between the model trained on the entire dataset and the one trained on the reduced dataset. This is the leave-one-out cross-validation method. Formally, this problem is often formulated as follows: how does removing a particular data point from the training dataset and subsequent model retraining affect its output? This change in output serves as a measure of the influence of the removed data point on this specific model output. This is done in two steps: training on the reduced set, and a measure quantifying the difference in quality between the two outputs (Maleki et al., 2014).

The main studies aimed at testing the feasibility of this approach first generate content with different ways of removing the training dataset. Thus, the real influence can be known, since we know which dataset was used to generate a type of content. Then, it uses the removal method to see if we recover the right dataset, i.e., the one that actually served to generate the output. By proceeding in this way, we can correlate test results to "reality" and see the quality of the method. These studies show significant results, but the degree of correlation between variables remains weak.

The leave-one-out method includes several limitations. First, it only examines the consequences of removing a data source from the complete dataset.

This approach might not accurately reflect the importance of a data point due to potential complex interactions between sources. Furthermore, duplicate data points, common in many machine learning datasets, may not credit the contribution of an item. For example, consider two dataset owners with almost identical characteristics. Removing either one from the dataset would probably result in minimal change

Furthermore, duplicated data points, common in many machine learning datasets, may not credit the contribution of an item. For example, consider two game dataset owners with nearly identical characteristics. Removing either one from the dataset would probably result in minimal change in the model's content generation probability, thus making the leave-one-out scores of each close to zero. This scenario could unfairly fail to attribute value to any of the contributors, despite the crucial role of their datasets in the model's performance. Moreover, in situations with numerous data sources, the leave-one-out score could diminish to nearly zero, failing to recognise the subtle contributions of individual sources.

Finally, the training phase on reduced datasets has a prohibitive computational cost (in terms of time, computing resources and energy) if one wants to understand the influence of each dataset, since it requires training a model on each data sub-type. Certain methods exist to reduce the cost of this operation (Hammoud & Lowd, 2024), which nonetheless remains high.

The "Shapley Value" Method

An alternative to leave-one-out is presented notably in a recent study (J. T. Wang, Deng, et al., 2024): measuring the incremental impact of incorporating a new data source in model training, considering all other possible combinations. **The approach is reversed compared to leave-one-out; instead of removing certain datasets from the model, they are added through successive iterations.** We thus measure the incremental impact of incorporating a new data source in model training, considering all other possible combinations. In other words, the idea is to first train the model on a small database. Then, the model is fine-tuned with the sequential addition, one by one, of new data. Each fine-tuning produces a new model, which contains a certain number of datasets.

We can proceed this way in different orders. For example, if we consider three datasets, A, B, and C, the model can be trained on dataset A, then fine-tuned on dataset B, then fine-tuned on dataset C. Finally, we can repeat the approach considering several orders (e.g. B, then C and A, etc.). This is more precise than training it on only A, only B, only C, A and B, B and C, etc. (cf. fictitious example in box 3 in the appendices). This approach allows one to derive a metric of the marginal contribution of a dataset to content creation in different orders, distributing compensation according to this marginal contribution (cf. box 4 Figure 1). The probability of generating a specific output after each adjustment step gives an estimation of the importance of each database.

This approach calculates the "Shapley value," a concept from cooperative game theory that leads to the fair distribution of rewards or costs among participants based on their individual contributions to a collective result. The Shapley value method takes into account the incremental impact of incorporating a data source alongside all possible combinations of other sources.

This method, announced as a more precise alternative than leave-one-out, nonetheless presents the same difficulties in terms of costs. To avoid the prohibitive cost of calculating the Shapley value on a large dataset, different proposals exist to approximate it. Here are some examples, among the abundant recent literature.

Some methods attempt to determine the influence of data on an output. A method called influence function, designed to approximate the Shapley value (Koh & Liang, 2017), consists of calculating the influence of a dataset on a model's parameters by overweighting a data point without retraining the model, rather than removing it then retraining the model, so as to evaluate the influence of the removed data point (leave-one-out). This method is either very imprecise, or very costly (if applied to each model subset) (Jia et al., 2019) (Koh & Liang, 2017). Another method consists of simplifying the model by reducing its number of parameters and simplifying the link between inputs and their output, in order to calculate the influence of each data point on the output without training the models (or only a few; methods called TRAK). Using this approach, (J. Deng et al., 2024) show on a relatively modest detailed musical dataset that this method—which is rapid and therefore inexpensive to implement at large scale—is correlated with the method of training models on data subsets (à la Shapley), though only modestly: 30%. This approximation is therefore imprecise.

Other approaches consist of using random or "intelligent" sampling methods. A team led by a researcher from the University of California at Berkeley compared the computational cost of different methods (Jia et al., 2019). The first consists of sampling models trained on data subsets (reduced data), with the goal of approximating the exact calculation of the Shapley value. This method is precise, faster than training the model on all data subsets, but remains unusable as it requires so much execution time. Another approach is inspired by group testing, which consists of determining the optimal test to determine if an object is defective. For example, if one wishes to determine which lightbulb among six is defective, one possibility is to test them individually, with the risk of having to test five times, whereas one could test the bulbs by group (first two groups of three, to identify the group containing the defective bulb). This allows efficient sampling. We can apply this principle by grouping data to identify not the defective quality but the utility of data for a model output, or its Shapley value (cf. box 4 figure 2). This approach is relatively precise but remains too costly to be deployable.

Other authors have proposed making the model unlearn the output generated by the model, and evaluating the images that are less well-represented in the new model. This approach allows one to quantify the contribution of training data without retraining (S.-Y. Wang et al., 2024). Regenerating data without these images leads to an image very different from that which was initially generated. This method has been used only with image generation models (not with text). Above all, nothing indicates at this stage that it could be used on large datasets.

Let us finally cite a last approach allowing approximation of the Shapley value by calculating the contribution of a data point during model training (J. T. Wang, Mittal, et al., 2024). While calculating the Shapley value requires retraining the model several times with different data subsets to calculate marginal contributions, the In-Run Data Shapley solves this problem by exploiting the iterative nature of training algorithms, which is thus done in stages. At each iteration of a model's training phase, a subset of training data points is used to update the model's parameters. The degree to which this update improves the model's prediction on validation data gives a measure of utility of these data points. This method is quite rapid and seems not to significantly affect computation time. However, it uses approximations and doesn't sample all subsets, which can affect attribution precision. Moreover, although promising, this approach lacks empirical validation at large scale, and is therefore not usable in its current state. Furthermore, it is not usable for each output generated by the user, which means it essentially allows measuring the contribution of a dataset to the model's general performance.

In conclusion, numerous methods are under development to quantify the contribution of different datasets to outputs generated by AI models on user requests, or on the model's overall performance. On the theoretical level, a method inspired by the Shapley value that gives precise results allowing fair quantification appears ideal; however, its practical implementation seems prohibitive in terms of costs (in time and computing resources) at large scale, in the current state of knowledge. They could be used to evaluate the overall performance of models on a large number of random or representative requests, with the risk that the marginal contribution of each data might be almost null. They could not be implemented to determine the contribution of each data to each model output, even if proceeding by sampling. Approximations exist but remain at the proof-of-concept stage. Such proof has, for example, been provided on a small database (80,000 images compared to hundreds of millions (DALL-E) or even billions (Midjourney) of images for the most used models (Bohacek & Farid, 2023)) and the solution is, at this stage, far from being deployable at larger scale (J. T. Wang, Deng, et al., 2024, Wang et al., 2024). This is why research has oriented itself not on a causality link via marginal utility but towards an approximation more deployable at scale.

4.1.2 Establishing similarity links between model output and training data – "passive" correlational method

A second approach—an imperfect alternative to Shapley—is in comparing characteristics of outputs with known training databases. This approach consists of extracting certain characteristics of generated content (such as, in the case of music, intensity, tonality, and duration) and data (e.g. training), and quantifying the similarity between the characteristics of certain training database data and those of generated content. In this case, there is no need to retrain an already existing model.

Researchers from the University of Illinois Urbana-Champaign used this metric to verify that training data with higher Shapley value were indeed more similar to sound generated by the AI model (J. Deng et al., 2024). Other researchers (S.-Y. Wang et al., 2023) first used a pre-trained model to fine-tune it on well-identified images (each fine-tuning is performed on a "source" image), based on which they generated synthetic images. In this way, synthetic images are influenced in construction (through fine-tuning) by the image on which the model is fine-tuned. Naturally, these synthetic images are not uniquely influenced by the example on which the model is fine-tuned, but this suffices to have an informative idea of the "source" image, despite image noise.

The idea is then to test different attribution methods, which, if they are effective, must be capable of attributing a higher score to the source image than to any other image in the training set. The authors then extracted image characteristics and measured similarity between training set images and the generated image, with similarity converted into probability that each image belongs to the training set. To quantify the capacity of pre-trained models (6 commonly used encoders: DINO, CLIP, ViT, MoCo, SSCD, ALADIN) to retrieve the training set, researchers evaluated the proportion of elements from the example dataset used for image generation in the top-10 retrieved images (Recall@10). The results obtained are largely above chance level and vary according to the method used. Critically, the same pre-trained models fine-tuned on attribution data are more influenced by source examples used for image generation (cf. box 4 figure 3). However, they attribute credit to many images that did not directly contribute to the genesis of the model output, which is problematic. This approach therefore doesn't seem applicable in its current state, even if rudimentary demonstration tools exist, essentially for illustration (e.g. *(GenAI Attribution Simulator - a Hugging Face Space by TheFrenchDemos*, 2025; Lorphelin, 2024).

This second approach, less costly than Shapley, seems to work on images and on a small dataset, as shown by work currently conducted by the Centre of Expertise for Digital Platform Regulation (PEReN). Its computational cost is not prohibitive but the result, in the current state of research, would be of (very) limited precision. Here again, the recent studies are at the proof-of-principle stage. It is not clear whether this approach could be deployed at large scale and on other types of data. An alternative would be to sample a limited number of requests by AI type (ChatGPT, LLaMA, etc.), assuming the sample is representative, then extrapolate to all requests. Finally, let us recall that the mathematical similarity approach—unrelated to that of counterfeiting in IP—assumes comparing a *known* input reference dataset and a *known* output dataset.

Another avenue consists of preventing difficulties in future models by marking training data with a computational watermark.

4.1.3 Training data marking - proactive causal method

For future models, a solution could be to associate a watermark with each training image, and to identify these markers in output images (Asnani et al., 2024). For example, a recent method developed notably by Adobe researchers, ProMark, performs causal attribution of synthetic images to predefined concepts present in training images.

Unlike previous work that establishes a correlation between synthetic images and training data, this method does not assume that similarity equates to a causal relationship (cf. box 4 figure 4). ProMark associates watermarks with training images and searches for these watermarks in generated images, which allows directly demonstrating causality rather than simply approximating or implying it. The principle is simple: if a specific watermark, unique to a training concept (i.e. an image of a foot, a laptop, etc.), can be detected in a generated image, this indicates that the generative model relies on this concept during the generation process.

Thus, ProMark is based on two steps: encrypting training data via watermarks, and training the generative model with watermarked images. To watermark training data, the dataset is first divided into N groups, where each group corresponds to a unique concept requiring attribution. These concepts can be semantic (for example, objects, scenes, patterns or stock image models) or abstract (such as stylistic elements or property information). Each training image in a group is encoded with a unique watermark, without significantly altering its perceptibility. Once training images are watermarked, they are used to train the generative model. During learning, the model learns to generate images from encrypted images. Ideally, generated images should contain traces of watermarks corresponding to the concepts from which they derive.

This method does allow one to retrieve the images used. Here again, however, this is at a proof-of-concept stage, and is therefore not directly applicable in its current state. However, it constitutes a solution for future data intended to serve for training.

Before examining how these methodologies could, in the future, be deployed to quantify value transfers with AI systems, we must return to the current valuation process for protected works.

4.2 Compensation for works under copyright protection: the process of valuation

Since copyright is currently the instrument through which work compensation is carried out, it is useful to recall the process for calculating this compensation, so as to better envision solutions in the case of AI. On the economic level, the economic component of copyright refers to two distinct prerogatives:

- The exclusive right to authorise or prohibit the exploitation of the work on a market;
- A form of compensation which is in principle proportional to the work's exploitation revenues, in order for the author to share in its success; in many cases, lump sums such as non-refundable advances or guaranteed minimums, often paid before any market revenue, complement proportional compensation.

Payment of compensation involves several calculation steps:

- knowing the compensation base (A);
- determining the share of this base devoted to the upstream part of the sector (creation - production) (P);
- distributing the amount received for creation among various works and rightholders (R).

Each step of this calculation faces various issues.

4.2.1 A compensation base linked to the exploiter's activity

In intellectual property matters, two main principles help define the compensation base:

- Compensation must be a function of the revenue realized by the work's exploiter.
- Compensation must be related to the exploitation of protected works.

In the "purest" cases of intellectual property, the two approaches overlap: the exploiter's revenue expresses, through a price paid by users, a clear relationship between uses and works. Revenues can be of very different nature (unit price paid by the public, subscriptions, advertising revenue, public resources, etc.).

Beyond revenue, the compensation base systematically relies on the existence of a strong link between activity and users' practices concerning the works in question. When revenue comes from advertising revenue or TV channel fees, amounts that include deductions on the compensation base are negotiated. Revenues that have no link or too distant a link with the use of works are not retained (sales of confectionery in movie theatres, for instance).

The existence of revenue attributable to certain companies or services therefore cannot be taken as an overall compensation base when the link with content exploitation is weak or very weak. Conversely, the fact that certain companies choose to monetise certain activities and not others—offered "free" to the user—is not relevant to the importance of content for users in the activities in question, and **the company cannot claim the absence of revenue to deduce a null compensation base.**

4.2.2 A share attributed upstream, adjusted according to professional habits and power relations between actors

The weight given to the upstream part of the sector (P) mainly results from professional habits and, occasionally, from past negotiations in the most established cultural industries. In publishing for example, authors receive on average 10% of bookstore exploitation revenues (Racine, 2020). Behind these averages hide a distribution with very different rates, notably according to segments of each sector and a given author's celebrity, negotiated in contracts between authors and publishers.

For more recent digital activities, this share equally reflects a negotiation and a power relationship established within the sector. For the Spotify digital streaming platform, the company's revenue excluding tax from subscription revenue and advertising revenue gives rise after deduction of the share for the company itself (about 30%) to an amount intended for creation and production (*Spotify Launches Revenue-Sharing Partner Program*, 2025); this amount is itself distributed among publishers, producers, authors and performing artists, according to signed contracts.

Professor Ernst Fehr's consulting firm examined the compensation of media (on the Swiss market) whose content is offered by Google search, which should correspond to advertising revenue lost by these media when Google search diverts users from their website (Johann et al., 2023). The firm attempts to establish the share of revenue that should be allocated to press publishers and, as such, proceeds in several steps:

1) the amount of revenue in the market (in this case, 1 billion CHF, drawn essentially from advertising revenue)
2) the market share relative to the press market (the number of requests on Google search relative to news, in this case about 55%, or 550 million CHF)
3) the share of people who would not have used the service without rightholders' content (the number of people who would not have used Google search if press article summaries had not been provided, or 70% of 550 million according to an experiment, thus corresponding to 345 million CHF)
4) revenue sharing between the search engine and media creators, (which corresponds to 40% of 345 million CHF, or 154 million CHF).

Finally, the amount (154 million CHF) which, according to this method, should be devoted to press publishers corresponds to about 16% of revenue.

The calculation methods and percentage obtained here are, however, far from being indisputable. Nor are they generalisable to all markets and all services. They nonetheless pose major questions. The example of press publishers' related rights, established by the European legislator then transposed into French law, is enlightening. The affirmation of a principle—without further examination of the delicate question of economic information necessary for quantifying the compensation base and the share of this base divided between publishers and journalists—shows how much this provision has led to many difficulties in implementation.

Moreover, in French law, proportional compensation does not apply if the nature or conditions of exploitation make impossible the application of this type of compensation, notably when the calculation base for proportional participation cannot be practically known or determined. In this case, one should turn towards lump-sum compensation (cf. supra) while determining, beforehand, the basic elements on which to rely when setting this lump sum. Even in the absence of proportional compensation based on the success of a given work, compensation must always be proportionate to the user's activity.

Power relations between actors therefore play an important role in calculating value sharing (the base from which to share, then percentages or lump sums according to specific cases). Consequently, the question that can legitimately be asked is one of competition law: to what extent, considering possible dominant positions, is value sharing negotiated under non-discriminatory conditions? Thus, ensuring that value sharing negotiations take place under fair competition conditions is an essential issue.

4.2.3 Distribution among works and rightholders

The lump-sum base granted upstream that we find in numerous situations, does not prevent, in occasional subsequent instances, an attempt to bring the final share closer to estimated success. In the simplest cases, distribution among works and among rightholders of these works can be carried out according to real, known data. In other cases, surveys are used; a sample of 120 discotheques in France equipped with a device allows the SPRE to know when a given piece of music is played and, thereafter, to establish associated equitable compensation for rightholders (*La SPRE collecte la rémunération équitable pour les artistes-interprètes et les producteurs de phonogrammes*, n.d.; Lorphelin, 2024).

In Spotify's case, in 2023, the firm generated 12.5 billion dollars in revenue and returned 9.5 billion to producers, publishers and composer authors (Loud and Clear by Spotify, 2023). The amount returned to the music sector by the streaming platform is thereafter distributed among different actors proportionally to users' listening frequency (Spotify Launches Revenue-Sharing Partner Program, 2025). But the calculation processes of associated distribution gave rise to intense debates between the "market-centric" approach chosen by Spotify, the "user-centric" view taken by Deezer or, more recently, the contractual "artist-centric" approach.

4.3 Quantifying value transfers in the case of AI

4.3.1 Guidelines

Quantifying value transfers thus requires proceeding in three complementary steps that we detail below.

1 - The transfer base (A)

Whether the distribution is lump-sum or proportional in nature, the preliminary calculation of the base relies on precise identification of companies and activities within these companies linked to the use of protected works, and on evaluation of the resulting valuation. Data-works are integrated into an ecosystem that is not limited to the training phase; it is in the deployment phases—notably allowing a user, on request, to produce a result—that the various valuation sources reside (cf. supra, figure 5).

This leads us to suggest that the compensation base could be both refocused and enlarged: refocused on value-creating deployment activities more than on development activities; enlarged to companies that are not only the few titans at the origin of foundation models, but also those that create other activities during deployment.

Since value chains are still under formation and the cascading links between services and applications are numerous (cf. part 3), the work of identification demands further study. The objective is both to identify the services concerned and their applications while avoiding duplication (counting twice). This task force therefore proposes that, with the help of the Ministry of Economics and Finance, a precise mapping of the sites of value creation and the relevant markets on which to base value sharing should be put on the agenda.

2 - The share attributed to cultural creation (P)

Regarding the calculation of the share attributed to the upstream part of the sector (P)—be it proportional or lump-sum in nature—no clear, unique rule based on a defined economic calculation seems to have emerged in the history of cultural industries, aside from negotiations, often complex, between the interested actors (see supra). In the case of AI systems, several points merit emphasis at this stage. First, **the question of AI/culture sharing must be negotiated for datasets (catalogues of works) and not work by work**. On the other hand, the question of the economic evaluation of compensation differs from the question of the legal principle of compensation.

Transparency on access to upstream inputs: insufficient foundation for evaluating compensation

Making transparency obligations—provided for by the AI act and the exercise of the opt-out prerogative as well as by the Directive on Copyright in the Digital Single Market—the bases for compensation evaluation appears largely insufficient on an economic level. Legal work conducted at the European level, under the aegis of the Commission and the AI office, shows that transparency requirements might not reach the level of granularity expected by certain rightholders, nor concern all AI actors. Above all, implementing transparency that is at once precise and respective of trade secrets, would certainly lead to identifying the presence of sites or catalogues of protected works within inputs; it would be one of the means of support for the indispensable exercise of rights to an effective remedy and proof on the legal level, and would open the way to legal actions and to better knowledge of partners with whom to negotiate (on alternative mechanisms of the right to proof, cf. the legal part). But once transparency is acquired, nothing would be clear about the economic value of "borrowing" a dataset for the model and appropriate compensation. In other words, to use a culinary metaphor, transparency shows the list of ingredients present in the kitchen cupboards, but doesn't say in what proportions each was used in a given dish—to say nothing of the recipe— and, consequently, how much the cook should pay to obtain these ingredients.

Downstream output destination: a complementary indicative foundation for evaluating compensation

Moving from the presence of a work-dataset to its economic valuation—at the "P" level where we position ourselves here—we propose focusing on the intended purpose of an AI system downstream in order to gauge the value of data upstream. This involves identifying within the value chain of AI systems and their applications those whose activity has a direct link with the use of works, and those who have access to works but whose AI system has an intended purpose that is not directly relevant. The compensation level P corresponding to the use of work-data at various stages of AI system development and deployment would be based on an economic presumption[3] of use according to the model's intended purpose.

[3] It should be noted that taking into account an economic presumption linked to an organisation's activity, for the sole purpose of evaluating the level of remuneration that should be granted to culture, is both different from and

The notion of "intended purpose" is explicitly introduced in the AI act; it is the use for which an AI system is intended by the provider, including the specific context and conditions of use, as specified in the information communicated by the provider in the user manual, advertising or sales indications and statements, as well as in technical documentation.

The intended purpose we refer to is not that of input data but the visible one of results produced by models, systems or applications. Indeed, in AI matters**, the idea of pricing by major data categories is sometimes mentioned, even according to development phases (pre-training, fine tuning, RAG) since the value of certain data is not the same depending on cases** (cf. part 3). **These mechanical solutions, though attractive a priori, seem to us undesirable for operational reasons**; they would undoubtedly lead to spillover effects and opportunistic behaviour. In the absence of transparency on the intended purpose of input data—the most common case, except when the concerned data is subject to systematic and generalised "marking" (cf. supra 4.1.3)—it would not be possible to avoid, for example, an actor having low-price access to "pre-training" data before selling, at a much higher price, the use of these same data for other uses like fine tuning.

Thus, valuation according to the intended purpose of the output does not correspond to the idea, sometimes advanced, of unique prices set for major categories of input data. This task force recommends that the rates practised be proposed by rightholders themselves to AI actors according to categories of uses and users. The intended purpose principle is, moreover, already habitually practised in intellectual property matters to determine compensation levels (Senftleben, 2024). Therefore, in the exercise of collective musical management, SACEM adjusts the rates for using its repertoire according to whether the exploiter's business model relies essentially or incidentally on the use of protected works: a hairdresser who plays background music for their clientele pays a much lower percentage of their revenue than, for instance, a discotheque would.

For feeding AI systems, it is appropriate to distinguish, as we have already noted, the fact of triggering compensation (cf. legal part) and the relative level of this compensation, which we focus on here. This task force's role is not to enter into precise calculation of each party's rates in all sectors and for all use cases. **Guidelines corresponding to three major categories of compensation levels can, however, be identified schematically, according to the criterion of intended purpose**. Based on these categories, a continuum of pricing levels could be established by cultural actors themselves according to annual licences, renegotiable each year with AI operators.

Intended purpose #1: basic compensation levels

The intended purpose suggests deploying a vast quantity of undifferentiated data in which mere access to protected data occupies an incidental place for model performance, because other data would be substitutable. The model does not learn on particular data or for specific intended purposes. Example: a model deployed literary texts during its training; one of its applications consists of responding to requests from insurance company clients.

Intended purpose #2: intermediate compensation levels

potentially complementary to the debate on establishing a possible legal presumption mechanism, which would require legislative changes (cf. legal section).

The intended purpose suggests deploying non-substitutable protected data upstream, but the model produces outputs that are not quasi-works likely to replace human works. Examples: an application provides tactile representation tools for musical pieces for the hearing impaired, or tools identifying on-the-fly works that are being listened to, or tools allowing phonogram producers to gain productivity.

The specific value of the result obtained thanks to work-data inputs leads to an "intermediate" level of compensation, including when this result does not directly harm a specific rightholder or does not appear "similar" to an input. Specialised and non-substitutable work-data have indeed been deployed to achieve the model's overall performance.

Intended purpose #3: higher compensation levels

The intended purpose suggests deploying non-substitutable protected data upstream *and* the model produces synthetic quasi-works likely to compete with human works. Examples: a model automatically generates synthetic illustration images, likely to replace works, or the style ("in the style of") of human illustrators.

3 - Distribution (R)

Regarding distribution, one method is "pay to train" currently used notably in the audiovisual or image domain. Major image banks (Shutterstock, Getty Images, etc.), after negotiating with AI model providers like Open AI or Google, distribute the amounts received among actors according to the "pay to train" method which links compensation to the number of works each holds in the dataset. An author having 200 photos in the dataset will thus receive double compensation from one with 100 photos. Shutterstock offers on average and every 6 months 0.0078 USD per image to content creators, with an average of 46 USD per portfolio (estimation made on a small sample of 58 people). With a rich database of 615 million images, compensating on average 0.0078 USD per unit, the amount returned would be approximately 4.797 million USD, or about 2.2% of revenue (215.3 million USD in winter 2023) and 15% of Shutterstock's profit (32 million USD) (Growcoot, 2023). This distribution method, as it is approximative, takes into account here neither image quality nor their respective marginal contribution for a given user's request. In the audiovisual domain, other individual agreements made with content aggregators such as the American company Calliope Networks value content based on their quality, nature & duration (more than $6 per minute for exclusive content, which increases if 4K or 3D quality, vs $1 for short formats), to then return collected sums to rightholders with whom the company has negotiated agreements (the company went from a catalogue of 17,000 hours of audiovisual content in August 2024 to 35,000 hours in early 2025 cf. https://calliopenetworks.ai/).

These distribution methods have the advantage of simplicity. Distribution modalities employing new quantification methods could afford refinement, by taking into account the marginal contribution of certain data to produced results.

4.3.2 Operational contribution of quantification methods

The different quantification methods presented in the previous part help approximate the contribution of a work-dataset and its valuation. Implementing these methods, whatever their differences, has limits that do not allow deploying them on all occasions. The economic question posed therefore becomes that of arbitrage, in terms of transaction costs, between, on the one side, the cost of quantification methods and, on the other, expected benefits in terms of compensation.

When the quantity of data is such that the expected marginal gain for a defined catalogue of works would be minimal, and potentially lower than the calculation cost, causality and similarity methods are inoperative. This will be the case if it involves evaluating the contribution of a dataset to the results of a generalist foundation model. For these models, each dataset rarely has decisive influence. This does not mean that the value of input data is null, but rather that current quantification techniques do not allow determining this value without ballooning costs. Concretely, if the cost of implementing an attribution method for each use of the model—or almost each use, if samples are employed—equals or exceeds profits, no profit is to be shared.

In cases for which well-identified catalogues of works have been made available to AI providers, **the objective is thereafter to evaluate the relative contribution of these protected works to specialised models. In this case, a limited number of data are deployed by the model, making solutions economically possible** in the sense that marginal revenue would likely be substantial.

Researchers have proposed using **causal attribution methods**, which consists of modifying the model by training it on truncated datasets to quantify their contribution either to producing a given output (for instance, "generate an image in the style of X") or to the model's overall performance (for example, is the model capable of generating outputs corresponding to user requests?). Ideally, a set of sub-models is retrained to calculate the Shapley value and determine, from marginal contributions of datasets, the final distribution of compensation expected by each rightholder (cf. Box 3). In the scenario in question, input data, which would be clearly known, are limited in number (so that the Shapley method can be deployed). Furthermore, in the presented example (using Shapley) revenues realized on markets must also be known in order to have a distribution base. Approximations of the Shapley method can also be deployed (using sampling of the sub-model space).

On larger datasets, **approximation methods less expensive than Shapley, by similarity** (between the training database and output) are possible, notably on images, but they are technically of lesser precision. Indeed, the contribution of a key image can be diluted (false negatives) or, conversely, the contribution of images that were little used or not used at all can be overestimated (false positives). In the current state, this technique is therefore not totally operational.

Given the limited research on quantification, this task force proposes that additional studies be conducted rapidly on input data to be determined and based on request sampling, in collaboration with work currently conducted by the Centre of Expertise for Digital Platform Regulation (PEReN).[4] These experiments will allow moving from proof of concept to operational methods on use cases. Let us note for now that content generation models for images, audio, texts or videos cannot be treated in exactly the same way.

[4] PEReN is an interministerial service in the field of data science, algorithms and artificial intelligence, which acts as a shared technical expertise base for State services and independent authorities responsible for regulating digital platforms.

***

The reminder of the typical stages for the valuation of protected works has highlighted the importance of thinking about value transfers in the case of AI, using different methodologies according to what one seeks to determine: the base of these transfers, i.e. sites of value creation by AI operators; the part apportioned to culture during sharing between AI operators and cultural actors; distribution among works and rightholders within cultural sectors. For the evaluation of such sharing, beyond the general principle of intended purpose and the presumption on which it is possible to rely, quantification techniques are operative in limited cases. These techniques will undoubtedly be most operational on one hand, at the distribution level, in order to ensure that all upstream creators benefit from value sharing; and on the other hand, to prove the use of works. The following two tables summarise these analyses.

Value transfer evaluation methodologies
1 – Level A - base *Objective*: to identify value creation by AI providers and deployers. *Method*: collaborating on mapping with services at the Ministry of Economics and Finances. **2 – Level P – share given to culture** *Objective*: to contribute to determining different levels of value sharing. *Methods:* • Mobilise the visible intended purpose criterion of activities that feed on protected data to effect a gradation of sharing. Deepen quantification techniques to prove and evaluate the relative contribution of a work-dataset to a specialised model's performance and/or to responses to specific requests. **3- Level R – distribution** *Objective:* to distribute P amounts among different works and rightholders. *Method:* deepen quantification techniques on use cases.

<table>
<tr><th colspan="4">Complementary contributions of model destination principle and quantification techniques to evaluate sharing (AI/culture) and distribution (within the cultural sector)</th></tr>
<tr><th rowspan="2">Intended purpose</th><th colspan="3">Method</th></tr>
<tr><th>CAUSALITY</th><th>SIMILARITY</th><th>LABELLING</th></tr>
<tr><td>Intended purpose 1

Generalist models

Sharing: basic compensation levels</td><td colspan="2">Inoperative</td><td rowspan="3">Envisaged objective: contribution of a dataset to a specific result

Operative only for future models and in the case of a lack of risk concerning circumvention

Potential operationality: proof of use for sharing and distribution</td></tr>
<tr><td>Intended purpose 2

Specialised culture-media models without output competition

Sharing: intermediate compensation levels</td><td rowspan="2">Envisioned objective: contribution of a dataset to the general performance of the model/to a specific result with a significant computational cost.

Operative on datasets limited in number

Potential operationality: sharing and distribution</td><td rowspan="2">Envisioned objective: Contribution of a dataset to the general performance of the model/a specific result with limited precision.

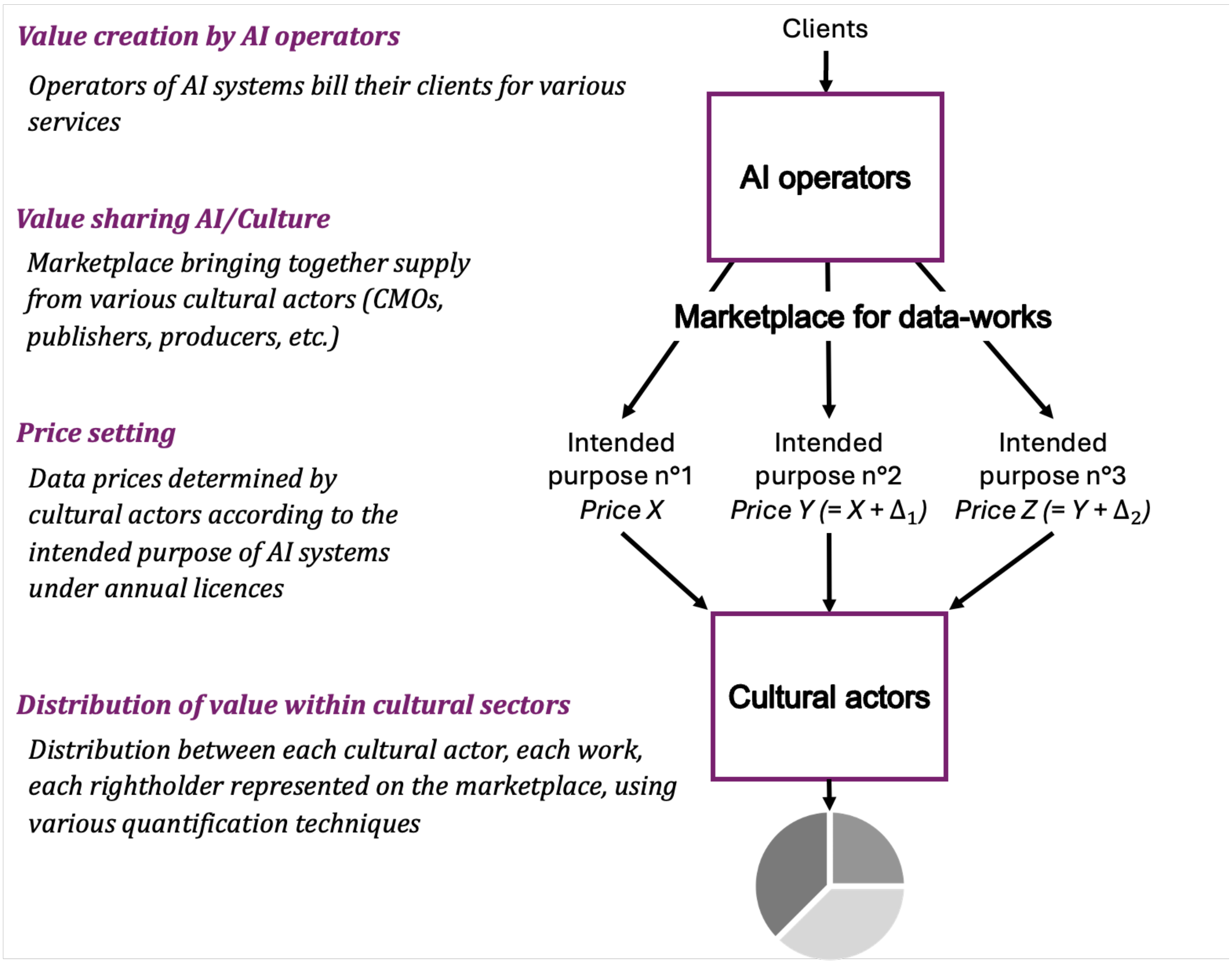
Operative on datasets limited in number

Potential operationality: sharing and distribution</td></tr>
<tr><td>Intended purpose 3

Specialised culture-media models with output competition

Sharing: high compensation levels</td></tr>
</table>

The following graph illustrates, as an example, the way in which value transfers could be implemented, in a scenario which combines a marketplace system (cf. part 2), the appraisal of value creation by AI operators (cf. part 3) and the use of the intended purpose criterion to share out value between cultural actors and AI providers (cf. part 4).



Recommendations from the economics report

1 – Explain and give publicity to the joint interest of cultural actors and AI operators in investing in a sustainable ecosystem that guarantees both the presence of European works in AI systems and the sustainability of their funding.

2 – Establish and/or consolidate appropriate support and training policies for the professions most directly impacted by the rise of AI.

3 – In the context of consultation between cultural actors and AI operators, consider the opportunity and feasibility of building a marketplace, a structured exchange space, enabling contractualization while respecting sectoral specificities.

4 – In the context of consultation between cultural actors and AI operators, explore the opportunity and feasibility of compensation mechanisms and value transfers in addition to those provided by intellectual property law.

5 – Carry out, with the services of the Ministry of Economics and Finance, a precise mapping of the sites of value creation and relevant markets, and monitor the circuits of valuation in the deployment phase, to provide the foundation for value sharing.

6 – Refine the operationality of the criterion of economic presumption of use according to the intended purpose of results produced by models, systems or applications that use protected data, to establish a value sharing scale.

7 – In collaboration with PEReN, further examine through case studies the operationality of scientific quantification methods to prove and/or evaluate the contribution of certain data-works to the results produced and/or to the overall performance of specialised models. Promote among cultural and AI operators the solutions deemed most relevant, according to a given case.

Organisations and individuals consulted for the report (legal and economic aspects)

Aday
Administration of Rights of Performing Artists and Musicians (ADAMI)
Adobe
AI Collaborative (Martin Tisné)
AI disclosure project (Ila, Stauss, Tim O'Reilly)
AIE (Italian Publishers Association)
General Information Press Alliance
French Alliance of Digital Industries (AFNUM)
Alt – Edic
Amazon
APIG
Association of Developers and Users of Free Software for Administrations and Local Authorities (ADULLACT)
Association of Literary Translators of France (ATLF)
Association Les voix
Competition Authority (ADLC)
Axel Springer
Bauer media
Bergaud Antonin (HEC)
National Library of France (BNF)
Bourreau Marc (Telecom Paris)
Brison Fabienne (HOYng Rokh Monégier Law Firm)
Cafeyn
Cairn (Thomas Parisot)
French Center for Copyright Exploitation (CFC)
National Center for Cinema and Animation (CNC)
Combé Julien (Ecole Polytechnique)
Condé Nast
Controv3rsé
Ekhoscènes
Emma ENPA
En chair et en os, collective
Eurocinema
Eviden
International Federation of the Phonographic Industry (IFPI)
Federation of European Publishers (FEP)
France Digitale
France télévisions
French Flair Entertainment
GEMA
Gesté
Ginsburg Jane (Columbia University)
Google
French Information Industries Group (GF2I)
Imatag
INRIA
National Audiovisual Institute (INA)
L'Express
Lagardère
Les Echos – Le parisien
League of Professional Authors

LinkUp
Ministry of Economy, Finance and Industry, State Secretariat for AI and Digital, Cabinet
Ministry of Economy, Finance and Industry, General Directorate of Treasury (DGT)
Ministry of Economy, Finance and Industry, General Directorate of Enterprises (DGE)
Ministry of Culture, Cabinet
Ministry of Culture, Service of Legal and International Affairs
Miso.ai (Lucky Gunasekara)
Mistral
Netflix
New Republic of Centre West (NRCO)
Panneau Fabienne (DLA Piper Law Firm)
Perchet Vianney
Digital Regulation Expertise Center (PEREN)
Prisma Media
Radio France
Rolling Stone magazine
Roux Steinkuhler
Civil Society of Multimedia Authors (SCAM)
Society for Perception and Distribution of Performing Artists' Rights (SPEDIDAM)
Society of Authors in Graphic and Plastic Arts (ADAGP)
Society of Authors of Visual Arts and Fixed Image (SAIF)
Society of Dramatic Authors and Composers (SACD)
Society of Authors, Composers and Music Publishers (SACEM)
Society of People of Letters (SGDL)
Society of Press and Display Industries (SIPA – Ouest France)
Society of Cinema and Television Producers (PROCIREP)
French Society for Written Authors' Interests (SOFIA)
Independent Online Information Press Union (SPIIL)
Heritage Film Catalog Union (SCFP)
French Union of Performing Artists (SFA-CGT)
National Publishing Union (SNE)
National Union of Authors and Composers (SNAC)
National union of phonographic publishers (SNEP)
National Union of Musical Artists (SNAM-CGT)
TF1
Treppoz Edouard (Paris 1)
TrustMyContent
United Voice Artists (UVA)

Appendices

Box 1 - Language models, diffusion models and measures of collapse

Collapse corresponds to the diminution of synthetic data quality when new generation models are trained on all or a strong proportion of synthetic data from the previous generation of models. It has been demonstrated in numerous AI models, notably large language models of the GPT type and in diffusion models for image generation. What metric is used for these two types of models? And how does it reflect collapse?

A) Language models and perplexity

A language model assigns probabilities to arbitrary symbol sequences such that the more likely a sequence is to exist in that language, the higher the assigned probability. A symbol can be a character, a word or a sub-unit (for example, the word "read" can be divided into two sub-units: "re" and "ad"). Most language models estimate this probability as a product of the probability of each symbol given its preceding symbols.

Given a sequence (w_1,w_2,…,w_n), the probability of this sequence is given by the following product:

$$P(w_1, w_2, \dots, w_n) = p(w_1)p(w_2|w_1)p(w_3|w_1, w_2) \dots p(w_n|w_1, w_2, \dots, w_{n-1}) =$$

$$\prod_{i=1}^{n} p(w_i|w_1, \dots, w_{i-1})$$

("$p(w_2|w_1)$" reads as "the probability of occurrence of w_2 given the occurrence of w_1". We calculate the probability of element w_2 knowing element w_1.)

In other words, the probability of the sequence "S = I like reading this report" can be calculated as follows:
P(S) = P(I) × P(like|I) × P(reading|I like) × P(this|I like reading) × P(report|I like reading this)

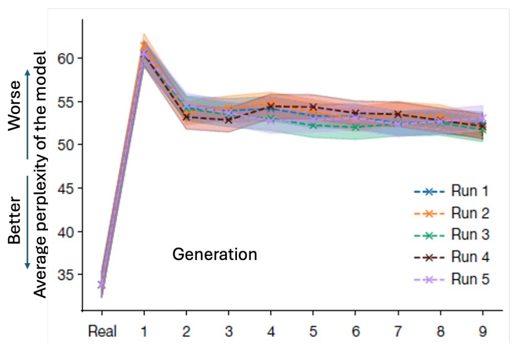
Perplexity is a measure quantifying how well a probability (e.g. generated by a language model) predicts a sample (e.g. the word "report" after "I like reading this"). Perplexity is a measure of uncertainty relative to the occurrence of the next symbol. Mathematically, it is represented as:

$$Perplexity\ (S) = exp\left\{-\frac{1}{n}\sum_{i}^{n} log\ p_{\theta}(w_i|w_{<i})\right\}$$

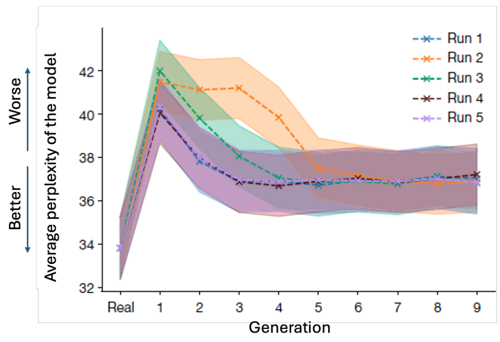
where log P_θ is the logarithm of the likelihood of obtaining symbol w_i given the preceding symbols. Intuitively, it's an evaluation of the model's ability to predict sequences from a corpus.

The graphs below illustrate the evolution of perplexity when a next-generation model is trained on synthetic data from the previous generation model exclusively (left) or on mixed data composed of 10% real data and 90% synthetic data from the previous generation model (right). These graphs are adapted from one of the studies presented in the report by (Shumailov et al., 2024).

Perplexity of the model without preserved real data



Perplexity of the model with 10% of real data preserved at each generation



The term "run" refers to an "experiment": the authors conducted five experiments, that is, they trained generations of models based on synthetic data from previous generations, and did so five times. The objective of this approach is to ensure statistical regularity of observations and replicability of results.

We observe that the perplexity of the model trained on real data ("real") is lower than the perplexity of models trained on purely synthetic or mixed data (from generation 1 to 9).

B) Diffusion models and Fréchet inception distance (FID)

Several types of models allow image generation. Diffusion models consist of inducing noise (i.e. "scrambling") in images through successive steps, and training the model to denoise images until retrieving the original image.

There are different measures of the quality of images generated by the model. One of them consists of quantifying how much the statistical distribution of images generated by models has moved away from the statistical distribution of images present in the data. This is the Wasserstein distance, which measures the minimum work required to move a probability density from one distribution to another. In practice, this distance is approximated by the Fréchet distance, applied to characteristics extracted for each image by another model already trained for this intended purpose (Inception, a convolutional neural network) rather than from each pixel. Hence the name of the metric (Fréchet Inception Distance, FID).

The idea is not to work on the pixels of the image, but to extract characteristics of the image, such as structures, shapes, textures, from angles to objects. A set of characteristics is thus obtained for each image. These characteristics are quantified and form statistical distributions at the database scale, for real data on one hand, and for synthetic data on the other hand. The distributions can be compared between these two types of data and can be characterised by their means (noted μ_r, μ_s for real data and synthetic data, respectively), and their covariance (noted Σ_r, Σ_s for real data and synthetic data, respectively). Mathematically:

$$FID = \|\mu_r - \mu_s\|^2 + Tr\left(\Sigma_r + \Sigma_s - 2\sqrt{\Sigma_r \Sigma_s}\right)$$

where $\|\mu_r - \mu_s\|^2$ measures the distance between the means of characteristics of real data and synthetic data; Tr(.) represents the trace of the matrix, and will evaluate the difference in variability between generated data and real data. Indeed, $Tr(\Sigma_r + \Sigma_s)$ reflects how much the two distributions are dispersed and $Tr(-2\sqrt{\Sigma_r \Sigma_s})$ quantifies the similarity of dispersion between the two distributions.

A low FID score indicates that the generated data is close to real data.

Data diversity is measured by recall. Recall estimates the fraction of samples in a reference distribution that are found in the support of the distribution learned by a generative model (i.e. real values). High recall scores indicate that the generative model captures a large part of the diverse samples from the reference distribution.

The graphs below illustrate the evolution of FID (left) when a next-generation model is trained on synthetic data from the previous generation model exclusively (in orange) or on mixed data composed of initial real data plus all synthetic data accumulated across generations. The right graph represents the evolution of image diversity (recall). These graphs are adapted from one of the studies presented in the report Alemohammad et al., (2023).

We observe that the quality (FID) of the model trained on synthetic data decreases (FID increases) over generations of models trained exclusively on synthetic data (in orange) and, albeit to a lesser extent, for the accumulation of synthetic data alongside initial real data (in blue). Diversity follows the same pattern.

Box 2 – Sources of error in the collapse process

The collapse process ***results from three sources of errors that accumulate across generations*** *and cause deviations from the original model (Shumailov et al., 2024).*

Statistical approximation error initiates the model collapse *process by eliminating extreme data, particularly affecting rare events (i.e., in statistical terms, by eliminating the tails of the data distribution and other fine statistical details of the distribution). When a model generates data to train the next generation of models, it may not include rare or low-probability events (such as unusual words or word sequences) that were present in the original data. Over time, these rare events disappear, and the model's understanding of the original data distribution narrows, leading to performance degradation. This error causes the disappearance of the data distribution tails, meaning the model focuses more on common events and forgets rare ones, thus leading to its* ***early collapse****.*

Functional expressivity error *affects the early stages by limiting the model's ability to capture the complexity of the original data. Even if a model receives perfect training data, it may not have sufficient expressive power to accurately capture the underlying real distribution. In other words, the model may be too simple to account for all statistical properties of the data. Attempting to approximate a complex distribution (for example, a mixture of two Gaussian distributions) with a simpler model (for example, a single Gaussian) will result in errors in how the model represents this distribution. This error occurs mainly during the first generation of the model training process and can lead to incorrect representations of real data, which then propagate through subsequent generations. While it cannot alone fully explain model collapse, it aggravates the problem in the early stages. Attempting to counter this effect by increasing model complexity can have a perverse effect of explaining noise in the data and thus generating generalisation errors. In other words, a model that is too simple fails to capture all the nuance of the data, while a model that is too complex risks capturing stochastic noise (for example, a statistical regularity like a particular sequence of sentences in a text specific to a dataset but which does not exist in general).*

Functional approximation error *results from limitations in model learning procedures, which can induce biases, for example during gradient descent or depending on the choice of objective. These biases lead to progressive deviation from the original data, particularly in later generations. Even when the model has sufficient expressive power and abundant training data, the learning process itself—for example, how the model optimises its parameters, or the choice of what is optimised (e.g., minimising average error, maximising likelihood)—introduces biases that can lead to additional deviation from the original data distribution. For example, the choice of what is optimised (the "objective function") plays a critical role. If the objective function used to train the model only minimises average error, the model estimation procedure may ignore important aspects, such as preserving rare events (the "tails" of the distribution). These biases accumulate over generations, as models are continually adjusted on data generated by previous models, leading to divergence from the original distribution and* ***late model collapse****. As models progressively fail to capture the original data distribution, they collapse and produce degenerate or simplified outputs with low variance.*

Box 3: The Shapley value method applied to protected cultural datasets

Wang et al.'s 2024 article uses the Shapley value to calculate the extent to which each copyright holder's data contributed to the success of a generative AI model in creating specific content. Based on this article, we propose to illustrate how the Shapley value could be calculated in a simple example, devised for the occasion.

Imagine a model trained on data from several copyright holders. We want to determine the extent to which each holder's data contributed to the model's ability to generate specific content. The Shapley value helps answer this question by taking into account all possible subsets and determining the extent to which each holder's data increases the model's utility when added to various combinations of data subsets.

Consider the step-by-step process with an example of three copyright holders, A, B, and C, each having contributed data to train the model. We can generate an image with AI like DALL-E then determine the extent to which each owner's data contributed to this work. Suppose that:

- A owns a dataset of landscapes.
- B owns a dataset of portraits.
- C owns a dataset of abstract art.

In Wang et al.'s 2024 article, utility is a measure of the model's performance in generating a specific result (for example, a specific artwork). Utility can be viewed as the probability that the model generates the same artwork using data from a given subset of owners.

Suppose we measure the utility of different subsets as follows:

- Utility of subset {A, B, C} (the complete set): 100 (this is the utility of the fully trained model).
- Utility of subset {A, B}: 80 (A and B together generate content well, but not as well as the complete model).
- Utility of subset {A, C}: 80.
- Utility of subset {B, C}: 60.
- Utility of subset {A}: 50.
- Utility of subset {B}: 40.
- Utility of subset {C}: 30.
- Utility of the empty set {}: 0 (no data, no model).

The key element for calculating the Shapley value is examining each data owner's marginal contribution to different subsets. The marginal contribution is the measure of how adding a particular owner's data improves the model's utility.

For example:

- Adding A to the empty set {A} increases utility from 0 to 50, so A's marginal contribution is 50 in this case.
- Adding B to {A} increases utility from 50 to 80, so B's marginal contribution is 30.
- Adding C to {A, B} increases utility from 80 to 100, so C's marginal contribution is 20.

The Shapley value is calculated by averaging each copyright holder's marginal contribution across all possible ways of combining the owners' data. The formula is:

$$\phi i = \frac{1}{n}\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} \sum_{S \subseteq N\{i\}} [v(S \cup \{i\} - v(S)]$$

Where:

- n is the total number of copyright holders.
- S represents a subset of copyright holders.
- v(S) is the utility of subset S.
- The sum is performed over all subsets of owners that do not include i.

The left sum indicates that we must sum across all possible combinations of contribution order for rightholder i.

The right sum corresponds to the marginal contribution of rightholder i (this is the difference between the utility of data with i minus the utility of data without i).

Now, let us calculate the Shapley value for each holder (A, B, C) by averaging their marginal contributions across all subsets.

For example:

For A:

- A's contribution when added to {}: 50 (utility goes from 0 to 50).
- A's contribution when added to {B}: 30 (from 40 to 80).
- A's contribution when added to {C}: 50 (from 30 to 80).
- A's contribution when added to {BC}: 40
- A's average marginal contribution: (50+30+50+40)/4 = 42.5

For B:

- B's contribution when added to {}: 40.
- B's contribution when added to {A}: 30.
- B's contribution when added to {C}: 30.
- B's contribution when added to {AC}: 20.
- B's average marginal contribution: (40+30+30+30)/4 = 32.5

For C:

- C's contribution when added to {}: 30.
- C's contribution when added to {A}: 30.
- C's contribution when added to {B}: 20.
- C's contribution when added to {A,B}: 20.
- C's average marginal contribution: (30+30+20+20)/4 = 25

Thus, the Shapley values would be:

- $\Phi(A) = 42.5$
- $\Phi(B) = 32.5$
- $\Phi(C) = 25$

Based on these Shapley values, copyright holders would receive compensation proportional to their contributions. For example, if the work generated by the model earns $100, the gains could be distributed as follows:

- A receives approximately $42.5.
- B receives approximately $32.5.
- C receives approximately $25.

Box 4: Different methods for estimating training data contribution to generative AI output

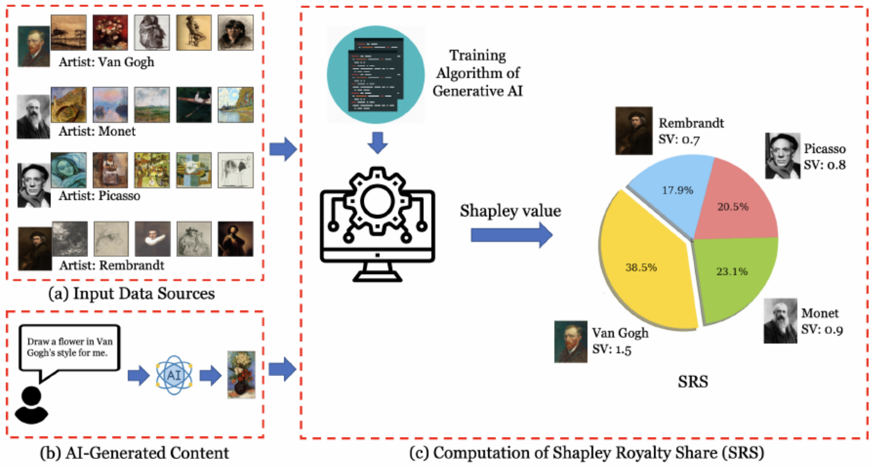


Figure 1. Overview of Shapley value calculation (causal approach)
We reproduce here the graphics from (J. T. Wang, Deng, et al. 2024) for methodological illustration purposes only. The painters cited in this article are in the public domain, and so are not concerned by the value-sharing question that is the subject of this task force.
a: suppose four datasets are used to train an AI model
b: a user uses an interface to submit a query to the generative AI model: "draw a flower in the style of Van Gogh".
c: The Shapley value can be calculated to determine the share due to each artist in the generated output. The model is first trained with Van Gogh's data, then fine-tuned on Monet's, then on Picasso's, then Rembrandt's, and also in various orders to estimate each author's marginal contribution for a given output. This method has a prohibitive cost and requires approximations.

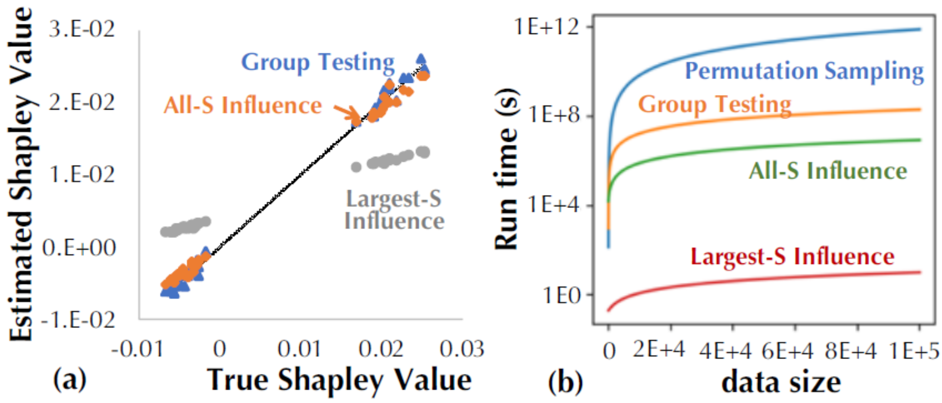


Figure 2. Comparison of different Shapley value approximation methods (causal approach)
From (Jia et al., 2019).

"Permutation sampling" corresponds to training models on a random portion of data subsets; "Group testing" corresponds to training models on several data subsets following "intelligent" sampling based on group theory; All-S and Largest-S Influence correspond to using the "Influence function" method consisting of studying the importance of a data point by overweighting it in models trained on all data (Largest-S) or on subsets (All-S).

Figure 2A shows the estimation of the Shapley value by the approximation method as a function of the true Shapley value. If the approximation is precise, we expect each point to be aligned with the black line. This is the case for the "group testing" and "All-S influence" methods, but not for "Largest-S influence."

Figure 2B shows the computation time for each method (on a logarithmic scale), as a function of dataset size. Overall, across the two figures, we observe that fast methods are imprecise, and vice versa.

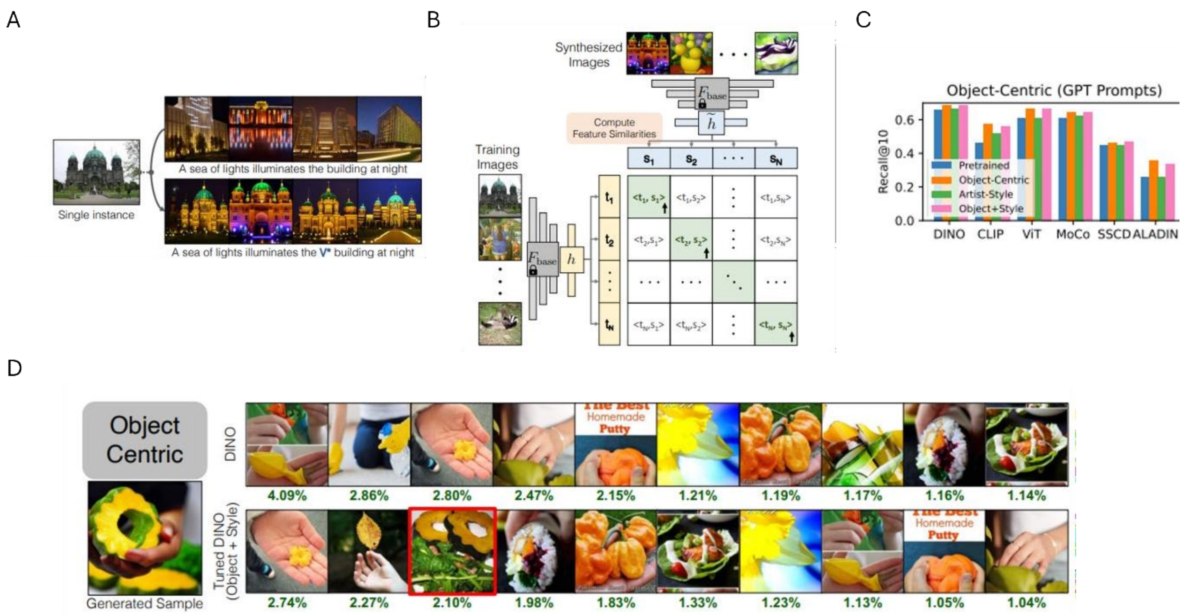


Figure 3. Example of using image characteristics to calculate similarity with model output (correlational approach). From (S.-Y. Wang et al., 2023).

A: Illustration of the fine-tuning phase on a source image, in this case a specific building. Asking the AI model for images of lit buildings produces various lit buildings, while asking it to generate buildings based on the source image generates different lit versions of this source building.
B: Calculation of similarity between generated images (table columns) based on the source image (table rows). We expect strong similarity on the table diagonals, i.e., generated images should resemble the source image on which the model is fine-tuned more than any other source image.
C: To verify this prediction, the authors calculate the frequency of source images in the top 10 images of the training + fine-tuning dataset (Recall@10), for each major model type (DINO, CLIP, etc.). The blue bar concerns the pre-trained model's performance, the orange bar is the one that interests us in this example: it should be higher than the blue bar for the attribution method to be considered correct.
D: In this example, we can see that the source image, framed in red, is indeed in the top 10 images most similar to the generated image. We also see that other images, though not used, are similar and would therefore (wrongly) be credited with compensation.

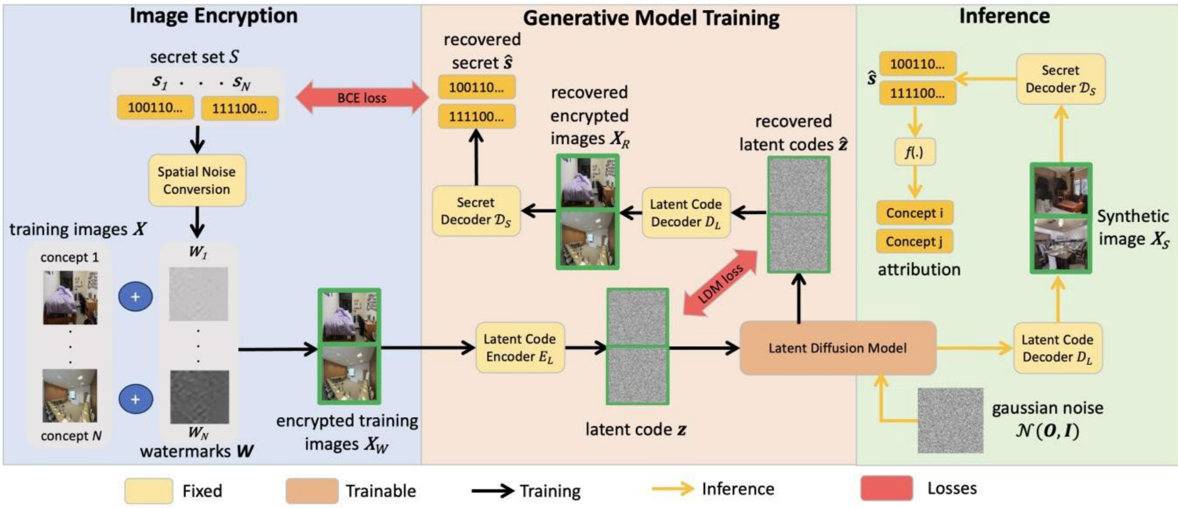


Figure 4. The watermarking method (causal approach)
From (Asnani et al., 2024).
In the first phase, a watermark "invisible" to the naked eye is associated with each image. In the second phase, the model is trained on marked training data; the marking is retrieved. In the inference phase, i.e., generation of an output image upon a user's request, it is possible to decode the images that contributed to the genesis of the synthesised images.

Bibliography

A three-step design pattern for specializing LLMs. (n.d.). Google Cloud Blog. Retrieved 12 March 2025, from https://cloud.google.com/blog/products/ai-machine-learning/three-step-design-pattern-for-specializing-llms

Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, *40*(S1), S293–S340. https://doi.org/10.1086/718327

Aghion, P., & Bouverot. (2024). *IA : notre ambition pour la France*. Comission de l'intellignece artificielle.

AI Foundation Models: Update paper. (2024, April 16). GOV.UK. https://www.gov.uk/government/publications/ai-foundation-models-update-paper

Alemohammad, S., Casco-Rodriguez, J., Luzi, L., Humayun, A. I., Babaei, H., LeJeune, D., Siahkoohi, A., & Baraniuk, R. G. (2023). *Self-Consuming Generative Models Go MAD* (arXiv:2307.01850). arXiv. https://doi.org/10.48550/arXiv.2307.01850

Asnani, V., Collomosse, J., Bui, T., Liu, X., & Agarwal, S. (2024). ProMark: Proactive Diffusion Watermarking for Causal Attribution. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10802–10811. https://openaccess.thecvf.com/content/CVPR2024/html/Asnani_ProMark_Proactive_Diffusion_Watermarking_for_Causal_Attribution_CVPR_2024_paper.html

Audiens data lab report. (n.d.). IA: mesure de l'évolution des métiers du doublage. Retrieved 10 November 2024, from https://app.powerbi.com/view?r=eyJrIjoiZTA1NGU0OTctMjRjNC00Y2VlLWExMDktNzk0YTRkMjg3NGVjIiwidCI6ImQwMGVkZjI4LTUyOTUtNGJmZS05YTU5LTM1NjA4NWMzZGU3YyJ9

Avis 24-A-05 du 28 juin 2024. (2024, June 28). Autorité de la concurrence. https://www.autoritedelaconcurrence.fr/fr/avis/relatif-au-fonctionnement-concurrentiel-du-secteur-de-lintelligence-artificielle-generative

Benabou, V.-L. (2018, November 14). *Mission du CSPLA sur les conséquences pour la propriété littéraire et artistique de l'avènement des notions de données et de contenus numériques*. https://www.culture.gouv.fr/nous-connaitre/organisation-du-ministere/Conseil-superieur-de-la-propriete-litteraire-et-artistique-CSPLA/Travaux-et-publications-du-CSPLA/Missions-du-CSPLA/Mission-du-CSPLA-sur-les-consequences-pour-la-propriete-litteraire-et-artistique-de-l-avenement-des-notions-de-donnees-et-de-contenus-numeriques

Benabou, V.-L. (2023, December 11). Du test en trois étapes au domaine public payant- Quelques idées pour mieux associer les titulaires de droit à la production des Intelligences Artificielles génératives dans le champ de la création intellectuelle. *Le 'chat' et la souris*. https://vlbenabou.blog/2023/12/11/168/

Bensamoun, A., & Farchy, J. (2020). *Intelligence artificielle et culture* [Rapport du CSPLA]. Ministère de la culture.

Bertrand, Q., Bose, A. J., Duplessis, A., Jiralerspong, M., & Gidel, G. (2024). *On the Stability of Iterative Retraining of Generative Models on their own Data* (arXiv:2310.00429). arXiv. https://doi.org/10.48550/arXiv.2310.00429

Besson, L., Dozias, A., Faivre, C., Gallezot, C., Gouy-Waz, J., & Vidalenc, B. (2024). *Les enjeux économiques de l'intelligence artificielle* (341; Trésor-Eco). Direction Générale du Trésor. https://www.tresor.economie.gouv.fr/Articles/2024/04/02/les-enjeux-economiques-de-l-intelligence-artificielle

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale. *2023 ACM Conference on Fairness, Accountability, and Transparency*, 1493–1504. https://doi.org/10.1145/3593013.3594095

Bohacek, M., & Farid, H. (2023). *Nepotistically Trained Generative-AI Models Collapse* (arXiv:2311.12202). arXiv. http://arxiv.org/abs/2311.12202

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems*, *29*. https://proceedings.neurips.cc/paper_files/paper/2016/hash/a486cd07e4ac3d270571622f4f316ec5-Abstract.html

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. https://proceedings.mlr.press/v81/buolamwini18a.html

Business Insights and Conditions Survey Team. (2024). *Business insights and impact on the UK economy—Office for National Statistics*. Office for National Statistics. https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/bulletins/businessinsightsandimpactontheukeconomy/3october2024

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. https://doi.org/10.1126/science.aal4230

Cantor, M. (2023). Nearly 50 news websites are 'AI-generated', a study says. Would I be able to tell. *The Guardian, May*.

Cazzaniga, M., Jaumotte, M. F., Li, L., Melina, M. G., Panton, A. J., Pizzinelli, C., Rockall, E. J., & Tavares, M. M. M. (2024). *Gen-AI: Artificial intelligence and the future of work*. International Monetary Fund. https://books.google.com/books?hl=en&lr=&id=YLXuEAAAQBAJ&oi=fnd&pg=PA2&dq=Gen-AI:+Artificial+Intelligence+and+the+Future+of+Work&ots=OOSfbJ70wz&sig=X5E0FQ-27qQLdeKPYMME4nh6Cek

Chin-Rothmann, C. (2023). *Navigating the Risks of Artificial Intelligence on the Digital News Landscape*. https://www.csis.org/analysis/navigating-risks-artificial-intelligence-digital-news-landscape

Cohen, C. (2024, May 26). *«L'IA a volé mon travail du jour au lendemain»: Ces traducteurs et doubleurs déjà remplacés dans le monde de l'édition*. Le Figaro. https://www.lefigaro.fr/medias/l-ia-a-vole-mon-travail-du-jour-au-lendemain-ces-traducteurs-et-doubleurs-que-l-ia-a-deja-remplaces-dans-le-monde-de-l-edition-20240526

Dastin, J. (2022). Amazon scraps secret AI recruiting tool that showed bias against women. In *Ethics of data and analytics* (pp. 296–299). Auerbach Publications. https://www.taylorfrancis.com/chapters/edit/10.1201/9781003278290-44/amazon-scraps-secret-ai-recruiting-tool-showed-bias-women-jeffrey-dastin

Deng, C., Feng, S., Wang, H., Zhang, X., Jin, P., Feng, Y., Zeng, Q., Chen, Y., & Lin, Y. (2022). OpenFWI: Large-scale multi-structural benchmark datasets for full waveform inversion. *Advances in Neural Information Processing Systems*, *35*, 6007–6020.

Deng, J., Zhang, S., & Ma, J. (2024). *Computational Copyright: Towards A Royalty Model for Music Generative AI* (arXiv:2312.06646). arXiv. https://doi.org/10.48550/arXiv.2312.06646

Directive (UE) 2019/790 du Parlement européen et du Conseil du 17 avril 2019 sur le droit d'auteur et les droits voisins dans le marché unique numérique et modifiant les directives 96/9/CE et 2001/29/CE (Texte présentant de l'intérêt pour l'EEE.)—Légifrance. (n.d.). Retrieved 22 November 2024, from https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000038481211

Distilling step-by-step: Outperforming larger language models with less training. (n.d.). Retrieved 12 March 2025, from https://research.google/blog/distilling-step-by-step-outperforming-larger-language-models-with-less-training-data-and-smaller-model-sizes/

Dohmatob, E., Feng, Y., Yang, P., Charton, F., & Kempe, J. (2024). *A Tale of Tails: Model Collapse as a*

Change of Scaling Laws (arXiv:2402.07043). arXiv. https://doi.org/10.48550/arXiv.2402.07043
Elias, G. (2024, June 11). *Profil de l'entreprise Runway ML - AI&YOU #44*. Skim AI. https://skimai.com/fr/aiyou-44-notre-profil-dentreprise-de-la-piste-ml/
Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2023). *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models* (arXiv:2303.10130). arXiv. http://arxiv.org/abs/2303.10130
European Commission. (2021). *EU Artificial Intelligence Act | Up-to-date developments and analyses of the EU AI Act*. European Commission. https://artificialintelligenceact.eu/
Gault, M. (2023). AI spam is already flooding the internet and it has an obvious tell. *VICE, April*.
Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., & Brendel, W. (2022). *ImageNet-trained CNNs are biased towards texture; increasing shape bias improves accuracy and robustness* (arXiv:1811.12231). arXiv. https://doi.org/10.48550/arXiv.1811.12231
GenAI Attribution Simulator—A Hugging Face Space by TheFrenchDemos. (2025). https://huggingface.co/spaces/TheFrenchDemos/genai-reward
Gerstgrasser, M., Schaeffer, R., Dey, A., Rafailov, R., Sleight, H., Hughes, J., Korbak, T., Agrawal, R., Pai, D., Gromov, A., Roberts, D. A., Yang, D., Donoho, D. L., & Koyejo, S. (2024). *Is Model Collapse Inevitable? Breaking the Curse of Recursion by Accumulating Real and Synthetic Data* (arXiv:2404.01413). arXiv. https://doi.org/10.48550/arXiv.2404.01413
Glickman, M., & Sharot, T. (2024). AI-induced hyper-learning in humans. *Current Opinion in Psychology*, *60*, 101900. https://doi.org/10.1016/j.copsyc.2024.101900
Growcoot, M. (2023, July 12). *Shutterstock May Have Paid Out Over $4 Million From its AI Contributor Fund*. PetaPixel. https://petapixel.com/2023/07/12/shutterstock-may-have-paid-out-over-4-million-from-its-ai-contributor-fund/
Hammoudeh, Z., & Lowd, D. (2024). Training data influence analysis and estimation: A survey. *Machine Learning*, *113*(5), 2351–2403. https://doi.org/10.1007/s10994-023-06495-7
Hoppner, T., & Streatfeild, L. (2023). ChatGPT, Bard & Co.: An introduction to AI for competition and regulatory lawyers. *An Introduction to AI for Competition and Regulatory Lawyers (February 23, 2023)*, *9*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4371681
House of Lords Communications and Digital Select Committee inquiry. (2023, December 5). *OpenAI—written evidence (LLM0113)*. https://committees.parliament.uk/writtenevidence/126981/pdf/
Hughes, A. (2023, December 12). Phi-2: The surprising power of small language models. *Microsoft Research*. https://www.microsoft.com/en-us/research/blog/phi-2-the-surprising-power-of-small-language-models/
Hughes, A. (2024, March 5). Orca-Math: Demonstrating the potential of SLMs with model specialization. *Microsoft Research*. https://www.microsoft.com/en-us/research/blog/orca-math-demonstrating-the-potential-of-slms-with-model-specialization/
IA générative et droit d'auteur: Quelle place pour les données européennes protégées à l'ère de l'IA ? (2024). France Digitale. https://francedigitale.org/publications/ia-generative-et-droit-dauteur
Inquiets de l'utilisation de l'IA, les acteurs et doubleurs de jeux vidéo vont faire grève en Californie. (2024, July 26). Le Figaro. https://www.lefigaro.fr/medias/inquiets-de-l-utilisation-de-l-ia-les-acteurs-et-doubleurs-de-jeux-video-vont-faire-greve-en-californie-20240726
Jia, R., Dao, D., Wang, B., Hubis, F. A., Hynes, N., Gürel, N. M., Li, B., Zhang, C., Song, D., & Spanos, C. J. (2019). Towards efficient data valuation based on the shapley value. *The 22nd International Conference on Artificial Intelligence and Statistics*, 1167–1176. https://proceedings.mlr.press/v89/jia19a.html
Johann, A., Drazilova, M., Treweller, S., Möhlen, J., Brusoni, S., & Fehr, E. (2023). The Value of Journalistic Content for the Google Search Engine in Switzerland. *FehrAdvice & Partners AG*.
Kazdan, J., Schaeffer, R., Dey, A., Gerstgrasser, M., Rafailov, R., Donoho, D. L., & Koyejo, S. (2024).

Collapse or Thrive? Perils and Promises of Synthetic Data in a Self-Generating World (arXiv:2410.16713). arXiv. https://doi.org/10.48550/arXiv.2410.16713

Khan, F. (2024, November 1). How OpenAI and Anthropic Are Cashing In on AI: A Look at Their Revenue Models. *Medium*. https://medium.com/@furqankhaan/how-openai-and-anthropic-are-cashing-in-on-ai-a-look-at-their-revenue-models-d9d9ae79dd28

Klemp, M., Rösch, K., Wagner, R., Quehl, J., & Lauer, M. (2023). LDFA: Latent diffusion face anonymization for self-driving applications. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3199–3205. https://openaccess.thecvf.com/content/CVPR2023W/E2EAD/html/Klemp_LDFA_Latent_Diffusion_Face_Anonymization_for_Self-Driving_Applications_CVPRW_2023_paper.html

Koh, P. W., & Liang, P. (2017). Understanding black-box predictions via influence functions. *International Conference on Machine Learning*, 1885–1894. http://proceedings.mlr.press/v70/koh17a?ref=https://githubhelp.com

La SPRE collecte la rémunération équitable pour les artistes-interprètes et les producteurs de phonogrammes. (n.d.). SPRE. Retrieved 12 March 2025, from https://www.spre.fr/

Laird, J. (2025, January 28). China's DeepSeek chatbot reportedly gets much more done with fewer GPUs but Nvidia still thinks it's 'excellent' news. *PC Gamer*. https://www.pcgamer.com/hardware/graphics-cards/chinas-deepseek-chatbot-reportedly-gets-much-more-done-with-fewer-gpus-but-nvidia-still-thinks-its-excellent-news/

L'ATLF a interrogé ses adhérents sur la post-édition. (2022). ATLF. https://atlf.org/latlf-a-interroge-ses-adherents-sur-la-post-edition/

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés—Légifrance. (n.d.). Retrieved 25 November 2024, from https://www.legifrance.gouv.fr/loda/id/LEGISCTA000006095896

Lorphelin, V. (2024). *Un modèle économique pour valoriser le capital culutrel et humain par l'IA générative*. Controv3rse.

Loud and Clear by Spotify. (2023). Loud and Clear. https://loudandclear.byspotify.com/

Lutes. (2025). *Identifying the Economic Implications of Artificial Intelligence for Copyright Policy*.

Luzi, L., Mayer, P. M., Casco-Rodriguez, J., Siahkoohi, A., & Baraniuk, R. G. (2024). *Boomerang: Local sampling on image manifolds using diffusion models* (arXiv:2210.12100). arXiv. https://doi.org/10.48550/arXiv.2210.12100

Maleki, S., Tran-Thanh, L., Hines, G., Rahwan, T., & Rogers, A. (2014). *Bounding the Estimation Error of Sampling-based Shapley Value Approximation* (arXiv:1306.4265). arXiv. https://doi.org/10.48550/arXiv.1306.4265

Martínez, G., Watson, L., Reviriego, P., Hernández, J. A., Juarez, M., & Sarkar, R. (2023). *Combining Generative Artificial Intelligence (AI) and the Internet: Heading towards Evolution or Degradation?* (arXiv:2303.01255). arXiv. https://doi.org/10.48550/arXiv.2303.01255

Nasiripour, S., & Natarajan, S. (2019). Apple co-founder says Goldman's Apple card algorithm discriminates. *Bloomberg. Com*.

Nellis, S., & read, K. F. U. 2 min. (2025, January 27). *Nvidia says DeepSeek advances prove need for more of its chips*. Yahoo Finance. https://uk.finance.yahoo.com/news/nvidia-says-deepseek-advances-prove-192658104.html

« Non, l'intelligence artificielle ne remplacera pas les traducteurs et les traductrices! ». (2024, September 9). https://www.lemonde.fr/idees/article/2024/09/09/non-l-intelligence-artificielle-ne-remplacera-pas-les-traducteurs-et-les-traductrices_6308656_3232.html

Online Nation 2023 Report. (2023). Ofcom.

Peukert, C., Abeillon, F., Haese, J., Kaiser, F., & Staub, A. (2024). *Strategic Behavior and AI Training Data* (arXiv:2404.18445). arXiv. http://arxiv.org/abs/2404.18445

Pinaya, W. H. L., Tudosiu, P.-D., Dafflon, J., Da Costa, P. F., Fernandez, V., Nachev, P., Ourselin, S., & Cardoso, M. J. (2022). Brain Imaging Generation with Latent Diffusion Models. In A. Mukhopadhyay, I. Oksuz, S. Engelhardt, D. Zhu, & Y. Yuan (Eds.), *Deep Generative Models* (Vol. 13609, pp. 117–126). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-18576-2_12

Racine, B. (2020). *L'auteur et l'acte de création*. Ministère de la culture.
Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10684–10695. http://openaccess.thecvf.com/content/CVPR2022/html/Rombach_High-Resolution_Image_Synthesis_With_Latent_Diffusion_Models_CVPR_2022_paper.html
Sacra. (2024). *Anthropic revenue, valuation & growth rate*. https://sacra.com/c/anthropic/
Sacra. (2024). *OpenAI revenue, valuation & growth rate*. Open AI. https://sacra.com/c/openai/
Sadeghi, M., Arvanitis, L., Padovese, V., Pozzi, G., Badilini, S., Varcellone, C., Wang, M., Brewster, J., Huet, N., Fishman, Z., Pfaller, L., Adams, N., & Wollen, M. (2024). *Tracking AI-enabled Misinformation: Over 1100 'Unreliable AI-Generated News' Websites (and Counting), Plus the Top False Narratives Generated by Artificial Intelligence Tools*. NewsGuard. https://www.newsguardtech.com/special-reports/ai-tracking-center
Senftleben, M. (2024). *AI Act and Author Remuneration—A Model for Other Regions?* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4740268
Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). AI models collapse when trained on recursively generated data. *Nature*, *631*(8022), 755–759. https://doi.org/10.1038/s41586-024-07566-y
Skjuve, M. (2023). Why people use chatgpt. *Available at SSRN*, *4376834*.
SoA survey reveals a third of translators and quarter of illustrators losing work to AI - The Society of Authors. (2024, April 11). https://societyofauthors.org/2024/04/11/soa-survey-reveals-a-third-of-translators-and-quarter-of-illustrators-losing-work-to-ai/
Spotify launches revenue-sharing Partner Program. (2025, January 3). https://podnews.net/update/spotify-partner-program-launch
Thomas, A. (2025, February 24). *The AI licensing economy – CREATe*. https://www.create.ac.uk/blog/2025/02/24/the-ai-licensing-economy/
Thomas, P.-L. (2023, May 25). *Menacés par l'intelligence artificielle, les comédiens de doublage s'unissent dans un collectif international*. https://www.lemonde.fr/economie/article/2023/05/25/menaces-par-l-intelligence-artificielle-les-comediens-de-doublage-s-unissent-dans-un-collectif-international_6174777_3234.html
Trésor, D. générale du. (2024, December 5). *La chaîne de valeur de l'intelligence artificielle: Enjeux économiques et place de la France*. Direction générale du Trésor. https://www.tresor.economie.gouv.fr/Articles/2024/12/05/la-chaine-de-valeur-de-l-intelligence-artificielle-enjeux-economiques-et-place-de-la-france
Troyanskaya, O., Trajanoski, Z., Carpenter, A., Thrun, S., Razavian, N., & Oliver, N. (2020). Artificial intelligence and cancer. *Nature Cancer*, *1*(2), 149–152.
Update on KDP Title Creation Limits. (2023). https://www.kdpcommunity.com/s/article/Update-on-KDP-Title-Creation-Limits?language=en_US&forum=KDP%20Forum
Veselovsky, V., Ribeiro, M. H., & West, R. (2023). *Artificial Artificial Artificial Intelligence: Crowd Workers Widely Use Large Language Models for Text Production Tasks* (arXiv:2306.07899). arXiv. https://doi.org/10.48550/arXiv.2306.07899
Vicente, L., & Matute, H. (2023). Humans inherit artificial intelligence biases. *Scientific Reports*, *13*(1), 15737.
Villalobos, P. (2022, November 10). *Will We Run Out of ML Data? Evidence From Projecting Dataset Size Trends*. Epoch AI. https://epochai.org/blog/will-we-run-out-of-ml-data-evidence-from-projecting-dataset
Vlasceanu, M., & Amodio, D. M. (2022). Propagation of societal gender inequality by internet search algorithms. *Proceedings of the National Academy of Sciences*, *119*(29), e2204529119.
Vulser, N. (2024, February 2). *Les traducteurs littéraires victimes de l'intelligence artificielle*. https://www.lemonde.fr/economie/article/2024/02/02/les-traducteurs-litteraires-victimes-de-l-intelligence-artificielle_6214361_3234.html

Wang, J. T., Deng, Z., Chiba-Okabe, H., Barak, B., & Su, W. J. (2024). *An Economic Solution to Copyright Challenges of Generative AI* (arXiv:2404.13964). arXiv. http://arxiv.org/abs/2404.13964

Wang, J. T., Mittal, P., Song, D., & Jia, R. (2024). *Data Shapley in One Training Run* (arXiv:2406.11011). arXiv. https://doi.org/10.48550/arXiv.2406.11011

Wang, S.-Y., Efros, A. A., Zhu, J.-Y., & Zhang, R. (2023). Evaluating data attribution for text-to-image models. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 7192–7203. http://openaccess.thecvf.com/content/ICCV2023/html/Wang_Evaluating_Data_Attribution_for_Text-to-Image_Models_ICCV_2023_paper.html

Wang, S.-Y., Hertzmann, A., Efros, A. A., Zhu, J.-Y., & Zhang, R. (2024). *Data Attribution for Text-to-Image Models by Unlearning Synthesized Images* (arXiv:2406.09408). arXiv. https://doi.org/10.48550/arXiv.2406.09408

Wenger, E. (2024). AI produces gibberish when trained on too much AI-generated data. *Nature*, *631*(8022), 742–743. https://doi.org/10.1038/d41586-024-02355-z