

Comment on Guidelines 1/2026 on Processing of Personal Data for Scientific Research Purposes

Submitted by: Meem Arafat Manab

Affiliation:

Marie Skłodowska-Curie Doctoral Fellow, Ontology Engineering Group, Universidad Politécnica de Madrid (HARNESS Doctoral Network, Horizon Europe).

Research Fellow (pro bono), School of Religion, Theology, and Peace Studies, Trinity College Dublin.

PhD research on data governance in European Data Spaces; prior research and applied work on machine learning systems, GDPR compliance, and the structural relationship between AI training and data protection rights (including arXiv:2403.05592 on the right to be forgotten and machine learning, and joint work with Dr. V. González-Castro on GDPR and linguistic data degradation, IFIP Summer School 2025).

Consultation: Guidelines 1/2026 on processing of personal data for scientific research purposes, adopted 15 April 2026

Deadline: 25 June 2026

Summary

This comment addresses three distinct gaps in the Guidelines. First, the Guidelines do not engage with the question of whether training of large language models and other foundation models on personal data constitutes "processing of personal data for scientific research purposes" within the scope of the GDPR, despite this being the most consequential current application of the scientific research framework in practice. Instead, the Guidelines' sole AI-related example addresses model evaluation, not training, and the harder case is left entirely unaddressed. Second, and separately, the six key-indicative factors in Section 2.1 are structured as cumulative and substitutable rather than individually necessary, a defect the Guidelines' own Example 2 demonstrates in practice rather than guards against. Third, and also separately, the factors are applied only once, at the point a controller first characterises its processing as research, with no mechanism requiring reassessment if its circumstances later change. The trajectory of OpenAI's own foundational documents between 2018 and 2026 illustrates concretely how this absence of a reassessment mechanism allows an initial research characterisation to persist long after the facts that justified it have changed, or worse, have become nullified.

Comment 1: The Guidelines do not address whether foundation model training is "scientific research"

Location: Section 2 generally; Section 2.1, factor v and vi (paras. 11, 12); Example 2 (pp. 10–11)

Issue:

The Guidelines were adopted in April 2026, with what we may only hope was a full awareness of the current state of AI development. They proceed to cite the AI Act once in a footnote (footnote 11) purely as a source for the general concept of "scientific research." Going against the tide of today, the Guidelines do not address the single most consequential and contested current application of the "scientific research" framework: whether training a foundation model (an LLM or other large-scale generative system) on personal data (scraped from the public web, or otherwise collected) qualifies as processing for scientific research purposes under the GDPR.

This omission is significant because (a) foundation model training involves personal data at unprecedented scale, and (b) foundation models trained on human-generated data, language, conversation, psychological and behavioural responses, are now a primary research method in the social and behavioural sciences. These are the fields where the data in question is most clearly personal data within the meaning of the GDPR. Researchers have fine-tuned foundation models on large corpora of human cognitive-psychology experiment data (Centaur, trained on Psych-101, comprising 10 million choices from 160 published cognitive-psychology experiments)¹ and on large-scale social-science survey response data (Socrates, trained on SocSci210, comprising 2.9 million social-science survey responses).² Meanwhile, a growing body of work on "behavioural foundation models," i.e., models trained to simulate human social behaviour at scale, draws on real dialogue corpora and social-interaction data to train systems intended to model affective expression, social cues, and conversational behaviour.³ This is a distinct and arguably more directly GDPR-relevant case than scientific foundation models trained on non-personal data such as protein structures or physical simulations, because the training data in these social and behavioural science applications is unambiguously about identifiable, or identifiable, human subjects, their choices, opinions, conversations, and psychological responses.

Given that foundation models of this kind are rapidly becoming central research infrastructure in psychology, sociology, and the broader behavioural sciences, the absence of any guidance on how the six key indicative factors apply to their training is not a narrow

¹Binz, M., Akata, E., Bethge, M., Brändle, F., Callaway, F., Coda-Forno, J., ... & Schulz, E. (2025). A foundation model to predict and capture human cognition. *Nature*, 644(8078), 1002-1009.

²Kolluri, A., Wu, S., Park, J. S., & Bernstein, M. S. (2025, November). Finetuning llms for human behavior prediction in social science experiments. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing* (pp. 30084-30099).

³See, e.g., Zhou, X., Sun, W., Du, W., Liu, J., Sun, H., Ma, Q., ... & Sap, M. (2026). OdysSim: Building Foundation Models for Human Behavior Simulation. *arXiv preprint arXiv:2606.14199*; Ong, K., Boughorbel, S., Xiao, L., Ekbote, C., Dai, W., Qu, A., ... & Liang, P. P. (2026). Omnisapiens: A foundation model for social behavior processing via heterogeneity-aware relative policy optimization. *arXiv preprint arXiv:2602.10635*.

gap but a substantial omission in a document that intends to bring legal clarity to the processing of personal data for scientific research.

This is not a hypothetical or marginal case either. Foundation model developers, including commercial entities, have argued in multiple jurisdictions that training runs constitute "scientific research" or research-adjacent activity, in order to invoke the presumption of purpose compatibility (Article 5(1)(b)), the storage limitation exception (Article 5(1)(e)), the legitimate interest balancing test (Article 6(1)(f) and para. 61 of the Guidelines, which states that "genuine scientific research is considered to be an important activity that is beneficial for the whole of society"), and the "manifestly made public" derogation for special categories of data (Article 9(2)(e) and Section 4.4.2 of the Guidelines).

Example 2 in the Guidelines (GSPM AI GmbH) does not address foundation model training at all. It addresses a narrower and easier case: a start-up evaluating bias in an already-trained generative AI model, not an entity collecting or processing personal data to train one. The harder and more consequential case, model training, is not addressed by Example 2 or anywhere else in the Guidelines. Lamentably, model training is the precise activity that determines what personal data a foundation model ingests, retains in its weights, and cannot straightforwardly erase.

The Guidelines should clarify, at least with one dedicated example, how the six key-indicative factors apply to the training, fine-tuning, and evaluation of foundation models on personal data, including:

- Whether bulk collection of personal data for pre-training (as distinct from a discrete, hypothesis-driven study of model behaviour) satisfies factor i (methodical and systematic approach), given that pre-training is generally not organised around a testable hypothesis but around maximising general model capability;
- Whether the "autonomy and independence" factor (factor iv) can be satisfied when the entity training the model is also the entity that will commercially deploy it, such that the research and product development functions are not meaningfully separable;
- Whether the "verifiability and transparency" factor (factor iii) can be satisfied when the training data, weights, and methodology are not disclosed, as is standard practice for most foundation model developers, notwithstanding the publication of an accompanying paper describing only partial aspects of the system;
- How the storage limitation exception (Section 3.2, paras. 27–31) applies to training datasets that are retained indefinitely to support continuous retraining and fine-tuning of a commercial product, rather than a defined future research project.

Suggested addition: A new example in Section 2, modelled on Example 2 but addressing pre-training or large-scale fine-tuning of a foundation model on personal data by a commercial developer, that works through the six factors with explicit attention to the

points above, and clarifies that satisfying the factors requires more than the procedural trappings of research (qualified staff, an ethics review, eventual publication) where the underlying activity is indistinguishable in substance from ordinary commercial AI development.

Comment 2: The key-indicative factors test does not adequately distinguish research from research-framed commercial data processing

Location: Section 2.1, paras. 11–12; Example 2 and Example 3 (pp. 10–12); Section 5.5, paras. 108, 110

This comment identifies two independent defects in how the six key-indicative factors operate, both arising from para. 12.

2(a). The factors are cumulative and substitutable, so a commercial entity can satisfy them while functioning, in substance, as a product company rather than an independent researcher

Issue:

Para. 12 states that "the more key-indicative factors are present, the more likely it is that the activities constitute scientific research." This treats the factors as cumulative and substitutable, rather than as individually necessary conditions. In practice, an entity can compensate for a weak claim on one factor (limited independence from commercial interests, for example) by strengthening another (obtaining ethics board approval, for example), even where the factor being strengthened does nothing to address the underlying concern. An ethics board review of a research protocol does not establish that the entity conducting the research is independent of its own commercial interest in the outcome. It establishes only that a separate body reviewed the protocol for compliance with research ethics standards, a narrower question entirely.

The Guidelines' own Example 2 demonstrates the result. A commercial AI start-up is found to satisfy factors i, ii, iii, iv and v, and is therefore conducting scientific research, on the strength of qualified staff, ethics review, and eventual publication, without the Guidelines requiring any separate showing that the start-up's process for selecting what to investigate, and how to report it, is insulated from its own commercial interest in the result. A company evaluating its own product for bias has an obvious interest in how that evaluation is designed, conducted, and characterised publicly. The six factors, applied cumulatively rather than with any factor treated as a necessary minimum, do not require that interest be examined at all.

Suggested addition: The Guidelines should state explicitly that the six factors are not a simple checklist in which strength on some factors offsets weakness on others, and that controllers relying on a marginal or contested case should be required to demonstrate, with particular rigour, factor iv (autonomy and independence) specifically. It is the factor most directly probative of whether the "research" framing is substantive rather than a means of accessing the GDPR's more permissive provisions for scientific research, and, not coincidentally, the factor most easily satisfied on paper (an ethics board, say) without being satisfied in substance.

2(b). OpenAI's trajectory shows that an entity satisfying the factors at one point in time may not satisfy them later, with no mechanism in the Guidelines to require reassessment

Issue:

The six factors are applied only once, at the point a controller first characterises its processing as scientific research. The Guidelines contain no mechanism requiring that characterisation to be reassessed as the controller's circumstances change. Entities evolve. A genuine research organisation can become a commercial product company without ever formally renouncing its original characterisation, and the Guidelines, as drafted, give such an entity no obligation to revisit whether that characterisation still holds.

The evolution of OpenAI is an instructive real-world illustration of why a point-in-time application of the six factors is insufficient, made unusually explicit by the organisation's own foundational documents. OpenAI's 2018 Charter⁴ committed the organisation to principles of "Broadly Distributed Benefits," "Long-Term Safety," "Technical Leadership," and "Cooperative Orientation," including a commitment to "use any influence we obtain over AGI's deployment to ensure it is used for the benefit of all," and an explicit undertaking that if a "value-aligned, safety-conscious project" came close to building AGI first, OpenAI would stop competing with it and assist it instead. This framing, emphasising independence from competitive commercial pressure and a stated willingness to subordinate the organisation's own position to a safety-driven mission, maps closely onto factor iv (autonomy and independence) and factor v (objectives of the research) as set out in the Guidelines, and was reinforced at the time by the organisation's name and its early practice of publishing research openly.

OpenAI's successor document, "Our Principles," published 26 April 2026,⁵ retains the stated mission that "AGI benefits all of humanity." The operative framing, however, has changed in ways relevant to the factors. The 2026 document substantially de-emphasises AGI-specific governance commitments relative to the 2018 Charter, and explicitly acknowledges that

⁴OpenAI. (9 April 2018). OpenAI Charter. *OpenAI*. <https://openai.com/charter>.

⁵Altman, S. (26 April 2026). Our Principles. *OpenAI*. <https://openai.com/index/our-principles/>.

OpenAI "is a much larger force in the world than it was a few years ago." Where the 2018 Charter included a concrete commitment to step aside for a better-positioned safety-conscious competitor, the 2026 document contains no equivalent undertaking. It states instead that OpenAI anticipates needing to "marshal substantial resources to fulfil our mission," language more at home in a fundraising deck than a research charter, while still invoking the language of broad societal benefit and democratisation that originally justified the organisation's claim to factors iv and v.

This is not presented as an allegation that OpenAI has violated the GDPR. The point is structural, and we do not wish to come across as accusatory. The contrast between the two documents, both publicly available on OpenAI's own website, illustrates concretely how an entity's own stated commitments on independence and mission can shift substantially over an eight-year period, while the underlying processing of personal data (training data collection, model fine-tuning, deployment) continues and scales, and while the organisation continues to describe its activity using research-oriented language. The change in authorship attribution between the two documents is itself notable: the 2018 Charter was presented as an institutional document reflecting input from OpenAI's research staff and leadership collectively, whereas "Our Principles" is published under the sole byline of CEO Sam Altman, even though it purports to state the organisation's governing commitments. A regulator applying the key-indicative factors to such an entity at two different points in time, using only the entity's own public self-description, could reasonably reach different conclusions about factor iv in 2018 and in 2026. The Guidelines provide no mechanism requiring that reassessment occur as part of a controller's ongoing accountability obligations.

At no single point in that trajectory is it obvious that the processing of personal data shifted from "scientific research" to ordinary commercial processing. The Guidelines, as drafted, give a controller no clear point at which the presumption of purpose compatibility (Section 3.1.1), the storage limitation exception (Section 3.2), or the legitimate interest balancing weight attached to scientific research (para. 61) would cease to apply to personal data collected or retained on a research basis. This is the structural problem with treating the six factors as a one-time threshold rather than a continuing condition: an entity can satisfy the factors at the point of initial data collection, then progressively repurpose both the data and the processing infrastructure toward commercial ends, without the Guidelines providing supervisory authorities or data subjects with any mechanism to reassess whether the original research justification still holds.

This connects to, but is not addressed by, the existing discussion of changes to processing operations in Section 5.5. That section's list of changes requiring renewed transparency to data subjects (paras. 108, 110) includes changes to the controller's identity (a laboratory merger, say) but does not include a change in the relationship between an organisation's research activities and its commercial product development, even though such a change

goes to the heart of factor iv, and may be more consequential to a data subject's reasonable expectations than the examples currently listed.

Suggested addition: The Guidelines should clarify that the key-indicative factors must be reassessed at the point of any material change in a controller's organisational structure, funding model, or the relationship between its research and commercial activities (conversion from a non-profit to a for-profit structure, say, or commercial deployment of models trained on data originally processed under a research framing), and that continued reliance on the scientific research provisions of the GDPR after such a change requires a fresh demonstration that the factors, in particular autonomy and independence, are still met. The dedicated example proposed in Comment 1 above, addressing a commercial foundation model developer, could usefully incorporate this point directly by showing how the controller's claim to the factors would need to be reassessed if its structure or commercial orientation changed after the data was first collected.

Respectfully submitted,
Meem Arafat Manab