EDPB public consultation on Guidelines 1/2026 on the processing of personal data for scientific research purposes

Provided by Fundación CARTIF on their behalf and related with their participation in LIFE LiveBetter project as responsable of its legal challenges (task 4.2.1).

Contribution from the perspective of an EU-funded project developing a digital building data space and AI-enabled services.

The Guidelines address a topic of direct relevance for EU-funded research and innovation actions, including LIFE and Horizon Europe projects, where personal data may be processed or further processed in order to validate innovative solutions, test AI-enabled services, improve data quality, develop interoperable infrastructures and produce evidence supporting public policy objectives.

From the perspective of LiveBetter, the Guidelines are particularly relevant because the project develops a Common European Digital Buildings Hub integrating a data space-compliant data-sharing mechanism and AI-based services for building renovation, dynamic EPCs, inspections, performance optimisation, renovation planning and financing, and energy flexibility. The project involves several partners across different Member States, business cases, pilot and validation activities, stakeholder engagement and potential processing or further use of personal or re-identifiable data, including dynamic building-related data and stakeholder data.

We consider that the Guidelines should preserve a clear, enabling approach to genuine scientific research while ensuring that safeguards, transparency, accountability and data subject rights are not weakened. The comments below focus on practical clarifications that would be helpful for EU-funded applied research projects using data spaces, AI services, smart buildings data and multi-partner governance models.

Scientific research in mission-oriented EU-funded projects

The Guidelines should expressly recognise that scientific research under the GDPR may include applied, mission-oriented and publicly funded research and innovation projects, provided that they follow recognised methodological and ethical standards and contribute to collective knowledge or societal benefit. In projects such as LiveBetter, research is not limited to laboratory or academic activities; it includes validation of data-driven services, development of common data models, performance evaluation, interoperability testing, evidence generation for policy and replication, and the assessment of innovative digital tools in real-world or pilot environments.

It would be useful for the Guidelines to include an example of an EU-funded LIFE or Horizon Europe project where research activities are carried out by a consortium combining research organisations, technology providers, public authorities, companies and end-user partners. Such an example should distinguish between:

- processing genuinely necessary for scientific research, validation, benchmarking, interoperability testing or impact assessment;
- ancillary processing supporting the research, such as participant management, workshops, surveys, stakeholder engagement or research panel management; and
- processing for operational deployment, commercial exploitation or service provision after the research phase, which may require a separate legal assessment and cannot automatically rely on the research framing.

This distinction is particularly important for projects that evolve from research and validation towards market uptake, exploitation or long-term platform operation.

Further processing, reuse of data and research infrastructures

We support the clarification that further processing for scientific research purposes may benefit from the GDPR presumption of compatibility, while still requiring a lawful basis and appropriate safeguards. This is essential for projects where data collected for one purpose may later be used to validate models, compare methodologies, produce aggregated indicators, train or evaluate AI services, or support future research in the same field.

However, for applied data-driven projects, additional guidance would be helpful on how to document the boundary between compatible further scientific research and other forms of reuse, such as product improvement, service optimisation, market validation or commercial exploitation. In LiveBetter-type projects, the same infrastructure may support research, service testing, policy reporting and future exploitation. A practical decision tree or example would help controllers determine when further processing remains within scientific research and when a new purpose assessment is needed.

We also suggest clarifying that research data infrastructures may include sectoral data spaces or research-oriented digital hubs, provided that access controls, governance rules, scientific purposes, ethical oversight and data minimisation safeguards are properly defined. For building-energy projects, this may include repositories or federated infrastructures that make harmonised building data available for future research, while avoiding unnecessary exposure of personal or re-identifiable data.

AI-enabled research and data-driven services

The Guidelines should include more explicit references to AI-enabled research (deeper mention to AI act). Many research projects use AI techniques such as machine learning for prediction, optimisation, recommendation, classification, anomaly detection, simulation or decision support. In LiveBetter, AI-enabled services may support energy performance forecasting, dynamic EPC-related outputs, digital inspections, renovation planning and financing, indoor environmental quality optimisation and energy flexibility.

Even where such AI systems are not classified as high-risk under the AI Act, the GDPR analysis for scientific research should consider transparency, fairness, explainability, human oversight, profiling, data quality and the possibility that outputs may influence decisions affecting individuals or households. The Guidelines could recommend that controllers document, at least at a high level, the AI use case, data sources, model purpose, expected outputs, human review, risk of bias or error, and whether the processing involves profiling or automated decision-making within the meaning of the GDPR.

This clarification would help ensure that Article 89 safeguards and DPIA screening are aligned with the realities of AI-enabled research without turning the research guidelines into a separate AI Act compliance document.