1. The six-factor test risks becoming a de facto legal definition of "scientific research"

The draft successfully tries to distinguish genuine scientific research from ordinary business analytics. However, it risks replacing the GDPR's intentionally broad concept of scientific research with a quasi-certification model based on publication, qualifications, independence, societal benefit, and external review. This could unintentionally exclude legitimate industrial R&D, AI research, citizen science, and emerging forms of scientific inquiry. The EDPB should clarify that the six factors are illustrative indicators and not cumulative substantive requirements for qualification as scientific research.

2. "Contribution to society's collective knowledge and wellbeing" is too normative

A regulator should not become the arbiter of whether research is sufficiently beneficial. Replace "contributing to society's general knowledge and wellbeing" with "generating, testing, validating, or applying knowledge using recognised scientific methods."

3. Researcher qualifications are under-specified and potentially exclusionary

The draft links scientific research to researchers possessing academic qualifications or scientific recognition. It is unclear whether formal credentials are required. Clarify that qualifications are merely one possible indicator and that demonstrated competence may suffice.

4. Broad consent remains unclear operationally

The draft is helpful in accepting broad consent for a "certain area of scientific research" and recognising combinations of broad and dynamic consent. Provide examples of acceptable research areas, overly broad areas and acceptable evolution of research objectives over time. This would substantially improve legal certainty.

5. The "reasonable expectations" test for broad consent is too vague

The guidelines require controllers to assess whether future projects remain within the reasonable expectations of participants. Reasonable expectations are notoriously difficult to assess consistently. Different DPAs may reach different conclusions about genetic research, AI model development, linkage with external datasets, or international collaboration. Create objective criteria like similarity of data categories, similarity of scientific field, similarity of participant population, or similarity of expected outputs.

6. Legitimate interest receives strong support but little operational guidance

One of the most consequential aspects is the statement that scientific research may constitute a legitimate interest and that significant weight can often be attributed to societal benefits in the balancing test. The draft does not explain when societal benefit

outweighs privacy intrusion and how public-benefit claims should be evidenced. Please provide sample legitimate-interest assessments.

7. The guidance on AI research is insufficient

The document repeatedly references AI and includes an example involving generative AI research. Yet it does not really address foundation-model training, web scraping, synthetic data generation, model evaluation datasets, or reuse of public internet content.

These are among the most contested research-compliance issues in Europe.

8. Transparency expectations may be unrealistic for longitudinal research

The draft stresses continuing transparency obligations and even suggests controllers should not knowingly delete contact details if future research processing is anticipated. This may conflict with data minimisation, pseudonymisation strategies, and long-term cohort studies. Researchers are often encouraged to separate identifiers from research datasets. Clarify that maintaining a trusted-contact mechanism or participant portal may satisfy transparency obligations without requiring continued retention of direct identifiers.

9. The storage-limitation section is helpful but still ambiguous

The guidelines recognise future research as a basis for retaining data and accept retention for future projects within a defined research area. The phrase "future scientific research activities should be reasonably foreseeable" is not operationalised. Provide examples of acceptable retention scenarios.

10. Anonymisation expectations may be unrealistically strong

The draft states that anonymised or alternatively pseudonymised data should be used whenever research objectives can be achieved that way and that directly identifying data should be processed only where strictly necessary and proportionate. Many modern research domains require longitudinal linkage - genomics, population health, AI fairness, or educational research. The language risks being interpreted by some DPAs as creating a presumption against identifiable processing. State explicitly that the necessity assessment should take account of scientific validity, reproducibility, quality assurance, linkage requirements and longitudinal follow-up.

11. The guidelines underplay Member State fragmentation

The document repeatedly notes that Member State law remains highly relevant. For cross-border research, this is arguably the biggest practical obstacle. Researchers need guidance on conflicting national research laws, multi-state consortia, or federated infrastructures. Add a dedicated section on cross-border research governance.