Comments on the draft Guidelines 02/2026 on Anonymisation

Submitted by Redactic (KS-Agents), Trieste, Italy

KS-Agents develops Redactic, a service that legal professionals and other organisations use to find personal data in PDF, Word, spreadsheet and image files, review each detection, and then either pseudonymise documents reversibly for work with AI services or redact them permanently before disclosure. We welcome the draft guidelines and their alignment with the judgment in C-413/23 P. Our comments come from the perspective of a provider that sees every day how organisations apply these concepts to documents rather than to datasets.

Summary of requests

1. Add a worked example applying the three criteria to a redacted text document, including its metadata.
2. Add an example on pseudonymised prompts sent to an AI service acting as processor.
3. Clarify how the location of the additional information weighs in the “means reasonably likely to be used” assessment.
4. Cross-reference Guidelines 01/2025 and list minimum documentation elements under paragraph 41.

1. Unstructured documents

The three criteria (No Record Isolation, No Linkage, No Inference) and most examples are framed around collections of records. A large share of everyday anonymisation concerns free text instead: contracts, pleadings, case files, emails, free-text survey answers. In such documents the information about one person is scattered across sentences, and the most common failure is not a residual name but a combination of role, date and place that singles someone out, or metadata that still carries the author’s and reviewers’ names.

Suggested addition: a worked example in which a controller redacts a text document for publication, showing how each of the three criteria is assessed on the body text, on indirect identifiers such as job titles and dates, and on the document metadata.

2. Processors and pseudonymised inputs to AI services

Paragraph 15 states that information which is personal data for a controller is also personal data for an entity processing it on the controller’s behalf. This is clear, but in practice it is easily misread: replacing names with placeholders is often taken to mean that an AI provider no longer receives personal data.

Suggested addition: an example on prompts sent to an AI service acting as processor, stating both that the pseudonymised prompt remains personal data for that service, and that pseudonymisation remains a valuable measure under Articles 25 and 32 GDPR because it limits what the processor can see and retain.

3. Where the additional information is held

Under the entity-relative approach, the location of the additional information matters. There is a meaningful difference between a mapping table stored by a service provider and one that never leaves the controller’s own environment, for example where real values are restored only on the user’s device. Paragraph 34 rightly notes that a contractual prohibition is not a prohibition by law and only complements technical measures.

Suggested addition: guidance on how such architectural choices, for instance keeping the mapping exclusively on the controller’s side, weigh in the assessment of the means reasonably likely to be used, so that controllers can compare solutions on technical rather than contractual grounds.

4. Consistency with Guidelines 01/2025 and documentation

Paragraph 22 of Guidelines 01/2025 states that pseudonymised data remains personal data even when the data and the additional information are not held by the same person, while the draft 02/2026, following C-413/23 P, adopts an entity-relative approach. Practitioners currently read both texts side by side.

Suggested addition: a cross-reference in 02/2026 explaining how the two documents relate, or a timely update of 01/2025. On paragraph 41, a short list of minimum documentation elements would make the accountability expectation easier to meet for small organisations: the techniques applied, the tests performed, the date, the reviewer, and an integrity reference for the output, such as a document hash.

Support for paragraph 40

We support paragraph 40. Describing reversibly pseudonymised output as “anonymous” is common in vendor marketing, and a clear statement from the Board helps users and providers alike.

Redactic (KS-Agents), Trieste, Italy