

SAMPLE. All figures are synthetic and for illustration only. They describe no real company or dataset.

KNOWLODE · ANONYMISATION REPORT

Dataset KL-SAMPLE-0001

§1 Summary

Source types	Support tickets; team chat
Sector / country	Accounting / Denmark
Period	2017–2025
Volume	48,210 tickets; 1,204,377 messages
Languages	Danish (82%), English (18%)
Decision	Approved for licensing (illustrative)

§2 Exclusions applied at source

- HR, payroll and health channels (7 channels)
- Direct messages and private channels
- Mailboxes of employees who opted out during the 30-day window (3)

§3 Methods

- Detection: Microsoft Presidio with a Danish transformer NER model, plus pattern rules for CPR, CVR, IBAN and phone numbers.
- Second pass: a language model reviewing residual entities, run on EU infrastructure.
- Replacement: consistent pseudonyms per entity, keeping thread and ticket relationships; pseudonym keys destroyed after packaging.
- Human review: 2% stratified sample (25,052 records) checked by two reviewers.

§4 Results

ENTITY TYPE	DETECTED	REPLACED	RESIDUAL IN HUMAN REVIEW
Person names	41,882	41,882	0
Email addresses	12,406	12,406	0
Phone numbers	3,117	3,117	0
CPR numbers (DK national ID)	214	214	0
IBAN / account numbers	1,032	1,032	0
Company names (customers, suppliers)	18,950	18,950	2 (fixed, re-tested)
Street addresses	2,644	2,644	0

§5 Re-identification tests

- **Singling out:** no record could be linked to a single real person or company using quasi-identifiers (role, date, city) in a 500-record adversarial test.
- **Linkability:** sampled records matched against public web sources and company registers: 0 confirmed links.
- **Inference:** rare events (e.g. named litigation, unique deals) generalised or removed: 37 records.

§6 Residual risk

In this example, residual risk is assessed as low from the point of view of a well-resourced recipient. The licence forbids re-identification attempts and onward publication.

§7 Deletion

Raw export received on day 0; deleted on day 25. Pseudonym keys deleted on day 25. Confirmation sent to the seller.