

IFG Decision: Transfer from KMS, Customer: [REDACTED] (A31-210-26)

English translation of the original German document. Legal terms are rendered functionally; the German original is authoritative.

From: Informationsfreiheitsgesetz (A31) <Informationsfreiheitsgesetz@destatis.de>

To: [REDACTED]

Cc: Informationsfreiheitsgesetz (A31) <Informationsfreiheitsgesetz@destatis.de>

Date: Wednesday, 26 August 2026, 13:33:40

Dear Mr [REDACTED],

By message dated 17 August 2026 (A31-210-26) you submitted a request under Section 1 of the Freedom of Information Act (Informationsfreiheitsgesetz, IFG) to the Federal Statistical Office.

With regard to the information requested, following consultation with the responsible organisational unit, we inform you as follows:

1. Which software or programme libraries does the Federal Statistical Office use for the record-linkage procedure? Is it an in-house development, a commercial product, or open-source components – such as Splink or a functionally equivalent probabilistic method?

Within the Population Methods Test (Methodentest Bevölkerung), the Python library Splink (open source) was used for the record-linkage procedure.

2. Does the procedure operate probabilistically or deterministically? Are decision thresholds set for match, non-match and manual review, and roughly what magnitude are they?

In the Population Methods Test, both probabilistic and deterministic search spaces were used. Since this was a methodological study, decision thresholds were regularly adjusted and changed. These investigations form the basis for further conceptual work on the 2031 Census. Two limit values were used per search space. If the matching score in the probabilistic record-linkage procedure (RLV) fell below the lower limit, the record pair was defined as a non-match. If a record pair lay above the upper limit, it was defined as a match. For each delivery source there are several different search spaces and, accordingly, thresholds that are continually adjusted. Since this was a trial of the methods in the population domain, it is not meaningful to state specific thresholds here. If the matching score fell between the lower and upper limits, the record pair could be passed to manual review. For the purposes of the methodological trial, selected cases were reviewed manually. The design of the record-linkage procedure for the 2031 Census is the subject of conceptual work on which no statement can be made at this time.

3. In how many cases did the comparison between the reports from the comparison registers and the population-register data yield no result, such that a residence analysis (Wohnsitzanalyse) was triggered? I request the absolute number of cases and the share of the total cases processed.

The number of unlinked delivery records varied greatly by delivery source. Since links exist at the level of register entries, whereas in the residence analysis it is persons who are surveyed, no direct inference is possible between linkage and residence analysis. In the Population Methods Test, the scope of the residence analysis was limited by law to 100,000 cases.

4. Are false matches and missed matches quantified? If so: what values were determined in the Methods Test? If not: for what reasons is such quantification dispensed with?

False links can only be determined with the aid of a reference dataset that represents the underlying “truth”. They therefore cannot readily be quantified. Unlinked records were quantified and differ by delivery source.

5. Are there evaluation or quality reports on the record-linkage procedure, and is their publication planned?

The reports on the Population Methods Test are currently being prepared. Initial publications on this are planned for 2027.

6. Was a data protection impact assessment under Article 35 GDPR carried out for these processing operations, and was the Federal Commissioner for Data Protection and Freedom of Information involved?

For the Population Methods Test, a data protection impact assessment under Article 35 GDPR was carried out in accordance with the guidelines of the Federal Commissioner for Data Protection and Freedom of Information and the data protection concept of the Federal Statistical Office.

We are happy to answer any further questions.

Notice of legal remedies

You may lodge an objection (Widerspruch) against this decision within one month of its notification. The objection must be lodged with the Federal Statistical Office, Wiesbaden.

The following options are available to you:

In writing or for the record: The objection may be lodged in writing or for the record. The address is: Gustav-Stresemann-Ring 11, 65189 Wiesbaden.

We thank you for your interest in official statistics and hope to have been of assistance.

Yours sincerely,
On behalf of

Josin Hauck
Referentin (Desk Officer)
Legal Department (Justizariat)
Federal Statistical Office (Destatis)
Tel.: +49 (0)611 75 2565 | Mobile: +49 (0)151 23569817
josin.hauck@destatis.de | www.destatis.de
Data protection notice: www.destatis.de/datenschutz

Original message: Acknowledgement of receipt

From: Informationsfreiheitsgesetz (A31) <Informationsfreiheitsgesetz@destatis.de>

Sent: Monday, 17 August 2026, 12:27

To: [REDACTED]

Cc: Informationsfreiheitsgesetz (A31) <Informationsfreiheitsgesetz@destatis.de>

Subject: RE: Transfer from KMS, Customer: [REDACTED] – Acknowledgement of receipt

Dear Mr [REDACTED],

We confirm receipt of your request under the Freedom of Information Act dated 17 August 2026.

Your request is being handled here under file reference A31_210-26. Please always quote this file reference in any further correspondence.

Your request is currently being processed. We therefore ask for your patience; we will get back to you without further prompting.

Yours sincerely,

On behalf of

Ms Thi Thu Ha Katzenmeier
Senior Case Officer (Hauptsachbearbeiterin)
Legal Department (Justizariat)
Federal Statistical Office (Destatis)
Tel.: +49 611 75 2369 | Mobile: +49 175 3444300
thithuha.katzenmeier@destatis.de | www.destatis.de
Data protection notice: www.destatis.de/datenschutz

Original request (17 August 2026)

Dear Sir or Madam,

I am an independent researcher working on a cross-border investigation into the use of probabilistic methods for linking register data (record linkage) in official statistics. The research currently covers the United Kingdom, Australia, New Zealand, Chile and Germany; the results will be published.

The reason for my request is that, to date, none of the statistical authorities examined has published quantified error rates for the linkage procedures in use. With regard to the register-based census, the Methods Test under the Register Census Trial Act (Registerzensuserprobungsgesetz) and the combination model, I would therefore like to put the following questions to you:

1. Which software or programme libraries does the Federal Statistical Office use for the record-linkage procedure? Is it an in-house development, a commercial product, or open-source components – such as Splink or a functionally equivalent probabilistic method?
2. Does the procedure operate probabilistically or deterministically? Are decision thresholds set for match, non-match and manual review, and roughly what magnitude are they?
3. In how many cases did the comparison between the reports from the comparison registers and the population-register data yield no result, such that a residence analysis was triggered? I request the absolute number of cases and the share of the total cases processed.
4. Are false matches and missed matches quantified? If so: what values were determined in the Methods Test? If not: for what reasons is such quantification dispensed with?
5. Are there evaluation or quality reports on the record-linkage procedure, and is their publication planned?
6. Was a data protection impact assessment under Article 35 GDPR carried out for these processing operations, and was the Federal Commissioner for Data Protection and Freedom of Information involved?

I would appreciate a reply by [DATE]. Should no statement be received by then, I will note this accordingly in the publication. Naturally, I offer you the opportunity to comment on the research findings in advance.

I am available for any questions at any time, including by telephone.

Yours sincerely,

[REDACTED]

Customer record (KMS): Email: [REDACTED] · Salutation: Mr · First name: [REDACTED] · Surname: [REDACTED] · Company: [REDACTED] [REDACTED].net · Country: Germany · User category: Media, communications and advertising agencies, graphic design · Channel: Email · Contact on: 17 August 2026 · Stored in customer management system under 575448 / 832397