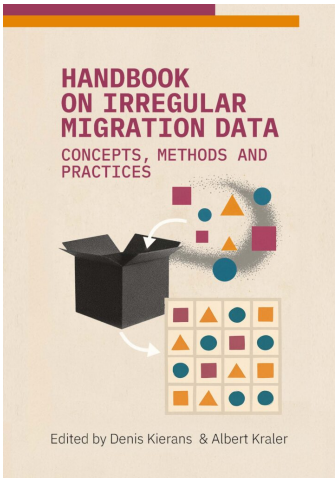

Textbox taken from D.Kierans & A.Kraler (eds), Handbook on Irregular Migration Data. Concepts, Methods and Practices. Krems: University of Krems Press

ISBN: 978-3-903470-24-8

Box 3.3 Linkage of administrative data in a data protection sensitive way – The case of Austria

Albert Kraler

To cite: Kraler, A. (2025). Linkage of administrative data in a data protection sensitive way – The case of Austria. In D. Kierans and A. Kraler (eds), *Handbook on Irregular Migration Data. Concepts, Methods and Practices.* Krems: University of Krems Press. https://doi.org/10.48341/g31s-vq79-box3.3

Keywords : Pseudonymisation, record linkage, Austrian Micro-Data Centre, Austria

their methods. Transparency here is substantive: it requires explaining potential sources of error, methodological assumptions, and the ways findings can and cannot be interpreted. For algorithmic methods, transparency can help other experts (and it is important to acknowledge this facet) to contest the assumptions and biases embedded into computational analysis. By doing so, policymaker and researchers help ensure that data supports informed, balanced policy decisions rather than fuelling sensationalism or punitive responses.

Considerations for data linkage and anonymisation strategies

Data protection law, especially the GDPR, imposes clear limits on how personal data may be collected, used, and shared. While these rules are crucial for protecting individual rights, they can also pose practical challenges for research, particularly in linking datasets across sources or countries.

It is necessary to respond to these challenges through careful anonymisation strategies. Pseudonymisation of individuals' identities is a standard practice, with participants given choices about the level of disclosure they are comfortable with. Researchers can use coded protocols for interviews, workshops, and surveys to minimise identifiability. Anonymisation should not be treated as a one-off exercise but as an ongoing obligation to protect participants' rights as data is processed, analysed, and shared. This also involves putting in place technical safeguards, for example: access controls that limit who can view or process data; and secure environments supported by encryption (see for an innovative example of pseudonymisation by design, Box 3.3).

Box 3.3: Linkage of administrative data in a data protection sensitive way – The case of Austria
Albert Kraler

On the national level, a wide range of statistical indicators on irregular migration are available from different administrative databases, including those on migration enforcement (apprehensions, return orders, rejections at the border, migrant smuggling, etc.), asylum databases, and residence permit databases. Despite some inherent limitations associated to their administrative purpose, the anchoring of measurement concepts in operational and legal categories and their specific scope linked to domain specific regulatory frameworks, administrative databases provide a rich source for scientific analysis. This is particularly true when they contain historical data and allow examining migrants' trajectories (chapter 7) or when they allow linkage of different databases (record linkage). In both cases, questions about data protection arise. For example, in compliance with the privacy regulations databases generally foresee a certain timeframe after which personal data needs to be deleted, if no longer necessary for the particular administrative purpose they are meant to serve. Sometimes, specific events will lead to the deletion of records from registers. For example acquisition of citizenship will result in the deletion of that person's records from residence permit registers). Similarly, record linkage can be restricted by law, as is the question of who has access to different types of data.

The case of Austria is a good example of database linkage and the preservation of historical records are possible in a data protection compliant way. In Austria, the pseudonymisation of register data for statistical purposes is achieved through the use of (encrypted) sector specific personal identifiers (verschlüsselte Bereichsspezifische Personenkennzahl Amtliche Statistik – bPK-AS). The bPK-AS is generated by the Stammzahlenregisterbehörde (Central Register Authority). It is a cryptographically derived identifier derived from the personal identifier used in a specific domain (for example social security, or the population register code) and a code for the domain.[3] It is unique to each individual

3 The principle of encryption used for the generation of the sector specific identifiers is described (in German) here : https://www.bundeskanzleramt.gv.at/agenda/digitalisierung/stammzahlenregisterbehoerde/bereichsspezifische-personenkennzeichen/beschreibung.html.
The encryption procedure is based on Central Register Authority Ordinance (Stammzahlenregisterbehördenverordnung) 2022, see https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20011934.

and serves as a key to match data from different registers with each other. Crucially, the bPK-AS is not reversible, meaning it cannot be traced back to the original personal identification number (Statistics Austria 2024). Statistics Austria uses these anonymised personal identifiers to link data from various sources – such as social insurance records, employment data, and education registers – through deterministic linkage and without revealing personal identities. While Statistics Austria gets updates from administrative databases in real time, it uses anonymized statistical mirror databases for statistical purposes (Fuchs et al. 2024). All register data is stored in a historicised way, allowing longitudinal analysis.

Since 2022, all statistical databases based on data collected by Statistics Austria itself (through surveys and other statistical reporting systems) as well as a wide range of administrative databases from different public bodies are assembled in the "Austria Micro Data Centre" (AMDC).[4] By mid-2026, all public administrative database – with the exception of security related databases – should be made available by the AMDC. In addition, researchers can link their own datasets to the AMDC by obtaining a sector specific identifier from the Central Register Authority for their own dataset, which in turn enables Statistics Austria to include this dataset in the AMDC, making it linkable to all datasets contained in the AMDC. A precondition for including a dataset in the AMDC is that researchers collect personal information (notably name, date of birth, place of residence) to enable pseudonymisation by the Central Register Authority. The AMDC is open for researchers in accredited institutions, which need to meet a number of criteria for accreditation (such as scientific purpose of the organisation, research quality, independence).

While immigration and migration enforcement related databases are not (yet) linked to the AMDC and therefore cannot be used to analyse legal status trajectories, the design of the system nevertheless can serve as a model for balancing data utility and privacy protection.

References

Fuchs, R., Göllner, T., Hartmann, S. & Thomas, T. (2024). "Fostering Excellent Research by the Austrian Micro Data Center (AMDC)" Jahrbücher für Nationalökonomie und Statistik, 244(4), 433-445. https://doi.org/10.1515/jbnst-2023-0043

Statistik Austria (2024): Standard-Dokumentation Metainformationen (Definitionen, Erläuterungen, Methoden, Qualität) zu den Registerbasierten Erwerbsverläufen. https://www.statistik.at/fileadmin/shared/QM/Standarddokumentationen/B_1/std_b_erv.pdf

Importantly, 'special categories' of personal data, such as 'race', ethnic origin or political opinions, carry heightened risks. The MIrreM project applies the 'data minimisation principle' by deliberately limiting data collection to what is strictly necessary, while ensuring individuals are fully informed of their rights and protections.

When applied carefully, these practices allow data to be used constructively: for example, to understand migration patterns, design inclusive services and improve resource allocation without compromising individual privacy.

4 https://www.statistik.at/en/services/tools/services/center-for-science/austrian-micro-data-center-amdc