Possible ways to claim that data are anonymous

(a taxonomy)

Disclaimer: This taxonomy solely attempts to make the **structure of reasoning** explicit by choosing possible complete **sets of necessary claims** (table rows); an assessment of the strength or validity of **arguments that support such claims is out of scope**. Without supporting argumentation, this taxonomy is insufficient to claim that data actually is anonymous. Lower in the table solely means structurally stronger argumentation, and depending on the strength and validity of supporting arguments, not necessarily "more anonymous".

Objective of Reasoning:

Notion of Anonymity: ☐ differential ☐ absolute

Time Horizon: _____ years

Structure of Reasoning: **Select type of argumentation** (table row); multiple checks indicate additional lines of defense, should the stronger ones fail. **Document facts** and **arguments** in support of claims.

For natively or reconstructed (i.e. the yellow part is an additional line of defense) individual-level data, **linkage is not possible because**:

1: The **data does not contain direct identifiers** and has been **protected against linkage** (syntactic methods) as follows:	**(i)** *Facts:* protection of **quasi-identifiers** *[selection criteria, list QIs, type of protection, strength (K-min, K-avg)]*	**(ii)** *Facts:* protection of **other attributes:** *[types of protections, resulting max. SUDA[1] score, ..]*	1.1: *Claim:* Linkage not possible based on **assumptions about potential attackers**	1.1.1: *Claim:* Attackers **lack motivation** (cost benefit)	☐
				1.1.2: *Claim:* Attackers **lack capability** (resources, skill)	☐
			1.2: *Claim:* Linkage impossible based on **assumptions about suitable auxiliary information** (with matching link targets) Consider inference! => fewer possible anchors (e.g., exclude spontan. recognition)	1.2.1: *Claim:* Suitable auxiliary information exists but is **not accessible to potential attackers**	☐
				1.2.2: *Claim:* Suitable auxiliary information **does not exist** (monopoly of observation, variation of values with each observation)	☐
			1.3: *Claim:* Linkage impossible since **data** provides **no unambiguous link anchors** (any unique combination of attributes) (with arbitrary auxiliary information) Consider inference!	1.3.1: *Claim:* **Modification of** potential **anchors** renders **matches uncertain** and **deniable** (noise, swapping, ..)	☐
				1.3.2: *Claim:* **No unique records** contained **in data** (all attributes treated as quasi-identifiers: classes of equal values or clusters of close values) Consider inference!	☐
			..join blue part here..		

stronger

For aggregate information, **reconstruction is not possible because:**

		..join yellow part here..		
2: The data is aggregated, and thus direct or indirect identification are only possible after successful (possibly partial) reconstruction *[type: statistics, AI-model, synthetic data, … ,minimal cell size or similar measure of aggregation level]*	**2.1:** *Fact:* **Data without mathematically guaranteed reconstruction protection** (e.g., • statistics without additional protection, • (empirical) rule-based disclosure control) *Reconstruction protection:* facts*: [type, properties]* *E.g. (none, …)*	2.1.1: *Claim*: Reconstruction is assumed to be impossible based on **assumptions about potential attackers**:	2.1.1.1: *Claim:* Attackers **lack motivation** (cost/benefit)	☐
			2.1.1.2: *Claim:* Attackers **lack capability** (skill, resources, ..)	☐
		2.1.2: *Claim*: Reconstruction is assumed to be impossible based on **assumptions about additional disclosures** and ..	2.1.2.1: *Claim*: **..addl. disclosures exist but are not accessible** by potential attackers	☐
			2.1.2.2: *Claim*: ..Significant additional disclosures **don't exist**	☐
		2.1.3: *Claim*: Reconstruction is assumed to be impossible based on current **state of the art** and .. **consider:** reconstruction protection, inferences	2.1.3.1: *Claim*: Known attacks *[enum]* fail **based on assumptions** about state of the art	☐
			2.1.3.2: *Claim*: Known attacks *[enum]* fail as **verified with own data**	☐
	2.2: *Fact:* **Data with mathematically guaranteed reconstruction protection** **Guarantee:** facts*: [type, strength] (e.g. eps-DP, eps)*	2.2.1: *Fact*: The privacy budget is managed **for own disclosures,** and..	2.2.1.1: Claim: ..no significant number of external disclosures are accessible to attackers	☐
			2.2.1.2: *Claim*: ..no significant number of external disclosures exists	☐
		2.2.2: *Claim*: The privacy budget is managed for both, **own and external disclosures**		☐

stronger

version 0.9

---

[1] Elliot, M. J., Manning, A. M., & Ford, R. W. (2002). A Computational Algorithm for Handling the Special Uniques Problem. International Journal of Uncertainty, Fuzziness and Knowledge Based System , 10 (5), 493-509. and also https://pypi.org/project/suda/.