Anonymizace je cesta ke zdravotnickým datům

Součást projektu Kaleidoskop zdravotnických dat

FBMI ČVUT
Aleš Tichopád

Data jsou aktivem informační společnosti 21. století

Je společensky žádoucí, aby zdravotnická data byla v rozumné míře přístupná odborné i širší veřejnosti pro sekundární využití

Princip rovnováhy mezi společenským benefitem a rizikem

Přístup ke zdravotním datům v EU
Právní základ v čl. 9 odst. 2 GDPR

Právní základ pro zpracování zdravotních dat pro výzkum	Výzkumní pracovníci veřejného sektoru	Výzkumní pracovníci třetích stran mimo veřejný sektor
Zpracování dat na základě souhlasu	AT, BE, DE, EE, FI, LV, IE, MT, NL	AT, BE, DE, DK, EE, FI, LV, IE, MT, NL
Zpracování dat na základě veřejného, nebo pro výzkumné účely	AT, BE, DK, DE, EE, FI, FR, HR, HU, LV, LU, MT, IE, IT, [UK] 15	AT, BE, DK, DE, EE, FI, FR, IE, IT, LV, LU, HU, MT, [UK] 14
Žádná konkrétní legislativa	BG, CY, **CZ**, ES, EL, LT, PL, PT, RO, SI, SK, SE 12	BG, CY, **CZ**, ES, EL, HR, LT, PL, PT, RO, SI, SK, SE 13

Hlavní principy podle GDPR

aneb není to beznadějné

⚠ **Zdravotnická data jsou osobní údaje (mnohdy citlivé)**

(4) Zpracování OÚ by mělo sloužit lidem. Právo na ochranu osobních údajů **není právem absolutním** …

(26) Zásady ochrany údajů by se měly uplatňovat na všechny informace týkající se **identifikované** nebo **identifikovatelné** fyzické osoby

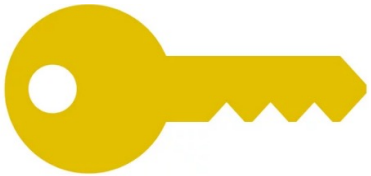
NICMÉNĚ!

(26) Anonymizovaná zdrav. data nejsou osobní údaje

„Zásady ochrany osobních údajů by se proto neměly vztahovat na anonymní informace, totiž informace, které se netýkají identifikované či identifikovatelné fyzické osoby, ani na osobní údaje anonymizované tak, že subjekt údajů není nebo již přestal být identifikovatelným. Toto nařízení se tedy netýká zpracování těchto anonymních informací, včetně zpracování pro statistické nebo výzkumné účely.“

Pozor na pseudonymizovaná data

- Pseudonymizace OÚ je proces nahrazení identifikátorů údajů **pseudonymem.**
- Pseudonym **umožnuje zpětně identifikací osoby za pomoci klíče**.
- Pseudonymizované OÚ **mohou byt spojené s konkrétní osobou** za pomoci dodatečné a odděleně uložené informace.
- **Pseudonymizované údaje jsou i nadále OÚ.**
- Psedonymizace na rozdíl od anonymizace poskytuje jen **částečnou ochranu** před ztotožněním.



Anonymizované osobní údaje

Anonymizace OÚ znamená takové zpracování OÚ, díky němuž dojde k **nevratné de-identifikaci** osoby.

U de-identifikovaných údajů dojde k nevratnému:
- odstranění přímých identifikačních údajů.
- Modifikaci nebo odstranění kvazi-identifikátorů

De-identifikace je nevratná => Anonymizace (jinak pseudonymizace)

U anonymizovaných OÚ se předpokládá, že je správce OÚ nebo jiná osoba nedokáží rozkrýt ani s využitím **prostředků u nichž rozumně předpokládáme jejich dostupnost**.

Anonymizované údaje **nejsou OÚ** a nevztahuje se na ně legislativa na ochranu OÚ.

Anonymizace je **předmětem intenzivního výzkumu a stále se rozvíjí**

Nepřímá identifikace osoby

aneb Jarmila je asi těhotná

- Osoba nemusí být uvedena jmenovitě, aby byla identifikovatelná
- Kvazi-identifikátory jsou údaje o osobě a jejich kombinace dokáže osobu (nebo malou skupinu osob) vyčlenit
  - Přímé identifikatory: Jméno a příjmení, rodné číslo (osoba je identifikovaná)
  - Nepřímé (kvazi-identifikátory): osobu lze stále identifikovat na základě nepřímých identifikátorů popř. jejich kombinací

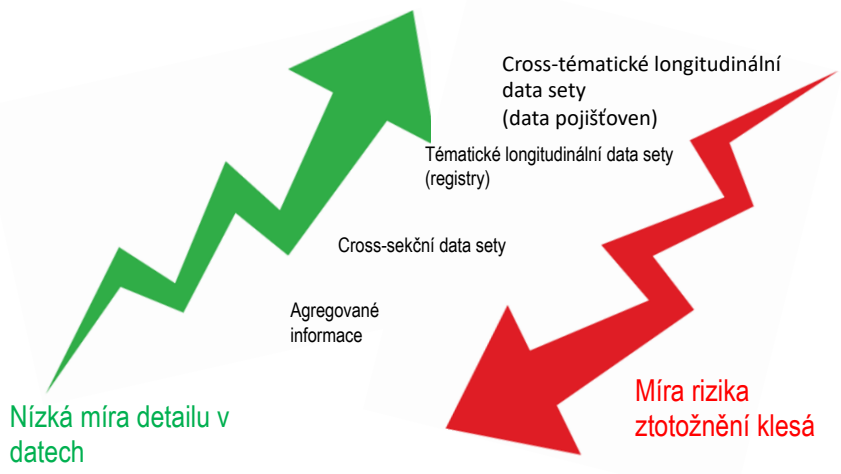
Příklad nepřímé identifikace: Datumy opakovaných vyšetření u gynekologa + pracoviště + datum narození + výkon: odběr plodové vody

Jarmila vedle z účtárny je asi těhotná!!!

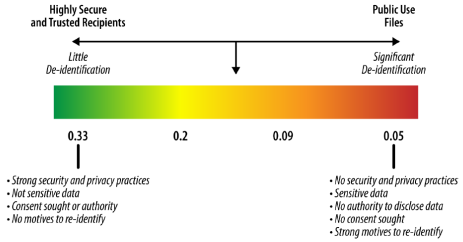
⚠ **Pozor na longitudinální data s přesnou časovou stopou**

TEST IDENTIFIKACE: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf

Riziko je škálovatelné



Riziko a jeho minimalizace podle žadatele



Mechanismy udělení přístupu ke zdravotním datům v EU

1. Přístup udělen na základě schválení etickou komisí (EK) nebo agenturou pro ochranu údajů (DPA)
2. Držitel dat uděluje přístup přímo, bez udělení souhlasu EK nebo DPA
3. Existence centralizovaného správního orgánu pro udělování přístupu k datům

Instituce	Mechanismus přístupu
Nizozemsko	Souhlas etické komise
Francie	Souhlas etické komise
Španělsko/Aragonie	Souhlas etické komise
Finsko	Přístup skrze centralizovaný správní orgán, v některých případech však souhlas EK
Dánsko	Držitel dat uděluje přístup, v některých případech nutný souhlas EK
Německo	Držitel dat uděluje přístup
Litva	Souhlas etické komise
Malta	Přístup skrze centralizovaný správní orgán, patient-level data – souhlas EK, agregovaná, čí cross-section data bez souhlasu EK

Zdroj: health.ec.europa.eu, tehdas.eu

Koncoví uživatelé způsobilí pro přístup ke zdravotním datům

Vybrané instituce

Instituce	Koncoví uživatelé							
	Národní výzkumníci	Evropští výzkumníci	Mezinárodní výzkumníci	Tvůrci národní politiky	Tvůrci evropské politiky	Tvůrci mezinárodní politiky	Jednotliví občané	Ostatní aktéři (soukromý sektor)
Nizozemsko CBS	✓	✓	✓	✓				
Finsko Findata	✓	✓	✓	✓			✓	✓
Francie HDH	✓	✓	✓	✓	✓	✓	✓	✓
Nizozemsko Health-RI	✓	✓	✓					✓
Španělsko IACS/BIGAN	✓	✓	✓	✓	✓	✓	✓	✓
Dánsko Sundhedsdata	✓	✓	✓*	✓	✓*	✓*	✓**	✓

Zdroj: health.ec.europa.eu, sundhedsdatastyrelsen.dk

Case study: Dánsko

"Dánsko má jedny z nejkomplexnějších zdravotních údajů na světě, které pokrývají mnoho oblastí zdraví po dlouhou dobu. Všechny údaje jsou propojitelné díky dánskému národnímu jedinečnému občanskému registračnímu systému. To umožňuje používat dánská zdravotní data k vytváření nových znalostí a poznatků."

Zdroj: sundhedsdatastyrelsen.dk

Typy dat dostupných pro výzkum
Pseudonymizace/anonymizace

- **Pseudonymizovaná data**
  - Data na úrovni osoby
  - Veřejné instituce shromažďují data občanů pomocí identifikačního čísla pacienta (PIN). Možnost propojení dat se Statistics Denmark a také s daty z různých sektorů.
- **Anonymizovaná/agregovaná data**
  - Pokud není potřeba pracovat s individual-level daty, lze požádat o extrakci statistik.
  - Tyto statistické výpisy jsou zcela anonymní.
  - Při práci s daty na úrovni osoby se využívá procesu anonymizace jsou **maskována pole s méně než pěti osobami/pozorováními**.
  - Pokud **Research Services** usoudí, že v tabulkách jsou jinak přítomny další osobní identifikátory, budou také maskovány.

Dostatečná anonymizace

aneb jak nevylít vanku s dítětem

- De-indentifikace nemusí být absolutní.
- Posuzujme žadatele
  - Akademická (systematicky vzdělávací) pracoviště
  - Kvalifikované fyzické a právnické osoby
  - Právní kanceláře
  - Marketingové agentury
  - Nekvalifikovaní žadatelé
  - Nejrizikovější jsou open data (snadnější vyčlenění nenáhodně vybrané osoby)
  - Co u žadatele posuzovat:
    - Existuje interní policy pro práci s daty?
    - Existuje předchozí zkušenost?
    - Hromadí se u něj data – pozor na re-identifikaci!
    - Existuje zjevný zájem?
    - Případy zneužití?

CESTA K DATŮM

Stanovisko a postupy

Plánované otištění v Časopise lékařů českých

7 hlavních principů cesty k datům

- ❖ Data jsou důležitý celospolečenským aktivem
- ❖ Ochrana soukromí by měla být posuzována v poměru k celospolečenským benefitům
- ❖ Riziko zosobnění je aktivně škálovatelné
- ❖ Deidentifikace poskytuje ochranu před ztotožněním u dat na úrovni fyzické osoby
- ❖ Aktivní prevence na straně žadatele je nezbytná
- ❖ Žadatel i poskytovatel musí v průběhu žádosti a poskytováni komunikovat a spolupracovat.
- ❖ Míry anonymity a rozsahu poskytovaných dat by měla být definovaná účelem a společenským přínosem

Děkuji za pozornost

ales.tichopad@fbmi.cvut.cz