

---

# Privacy-Preserving Record Linkage: Past, Present and Yet-to-Come

EDBT 2026

Tampere, Finland

L. Stetsikas, D. Karapiperis, G. Papadakis,  
and M. Koubarakis

---



Funded by  
the European Union

# anonlink

**Creators:** CSIRO's Data61

**Organisation:** CSIRO's Data61

**Git Repo:** <https://github.com/dat61/anonlink>

**Language:** python

**Installation:** pip

**Type:** Library

**Input-Files:** .csv

Pipeline Feature	Status	Supported Methods / Notes
Type of Input	●	ANY
Preprocessing	●	Standard
Encoding	●	Cryptographic Long Keys
Blocking	●	LSH (Hamming)
Matching	●	Dice, Simple Matching Coefficient
Clustering	●	
Multiparty	●	
Operational Status	●	
Documentation	●	

# AMPPERE

**Creators:** Yao, Yixiang and Ghai,  
Tanmay and Ravi, Srivatsan and  
Szekely, Pedro

**Organisation:** University of Southern  
California Information Sciences Institute,  
California

**Git Repo:** <https://github.com/usc-isi-i2/amppere>

**Language:** C++

**Installation:** Source Files

**Type:** CLI

**Input-Files:** .csv

Pipeline Feature	Status	Supported Methods / Notes
Type of Input	●	ANY
Preprocessing	●	febrl.py
Encoding	●	Homomorphic Encryption
Cryptographic Schemes	●	BGV, CKKS
Blocking	●	LSH
Matching	●	Vector Rotation, Vector Extraction
Clustering	●	
Evaluation	●	
Multiparty	●	3-Party Support
Operational Status	●	Done some tweaking
Documentation	●	

# Linkja

**Creators:** William Trick, Abel Kho, Francisco Angulo, Luke Rasmussen, Kruti Doshi, Larry Lemmon

**Organisation:** Linkja

**Git Repo:** All Java programs to run linkja can be found here: <https://linkja.github.io/>

**Language:** Java

**Installation:** Source Files

**Type:** CLI

**Input-Files:** .csv

Pipeline Feature	Status	Supported Methods / Notes
Type of Input	●	Only PHI data
Preprocessing	●	
Encoding	●	Bloom Filters
Blocking	●	
Matching	●	Exact Match
Clustering	●	
Evaluation	●	
Multiparty	●	
Operational Status	●	
Documentation	●	

# PPRL Toolkit

**Creators:** Henry Wilde, Mat Weldon

**Organisation:** Data Science  
Campus, UK

**Git Repo:** [https://github.com/datasciencecampus/pprl\\_toolkit](https://github.com/datasciencecampus/pprl_toolkit)

**Language:** python

**Installation:** Source Files

**Type:** Library

**Input-Files:** .csv

Pipeline Feature	Status	Supported Methods / Notes
Type of Input	●	ANY
Preprocessing	●	Standard, Feature-based
Encoding	●	Bloom Filters
Blocking	●	
Matching	●	Soft Cosine
Clustering	●	
Evaluation	●	
Multiparty	●	
Operational Status	●	Archived in Github
Documentation	●	

# PRIMAT

**Creators:** Florens Rohde, Victor Christen, Ziad Sehili, Thomas Hoppe, Duc Dung Dao, Marcel Gladbach

**Organisation:** University Of Leipzig, Germany

**Git Repo:** <https://git.informatik.uni-leipzig.de/dbs/pprl/primat>

**Language:** Java

**Installation:** Maven

**Type:** Library

**Input-Files:** .csv

Pipeline Feature	Status	Supported Methods / Notes
Type of Input	●	ANY, seperator always ,
Preprocessing	●	Standard
Encoding	●	Bloom Filters
Blocking	●	Standard, LSH
Matching	●	Dice, Hamming, Jaccard
Clustering	●	Removes Conflicting matches by keeping the most similar pairs
Evaluation	●	Only for matching
Multiparty	●	
Operational Status	●	
Documentation	●	

# privJedAI

**Creators:** L. Stetsikas, D. Karapiperis, G. Papadakis, M. Koubarakis

**Organisation:** National and Kapodistrian University Of Athens, Greece

**Git Repo:** <https://github.com/AI-team-UoA/privJedAI>

**Language:** python

**Installation:** pip

**Type:** Library

**Input-Files:** .csv

Pipeline Feature	Status	Supported Methods / Notes
Type of Input	●	ANY
Preprocessing	●	Standard
Encoding	●	Bloom Filters
Blocking	●	LSH, ANNS, Metablocking
Matching	●	Dice, Hamming Distance, Jaccard
Clustering	●	Unique Mapping, Connected Components, Kiraly
Evaluation	●	Produce metrics for each method
Multiparty	●	(Not Supported)
Operational Status	●	
Documentation	●	privjedai.rfd.io

# Summary

Feature	Anonlink	AMPPERE	Linkja	PPRL Toolkit	PRIMAT	PrivJedAI
<b>Language</b>	Python	C++	Java	Python (C++ Core)	Java	Python
<b>Program Type</b>	Open-Source Library	Implementation For Research Purposes	CLI & GUI	Open-Source Library	Open-Source Library	Open-Source Library
<b>Encoding Methods</b>	Cryptographic Long Keys ICLK)	Homomorphic Encryption (BGV, CKKS)	SHA-256, Bloom Filters	Bloom Filters, Specified Hashing Methods for Different types of Features	Bloom Filters (CLK, RBF), Hashing with Multiple Functions	SHA-256, Bloom Filters (CLK, RBF), Hashing at Different Levels
<b>Blocking Methods</b>	Hamming LSH, Radius Blocking, Explicit Blocking	LSH			Hamming LSH, Standard Blocking	Hamming-LSH, ANNS, Meta-blocking Methods
<b>Matching Similarities</b>	Dice Coefficient, Hamming Similarity, Simple Matching Coefficient	Vector Rotation, Vector Extraction	Exact Matching	Soft Cosine Measure	Hamming Distance, Jaccard Coefficient	Jaccard, Dice, Hamming Distance
<b>Additional Methods</b>	Multiparty				Concurrent Matching, Evaluation Methods, Multiparty, Threshold-Based Classification, Incremental Matching	GPU-Acceleration, Concurrent Matching, Evaluation Methods for each step, Clustering Methods

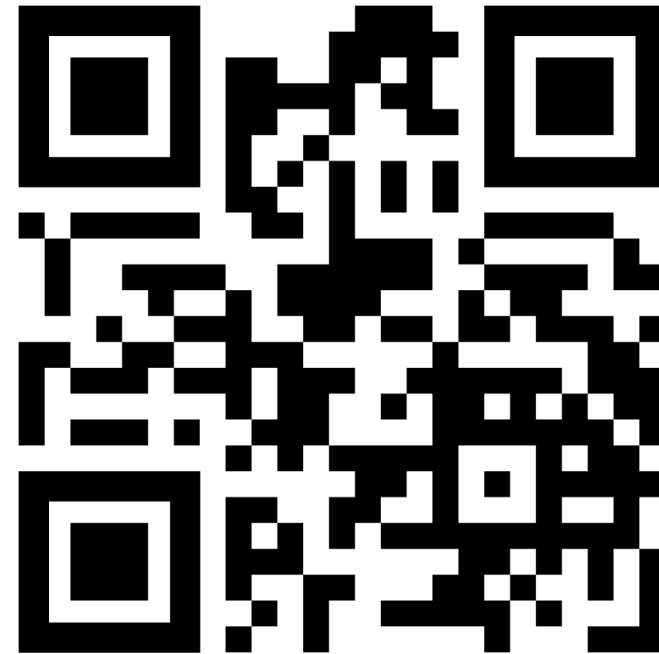
---

## Hands On Tutorial Github Repo



<https://github.com/AI-team-UoA/PPRL-tutorial>

## privJedAI Github Repo



<https://github.com/AI-team-UoA/privJedAI>

# THANK YOU!

Istetsikas@di.uoa.gr

---



**RECITALS**

Privacy-Preserving Data  
Sharing & ID Management



**Funded by  
the European Union**

Q&A