

Tackling Privacy-aware Data Synthesis



Antheas Kapenekakis @ Aalborg University, antheas@cs.aau.dk, eBISS 2023

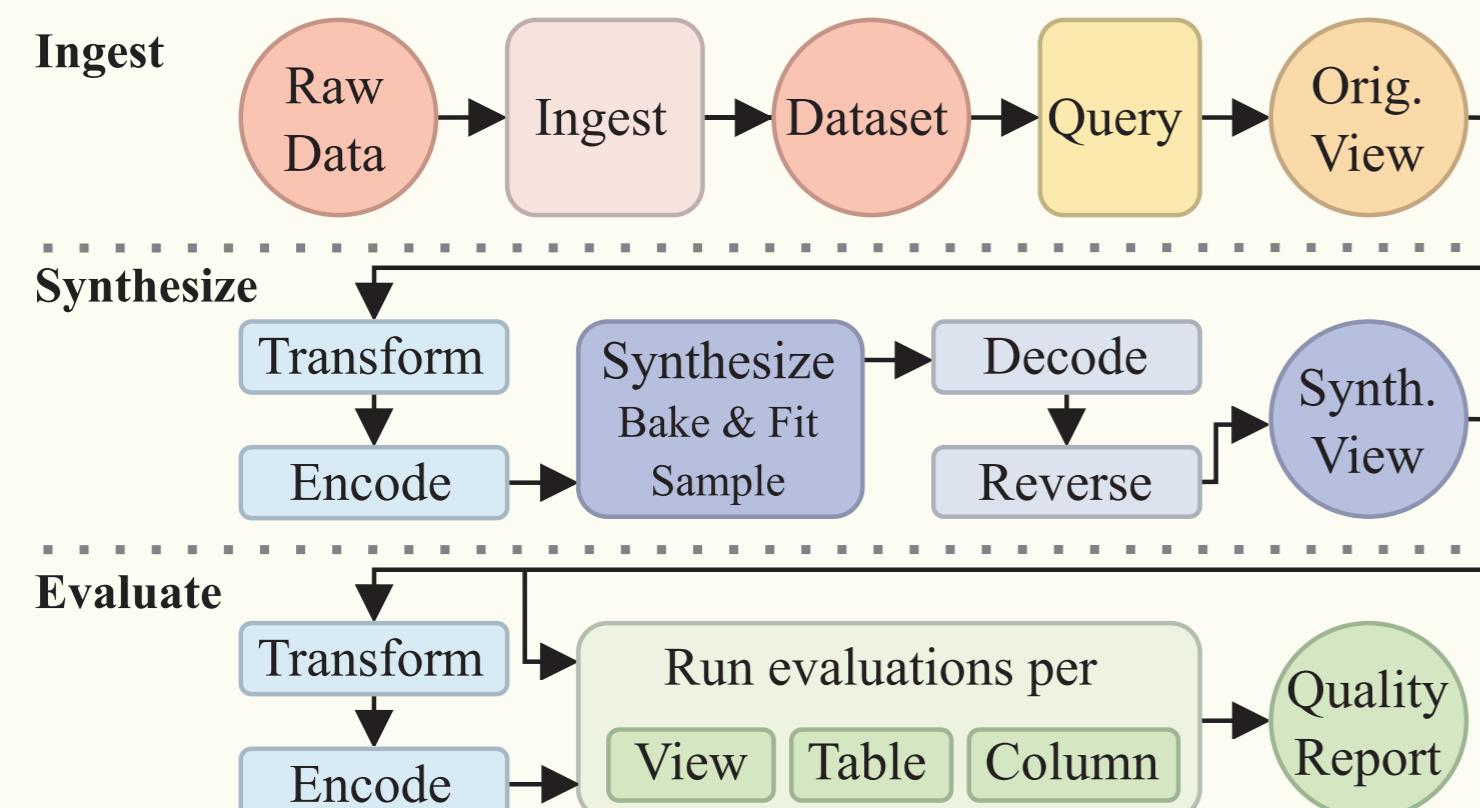
Motivation for Data Synthesis

Medical data has **utility** in the context of **research** and **commercialization**. However, the **sensitivity** inherent in it complicates the **sharing** and **extraction of insights**.

Data synthesis is a **promising** solution. Our research focuses on **structuring** and **scaling** the synthesis process and on proposing new **privacy-aware** methods.

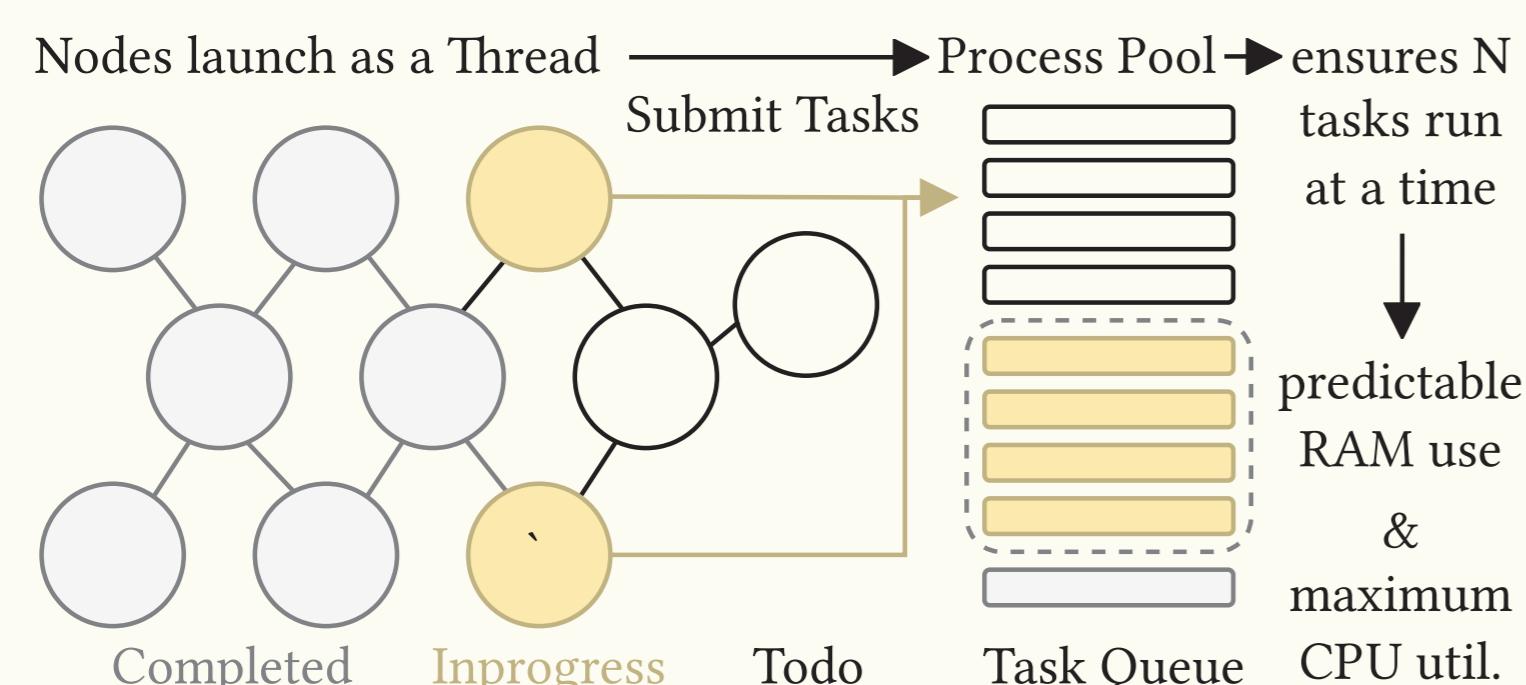
ISE Methodology

We introduce the **Ingest, Synthesize, and Evaluate (ISE)** methodology that **structuring** Data Synthesis.



The Pasteur System

To use the methodology, we create **Pasteur**, a **system** for **larger-than-memory, parallelized** Data Synthesis.



MARE: Event-level Data Synthesis

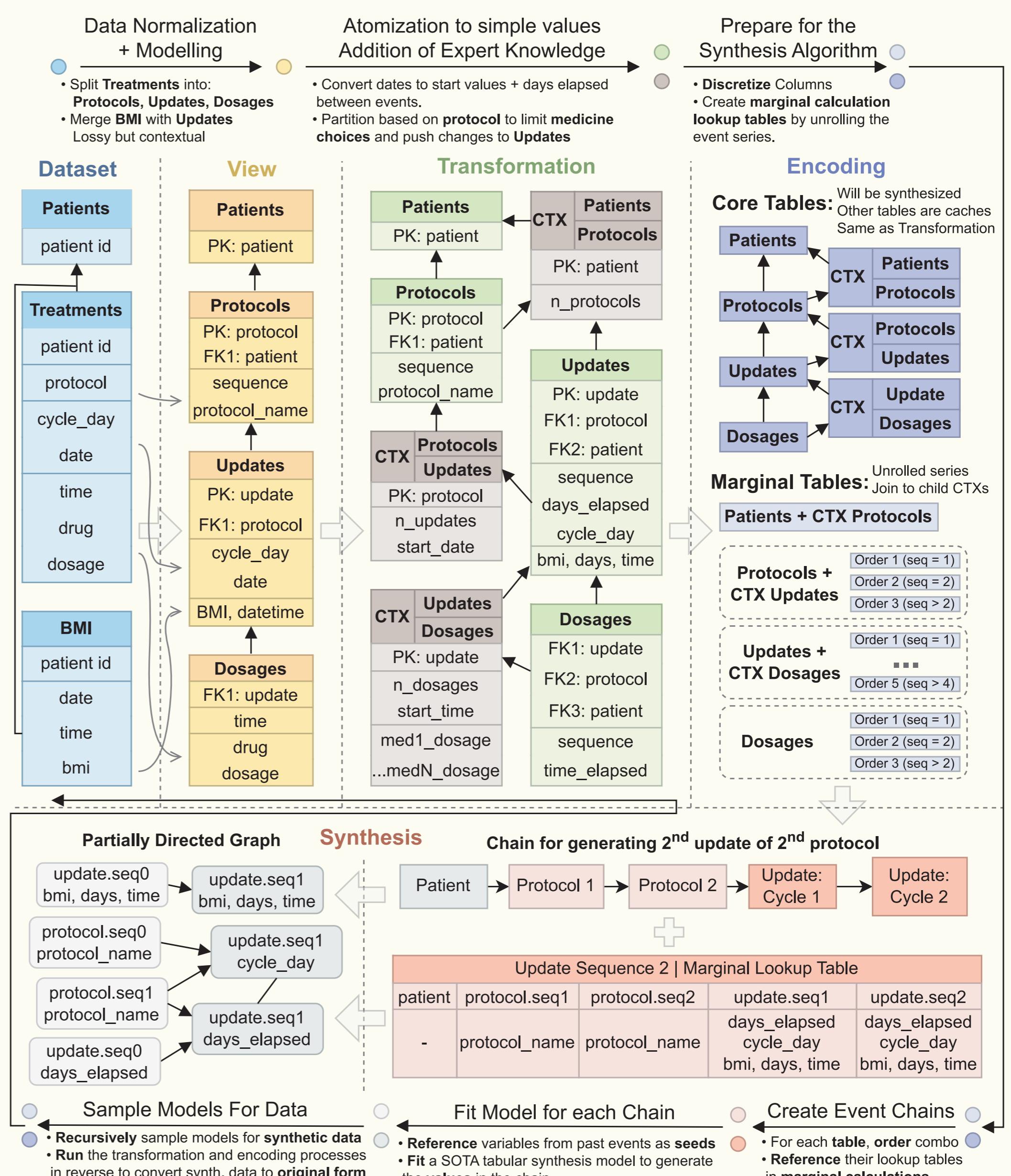
"**MARE: Markov-based Algorithm for Recursive Events**" is a method being co-developed with Aalborg University Hospital for accurate and privacy-aware generation of health events.

aDP: Attributional Differential Privacy

Differential Privacy (DP) by design limits the amount of data that can be generated from each person and protects it equally, causing issues when the **cardinality of input data is large** or **only a portion is sensitive**.

aDP applies to each attribute of input data **individually**, adapting to the **sensitivity** of input data and allowing the **scaling** of input cardinality.

ISE Methodology in MARE



Medical Data used in MARE

Date	Time	Protocol	Cycle_day	Drug	Dosage Information	Medicine
2009-01-18	10:00:00	EXE preoperation	C0 D0	primerper	with X Tablet Oral x Y dg1 for X days PRN	epirubicin
2009-01-18	10:00:00	EXE preoperation	C0 D0	capecitabin	with X Tablet Oral x Y dg1 for X days	oxaliplatin
2009-01-18	10:00:00	EXE preoperation	C0 D0	zofran	with X Tablet Oral x Y dg1 for X days	primperan
2009-01-18	10:00:00	EXE preoperation	C0 D0	prednisolon	with X Tablet Oral daily for X days	zofran
2009-01-18	10:02:00	EXE preoperation	C1 D1	epirubicin	with X Intravenous daily over X for X day	epirubicin
2009-01-18	10:28:00	EXE preoperation	C1 D1	oxaliplatin	with X Intravenous daily over X for X day	oxaliplatin
2009-02-11	10:27:00	EXE preoperation	C2 D1	epirubicin	with X Intravenous daily over X for X day	epirubicin
2009-02-11	11:00:00	EXE preoperation	C0 D0	zofran	with X Tablet Oral x Y dg1 for X days	zofran
2009-02-11	11:00:00	EXE preoperation	C0 D0	primerper	with X Tablet Oral x Y dg1 for X days PRN	primperan
2009-02-11	11:00:00	EXE preoperation	C0 D0	prednisolon	with X Tablet Oral daily for X days	prednisolon
2009-02-11	11:00:00	EXE preoperation	C0 D0	capecitabin	with X Tablet Oral x Y dg1 for X days	capecitabin
2009-02-11	11:03:00	EXE preoperation	C2 D1	oxaliplatin	with X Intravenous daily over X for X day	oxaliplatin
2009-03-01	09:30:00	EXE preoperation	C3 D1	epirubicin	with X Intravenous daily over X for X day	epirubicin
2009-03-01	09:30:00	EXE preoperation	C0 D0	zofran	with X Tablet Oral x Y dg1 for X days	zofran
2009-03-01	09:30:00	EXE preoperation	C0 D0	primerper	with X Tablet Oral x Y dg1 for X days PRN	primperan
2009-03-01	09:30:00	EXE preoperation	C0 D0	capecitabin	with X Tablet Oral x Y dg1 for X days	capecitabin
2009-03-01	09:30:00	EXE preoperation	C0 D0	prednisolon	with X Tablet Oral daily for X days	prednisolon
2009-03-01	09:55:00	EXE preoperation	C3 D1	oxaliplatin	with X Intravenous daily over X for X day	oxaliplatin
2009-05-17	12:00:00	EXE postoperation	C0 D0	capecitabin	with X Tablet Oral x Y dg1 for X days	epirubicin
2009-05-17	12:00:00	EXE postoperation	C0 D0	primerper	with X Tablet Oral x Y dg1 for X days PRN	oxaliplatin
2009-05-17	12:00:00	EXE postoperation	C0 D0	prednisolon	with X Tablet Oral daily for X days	primperan
2009-05-17	12:20:00	EXE postoperation	C1 D1	epirubicin	with X Intravenous daily over X for X day	zofran
2009-05-17	12:42:00	EXE postoperation	C1 D1	oxaliplatin	with X Intravenous daily over X for X day	epirubicin
2009-06-08	09:27:00	EXE postoperation	C2 D1	epirubicin	with X Intravenous daily over X for X day	oxaliplatin
...						
2009-07-05	11:20:00	EXE postoperation	C3 D1	epirubicin	with X Intravenous daily over X for X day	epirubicin
2009-07-05	11:47:00	EXE postoperation	C3 D1	oxaliplatin	with X Intravenous daily over X for X day	oxaliplatin

Protocol 1: EXE preoperation
Protocol 2: EXE postoperation
Protocol Update: Same Day
Medicine Updates: Within 2 hours



Aalborg University
Katja Hose, Daniele Dell'Aglio



Athena Research Center
Minos Garofalakis

