



As health does not happen in a vacuum, **EXTERNAL_EXPOSURE** helps OMOP represent the world around the patient

This extension links where people lived, when they lived there, and what exposures they experienced – while preserving standardization, provenance, and compatibility with federated analytics.

Title: Achievements, challenges, and future directions of the EXTERNAL_EXPOSURE table as an OMOP extension for GIS research

Background

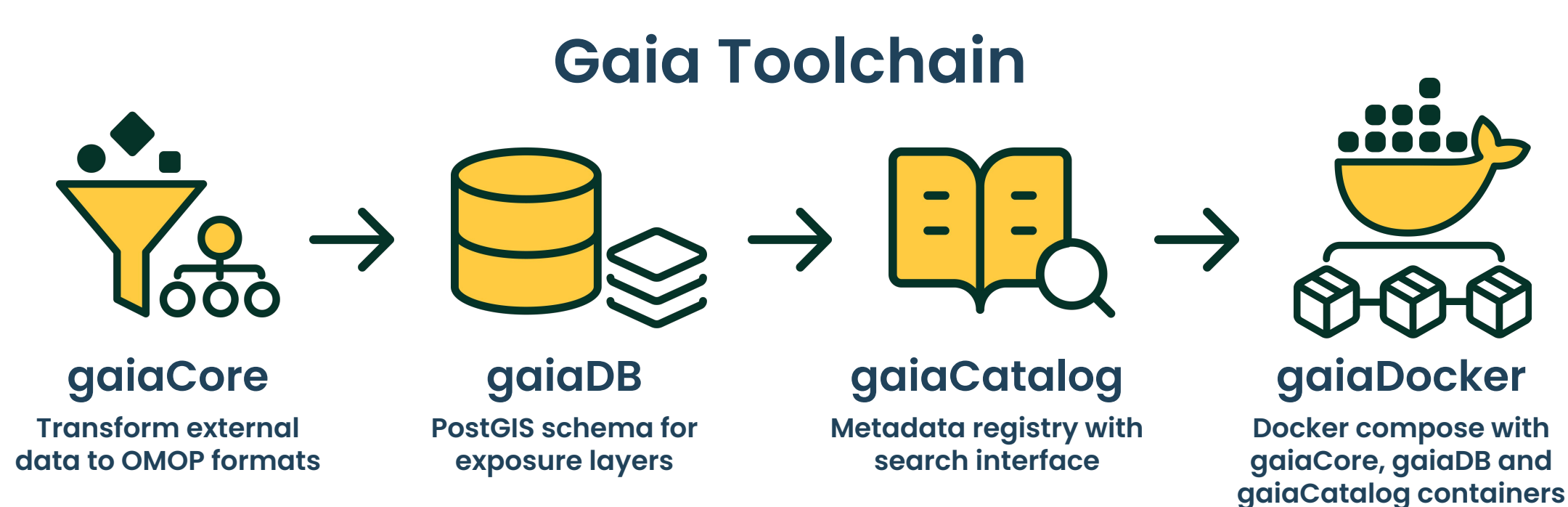
- Environmental and social exposures shape health, but OMOP historically lacked native support for representing them.
- Exposome research requires linking place-based context to person-level clinical timelines.
- The OHDSI GIS Working Group addressed this gap with EXTERNAL_EXPOSURE and LOCATION_HISTORY.
- The goal is a standardized, privacy-preserving, analytically tractable way to treat external exposures as first-class data elements in OMOP.

Objectives

- To implement a standardized OMOP-based framework for representing person-level environmental and social exposures across space and time.
- To support reproducible, federated observational research on exposome and SDOH effects alongside clinical data.

Methods

- Designed EXTERNAL_EXPOSURE as an OMOP-style event table for person- and location-linked exposures with standardized concepts, time intervals, provenance, spatial relationship, and numeric or categorical values.
- Used LOCATION_HISTORY to model residential trajectories as location_id plus date intervals.
- Assigned exposures only when residential and exposure time windows overlapped.
- Built the Gaia toolchain for implementation:

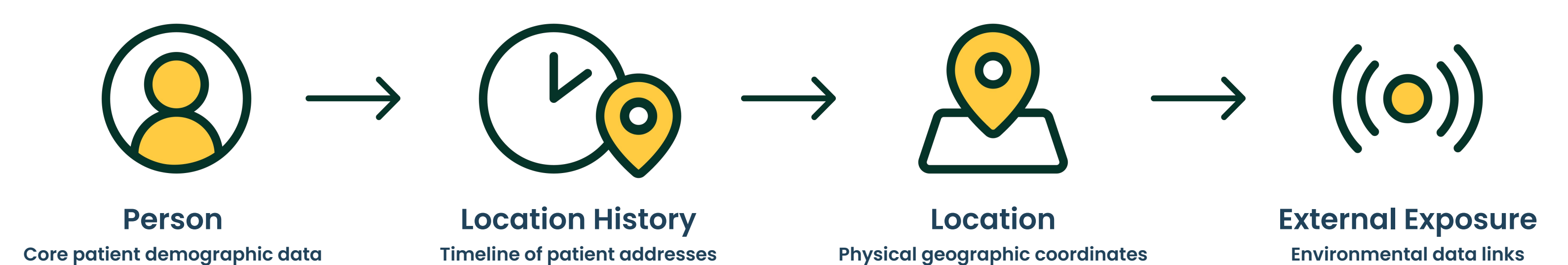


- Ran metadata-driven ETL pipelines to harmonize external datasets, perform spatial joins to residential locations, and load provenance-traceable records into EXTERNAL_EXPOSURE.
- Applied the GIS Vocabulary Package to standardize environmental and SDOH concepts, geographies, and spatial relationships.

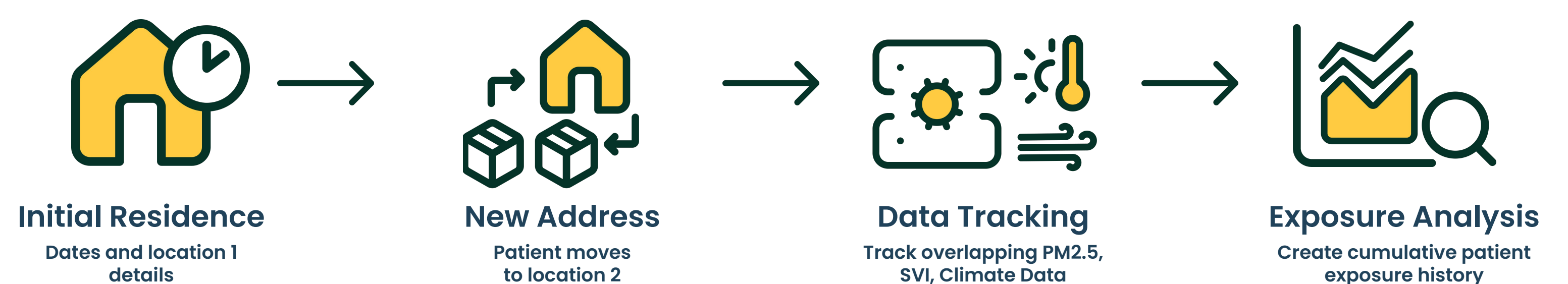
Results

1. In 2024, the team released a fully documented, publicly available first version of the schema, vocabulary extensions, ETL conventions, and toolchain components.
2. The pipeline was demonstrated on synthetic data, successfully loading temporally aligned exposures such as PM2.5, deprivation indices, and climate exposures into EXTERNAL_EXPOSURE.
3. The NIH Bridge2AI CHoRUS initiative has adopted the schema for community-level SDOH and environmental context integration. CHoRUS is mapping variables such as ZIP-code-level poverty, housing conditions, and pollutant levels to critical care patients using this model.
4. The schema successfully represented composite indicators such as SVI using multiple rows per person with distinct concept_id values and value_as_number for quantile rank.
5. Analytical results from CHoRUS are forthcoming, but the implementation supports the schema's fitness for large-scale multimodal data.

Schema Relationship Flow



Patient Environmental Exposure Timeline



Why it is metter?

This moves OMOP closer to representing not only what happened in care, but also what happened around the patient.

The design is structurally OMOP-native enough to support ETL consistency, provenance tracking, and future analytical reuse.

External exposures can now participate in cohort definitions, temporal logic, and stratification—not just sit in a sidebar file.

For OHDSI, this opens a practical path toward exposome-aware real-world evidence generation across networks.

Limitations

- Environmental datasets vary in spatial resolution, temporal resolution, coverage, and units, requiring strong metadata and conversion logic.
- Spatial joins and temporal alignment increase ETL complexity and require geospatial expertise plus performant infrastructure.
- Scalability remains an issue for repeated measures, especially daily exposures across millions of individuals.
- Some use cases need better support for aggregated indicators and categorical encodings, prompting proposed schema revisions.
- Broader adoption is constrained because the table is not yet part of the OMOP core specification.
- Full analytical use may require enhancements in ATLAS and HADES to better support exposure-specific semantics.

Conclusion

EXTERNAL_EXPOSURE provides a standardized mechanism to represent person-level environmental and social exposures in OMOP CDM. The schema is already implemented, usable, and evolving – positioning OHDSI networks to study how place and health interact with more rigor and less improvisation.

- 1 Polina Talapova
- 2 Andrew Williams
- 2 Jared Houghtaling
- 3 Kyle Zollo-Venecsek
- 1 Maksym Trofymenko
- 4 Jay Greenfield
- 5 Timothy Norris
- 2 Robert Miller
- 6 Jimmy Phuong
- 7 Anne Thessen
- 8 Timothy D. Rossi

- 1 IT company SciForce, Kharkiv, Ukraine
- 2 Tufts Clinical and Translational Science Institute, Boston, MA, United States
- 3 Valo Health, MA, USA
- 4 CODATA, Paris, France
- 5 The University of Miami, FL, USA
- 6 University of Washington, WA, USA
- 7 University of North Carolina Chapel Hill, NC, USA
- 8 Architectural Medicine LLC, NM, USA

