

DDI Lifecycle and Qualitative Data: Development of a Formal Model

Arofan Gregory

Joachim Wackerow

Background

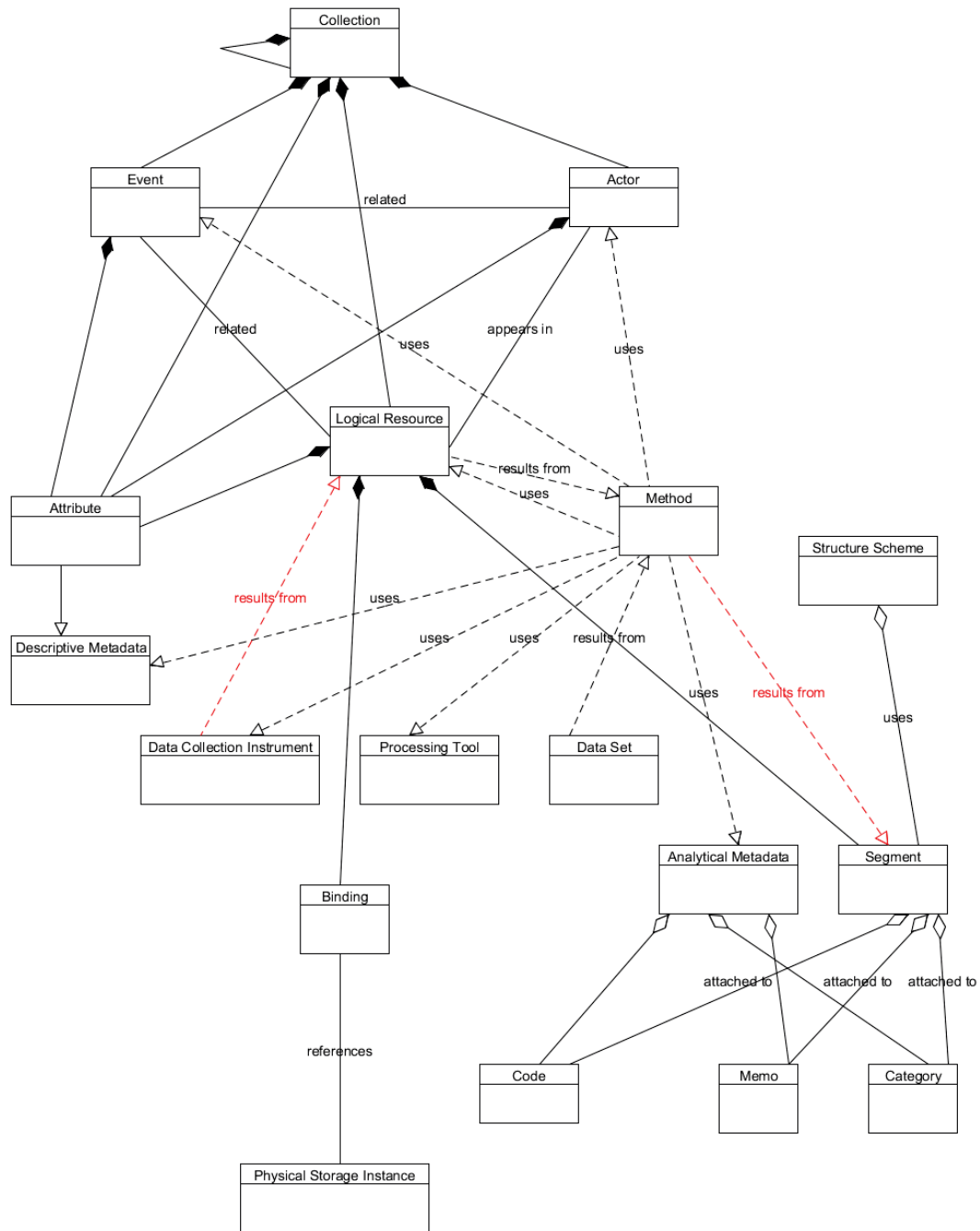
- We have been talking about metadata for qualitative data for a long time
 - QuDEX
 - Other work
- The model presented here is one which came out of our working session at EDDI 2011 in Gothenburg
 - 25 people
 - Several outputs

The Problem Space

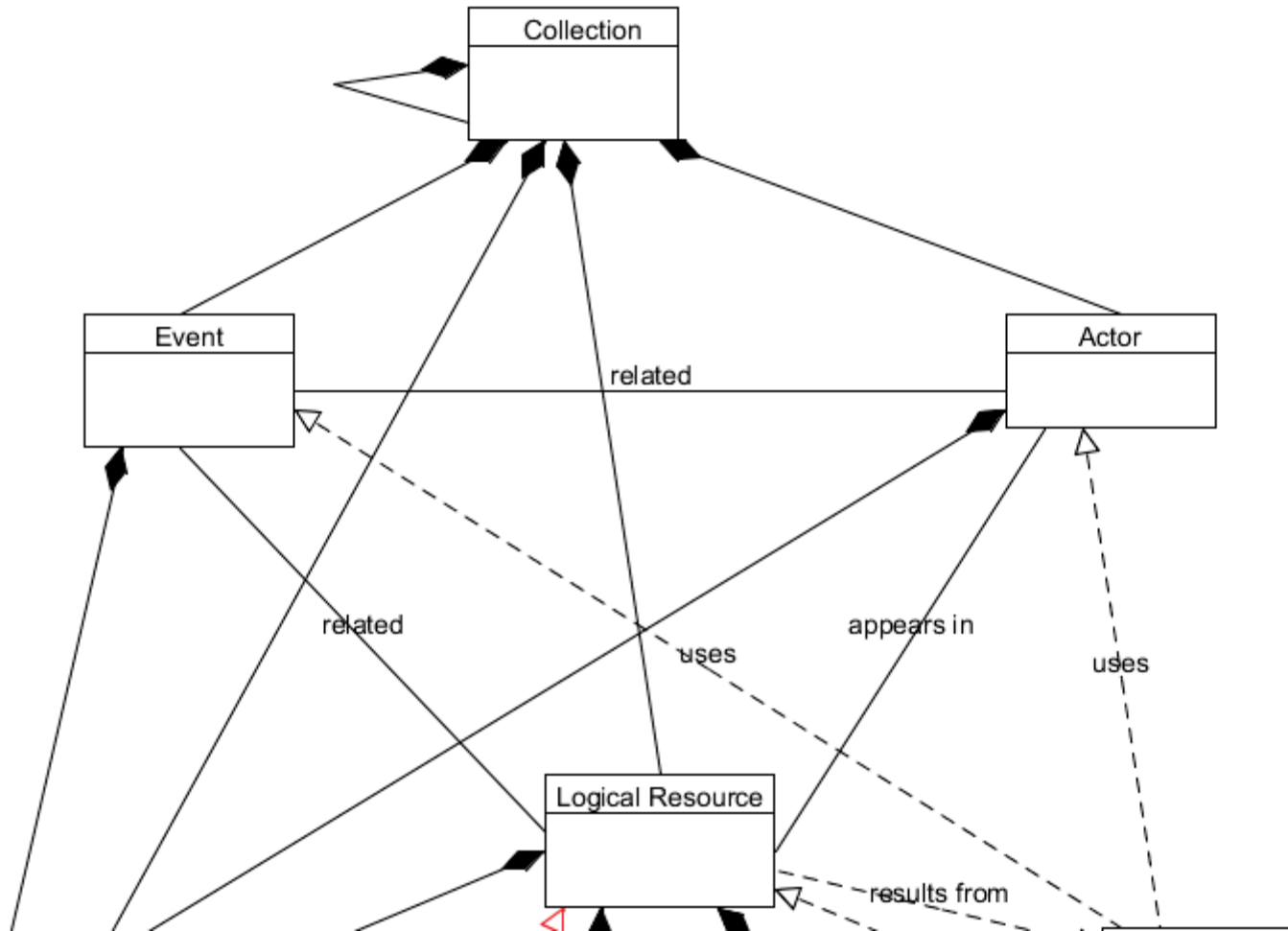
- Qualitative methods are still being explored and developed
 - Text analysis
 - Multi-media tools (CAQDAS)
 - Hand-crafted methods
- But mixed-method and qualitative studies should be manageable like quantitative studies

The Model

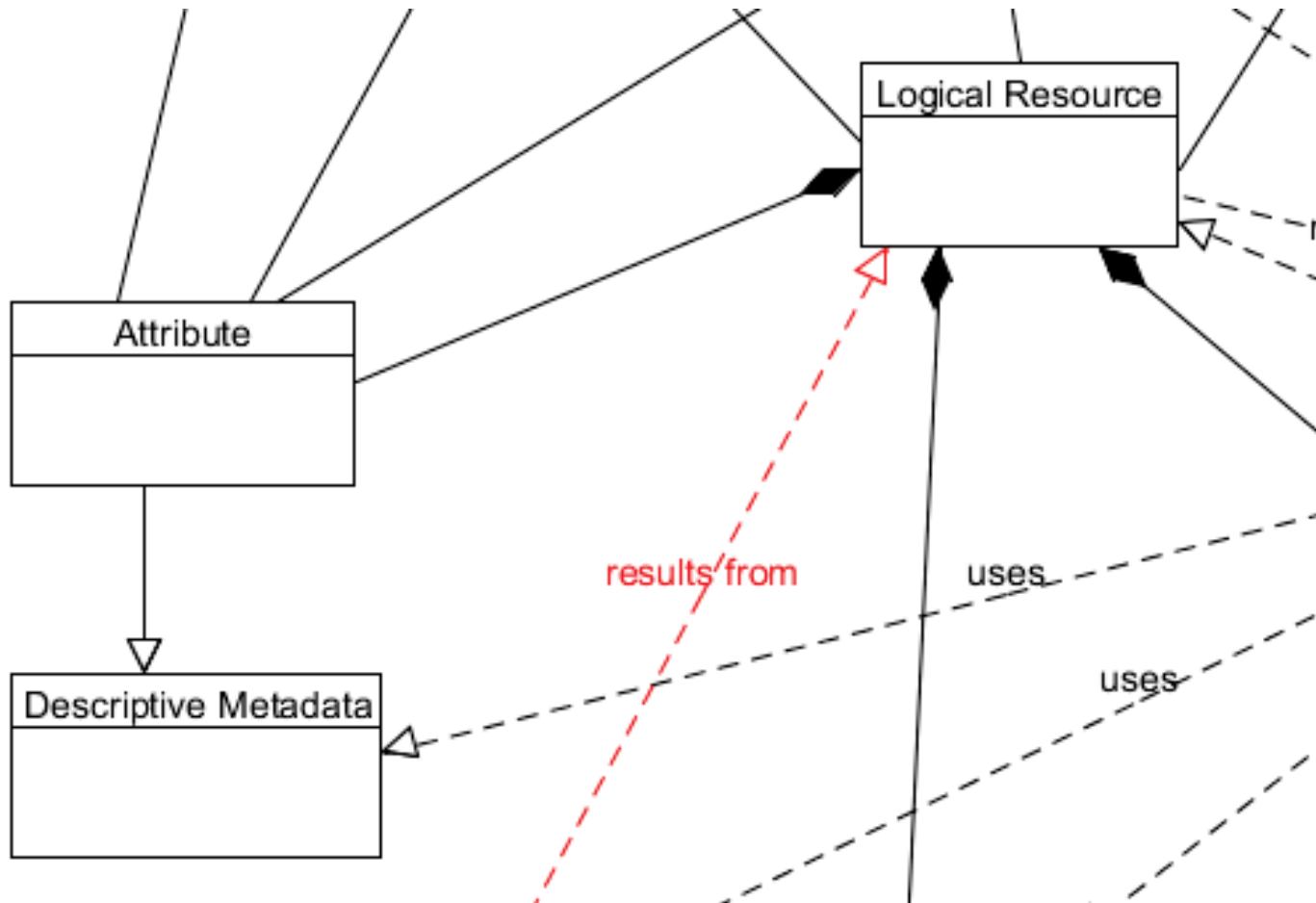
- A high-level conceptual model
 - DDI will be model-driven in future versions
- Still a draft
 - This version is the cleaned-up one done by Larry Hoyle



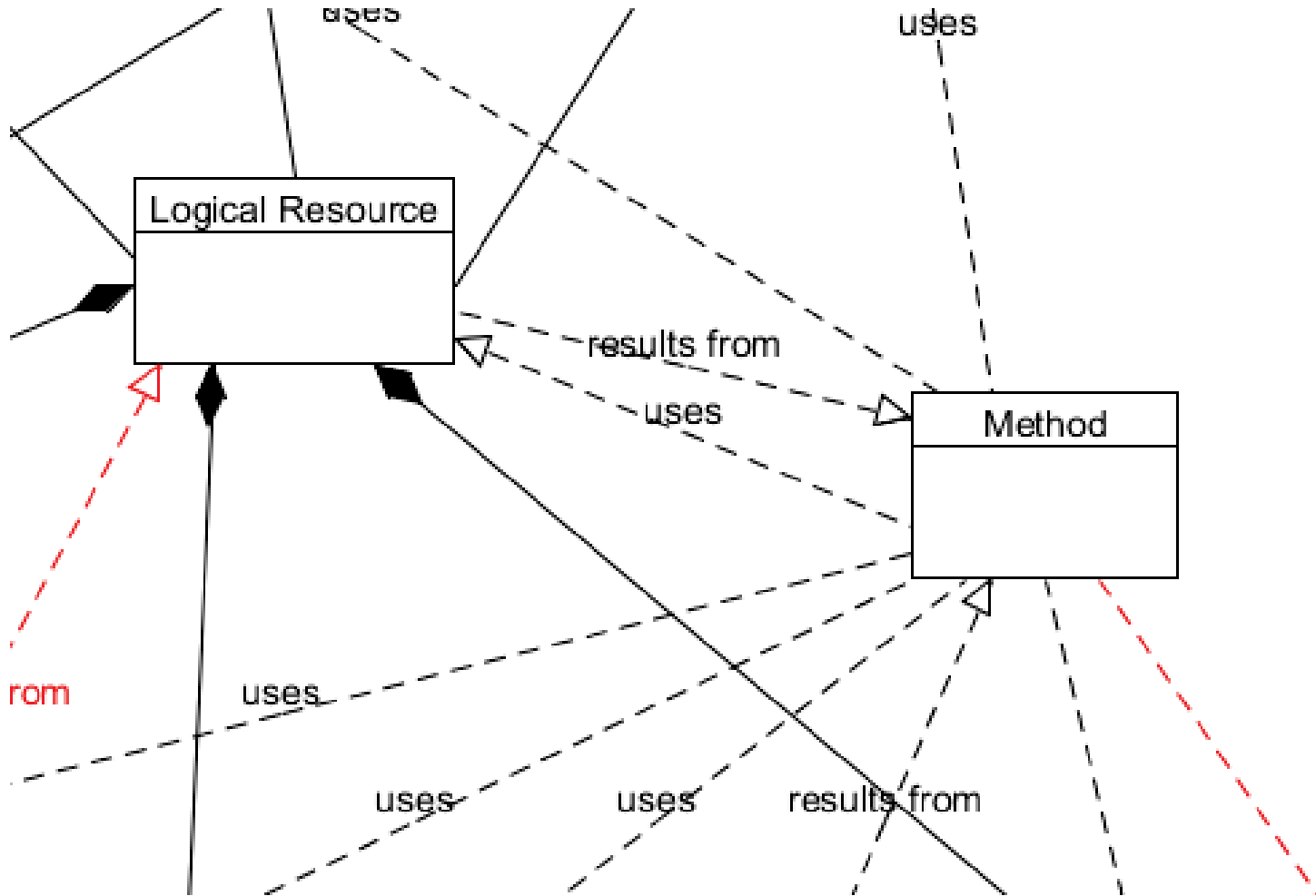
Collections



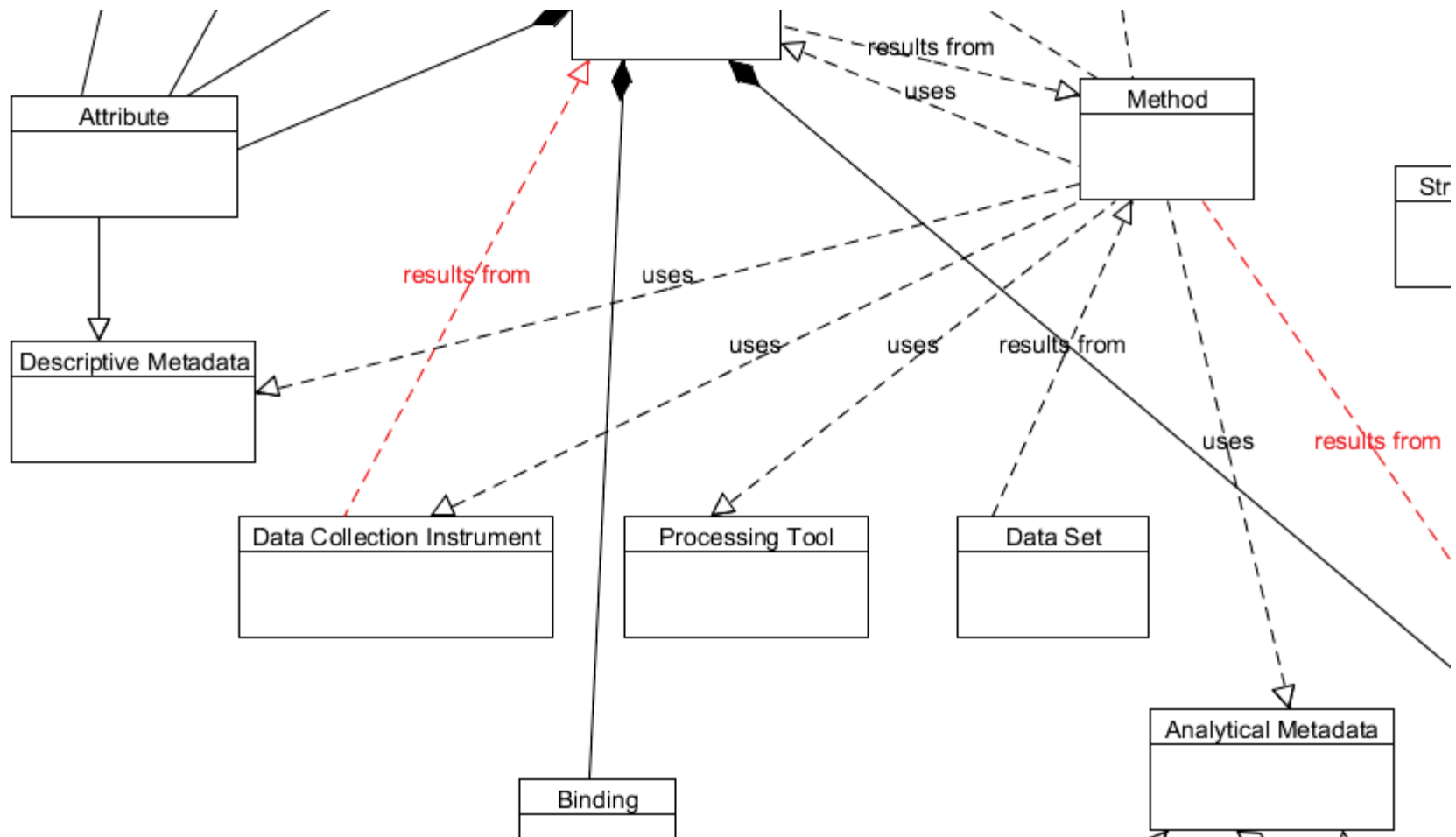
Collections, Actors, Events, and Logical Resources Have Attributes



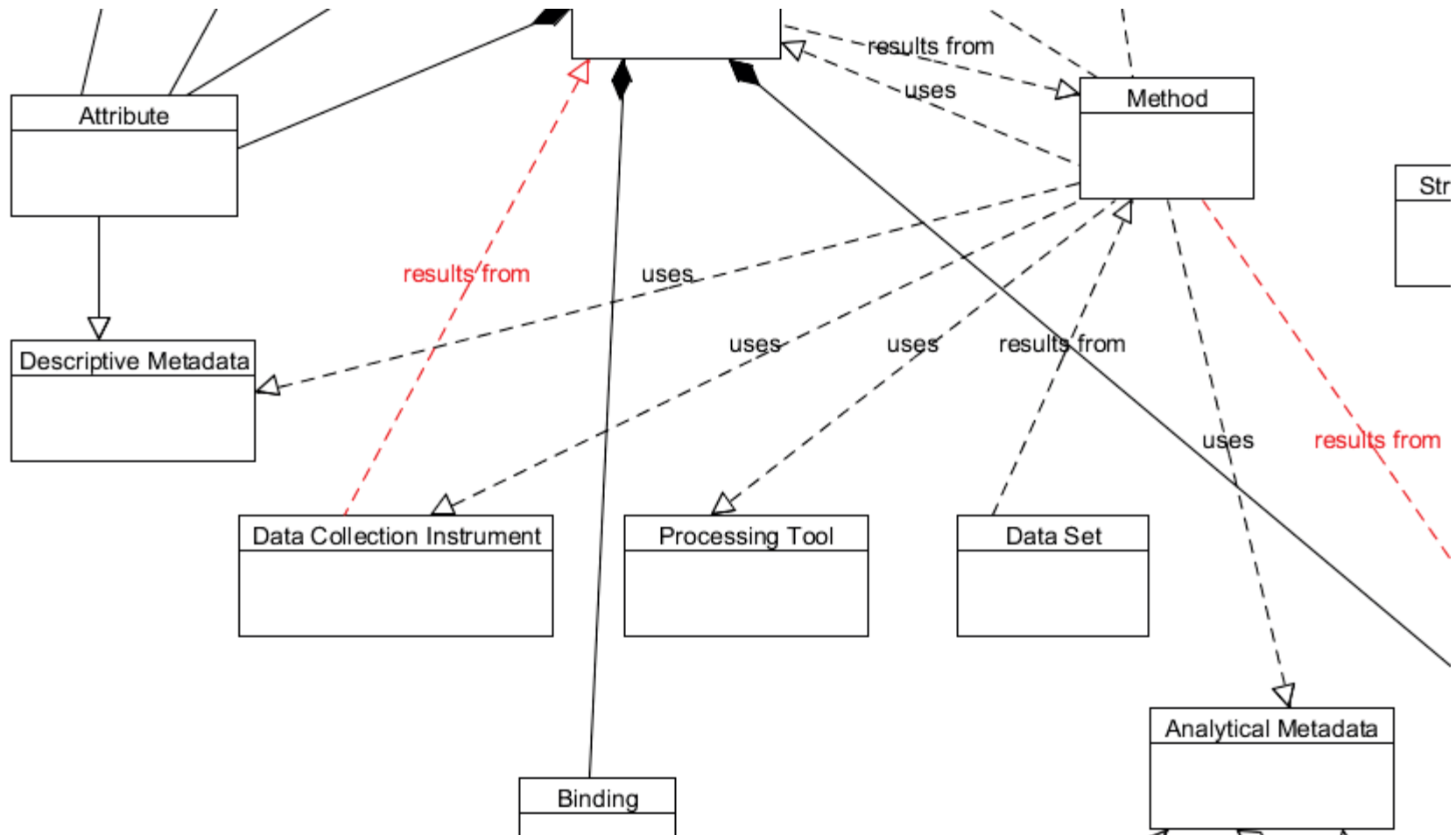
Logical Resources Use and Are Produced by Methods



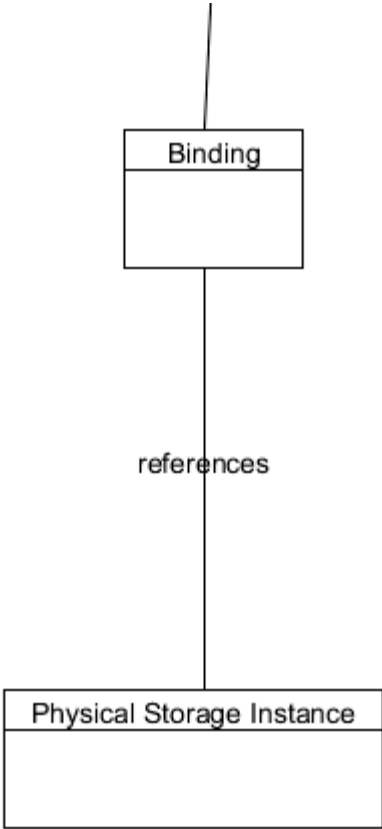
Methods Use Instruments and Processing Tools



Methods Produce Descriptive and Analytical Metadata and (Quantitative) Data Sets



Logical Resources Have Bindings to Physical Storage Instances

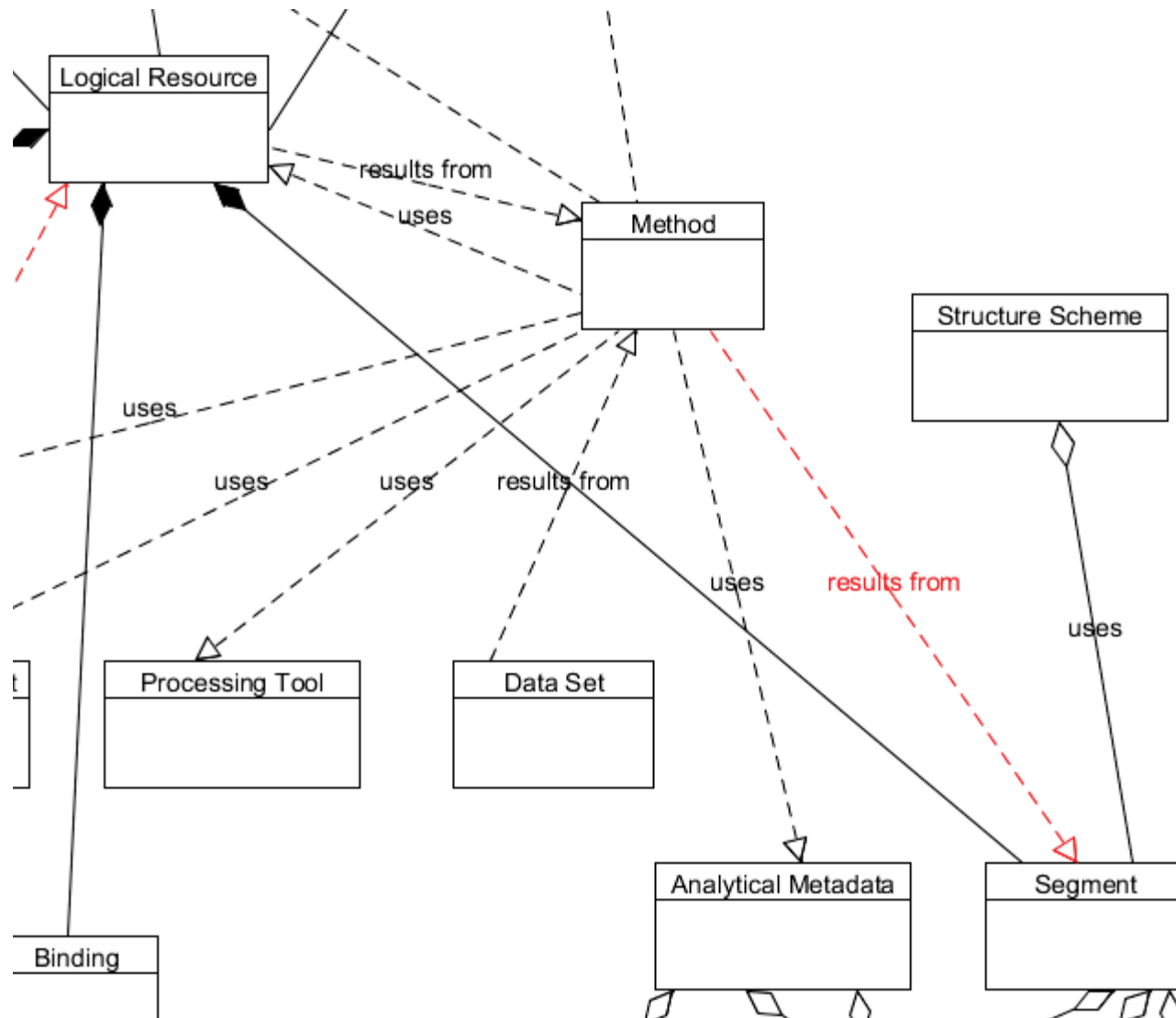


Separation of Logical and Physical

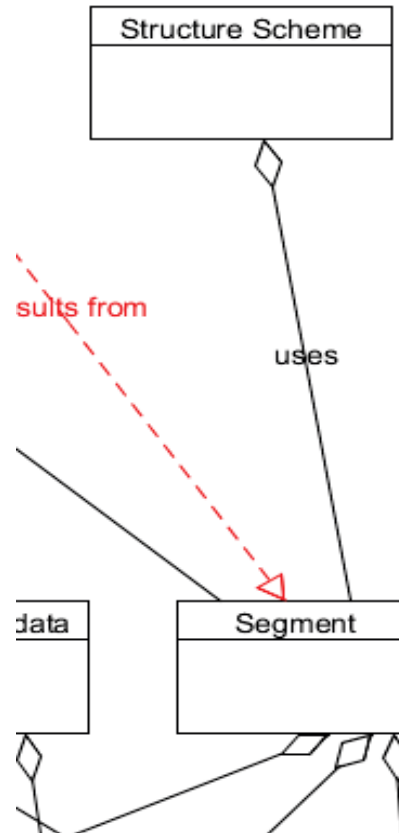
- Like DDI-Lifecycle, the logical and physical aspects of a resource are modeled separately

Logical Resources Can Have Segments

- All or Part of the Resource



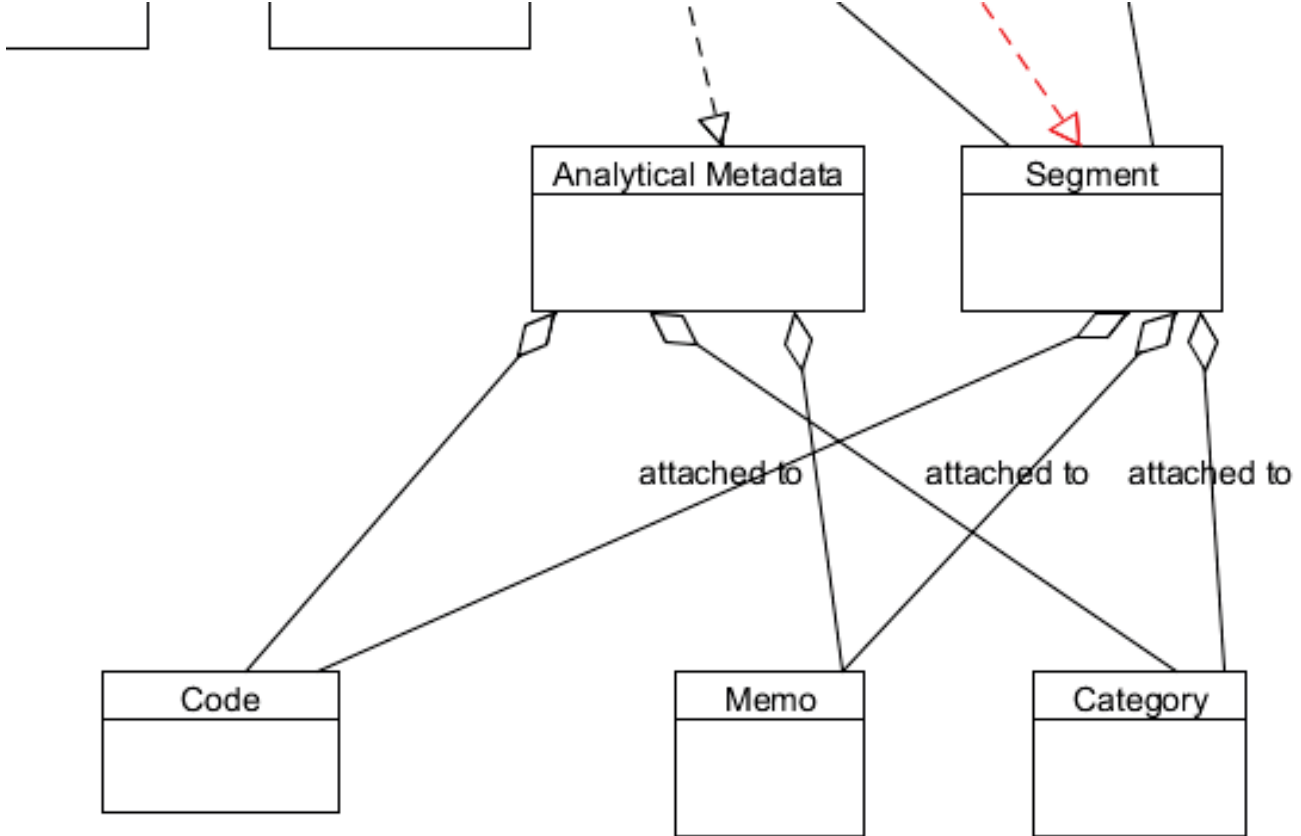
Segments Are Defined According to A Structural Scheme – Different for Different Types of Files



Structural Schemes Vary by Type

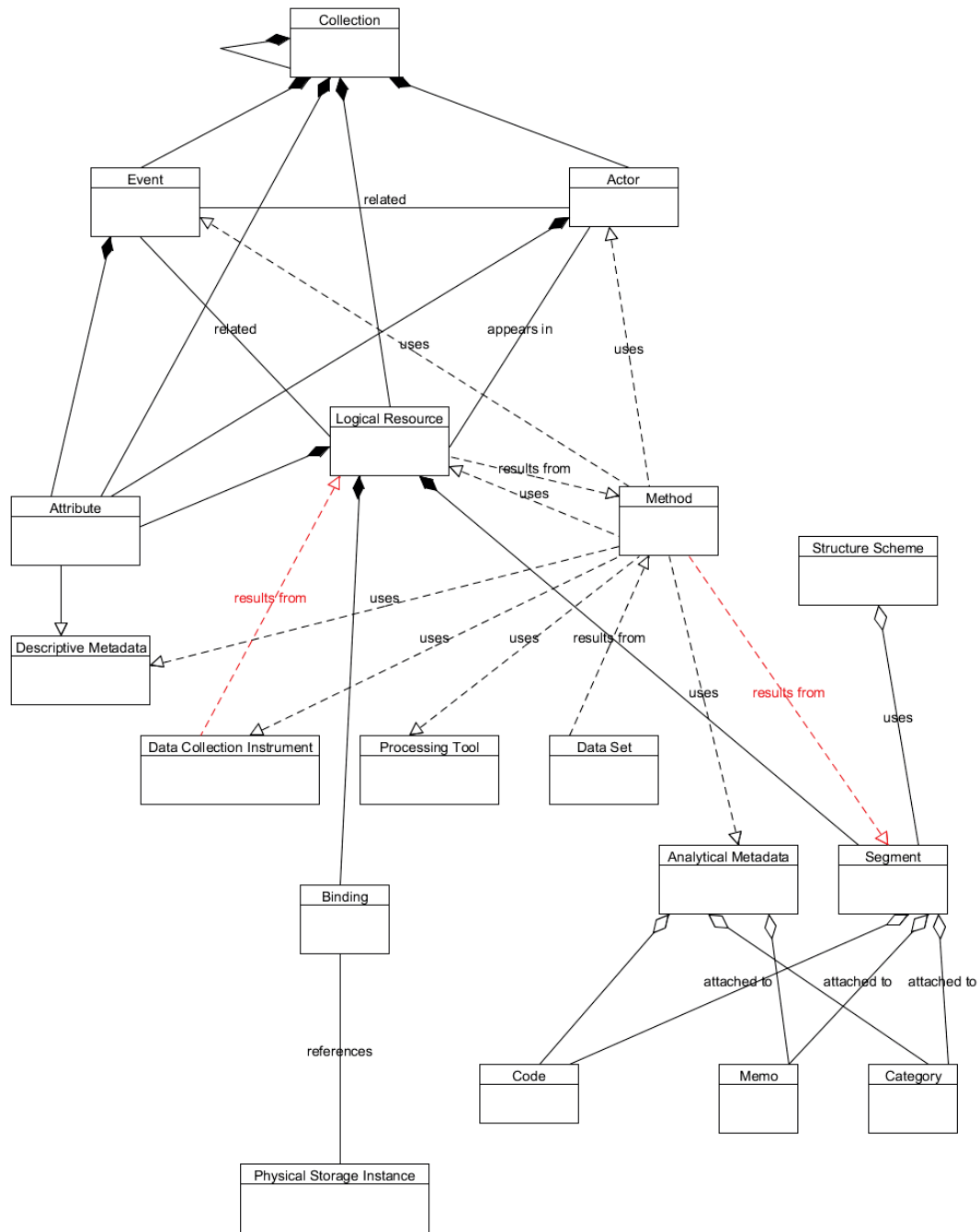
- Text has lines and characters
- Images have coordinates
- Sound files have start and end time
- Video has coordinates and start and end time
- XML has nodes
- Note that this model is already in DDI-Lifecycle
 - It was taken from QuDEX

Segments Can Have Analytical Metadata Attached – Memos, Codes, and Categories



Analytical Metadata

- Memos are textual annotations
- Codes are codes assigned at analysis time (no formal scheme)
- Categories are codes coming from a formalized scheme
- You can have arbitrary relationships between segments and these other bits of metadata (not shown in model)



Connections to DDI - Lifecycle

- Actors are Organizations/People (Organization Scheme)
- Data Sets are Data Sets (with all regular DDI-Lifecycle metadata)
- Data Collection Instruments are instruments/collection events
- Codes and Categories are Codes and Categories

To Learn More...

- Panel: Preserving and organising qualitative social science data resources for sharing and re-use
 - 13:45 – 15:15, Thursday