Natural Information Systems

Nick Rossiter
Computer Science and Digital Technologies
Northumbria University

With thanks to Michael Heather, Paul Vickers,
Michael Brockway

Natural

- An object/arrow is defined uniquely over 3 levels up to natural isomorphism
- 3 levels of arrows:
  - Object A in Category C, $1_A: A \rightarrow A$
  - Functor F: C → C; G: C → C
  - Natural Transformation α: F → G
- Information System:
  - As in Universe

Current Information Systems

Existing approach is models based on a standard mathematical structure:

Hierarchical	Trees
Network	Graphs
Relational	Sets
Functional	Functions
Object-oriented	Objects/methods

What about Categories?

- Previous work:
  - Rosebrugh and co-workers
  - Spivak
  - Baclawski
- Has largely concentrated on:
  - Representing relational databases or the entity-relationship model in categories
  - Sketches popular
  - Developed by Charles Ehresmann

Sketch outline

- A sketch is a
  - Graph with
    - a set of diagrams
    - a set of cones defining which diagrams have limits
    - a set of cocones defining which diagrams have colimits
- Compared to Cartesian closed, a sketch relaxes
  - The terminal object requirement
  - The need for all diagrams to have limits and colimits
- But a sketch is limited to graphs

Categorification

- Use of sketches is categorification
  - Transforming existing techniques on a 1:1 basis from application to categories
  - May provide useful support for current database model research.
- That does not realise the full potential for category theory in advancing database techniques

Fundamental Approach

- Start from basics
  - a clean sheet
- Decide on requirements for an information system
- Identify features of category theory that help to meet requirements
- Produce a framework which satisfies software engineering principles:
  - High cohesion
  - Low coupling

Candidate - Topos

- Categorical structure
- Attracts much attention in standard texts
  - Mac Lane (CWM), Goldblatt (Topoi), Johnstone (Topos Theory)
- Captures properties of sets
- Based on Cartesian closed categories (CCC)
  - Basis of much of our recent work
  - Also dealt with fully in standard texts

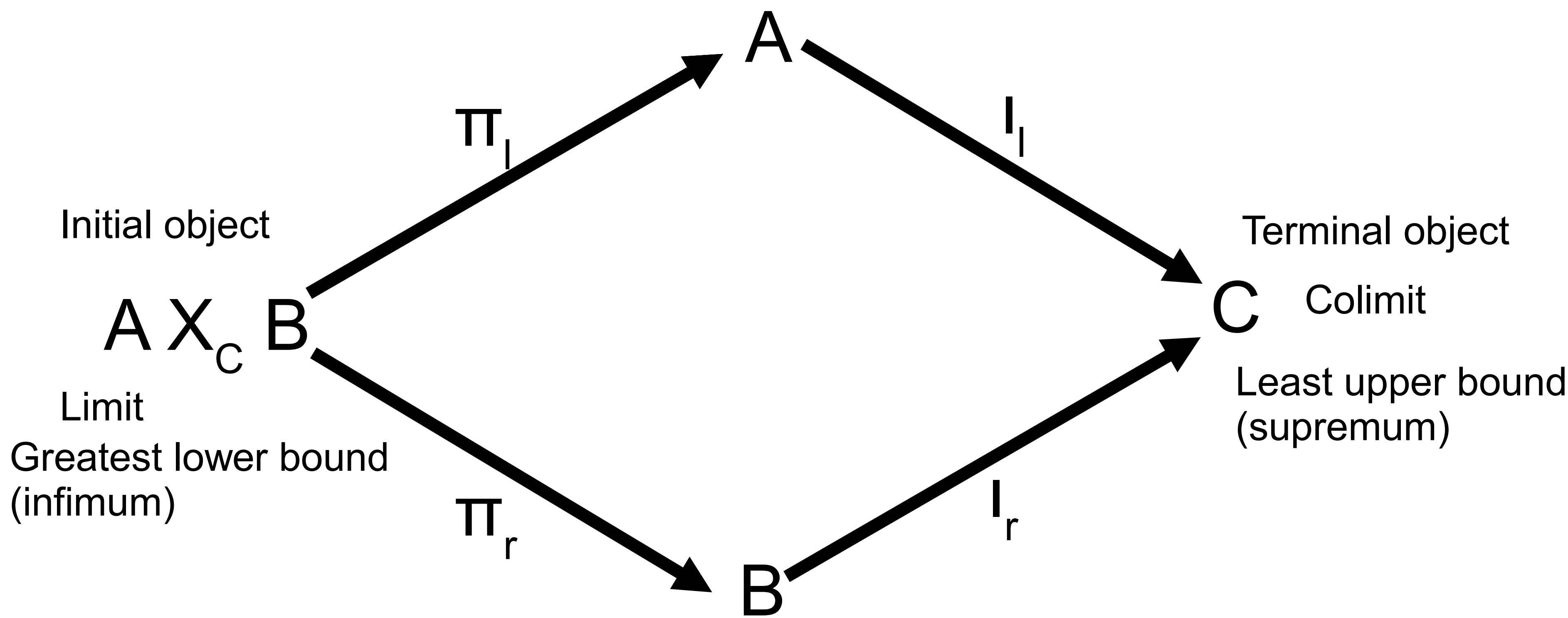
The Topos - Definition

- A category $\mathcal{E}$ that
  - is finitely complete
    - Limits of all finite diagrams (cones)
  - is finitely co-complete
    - Colimits of all finite diagrams (cocones)
    - Follows automatically if finitely complete and CCC
  - is Cartesian closed
  - has a subobject classifier

Example of CCC

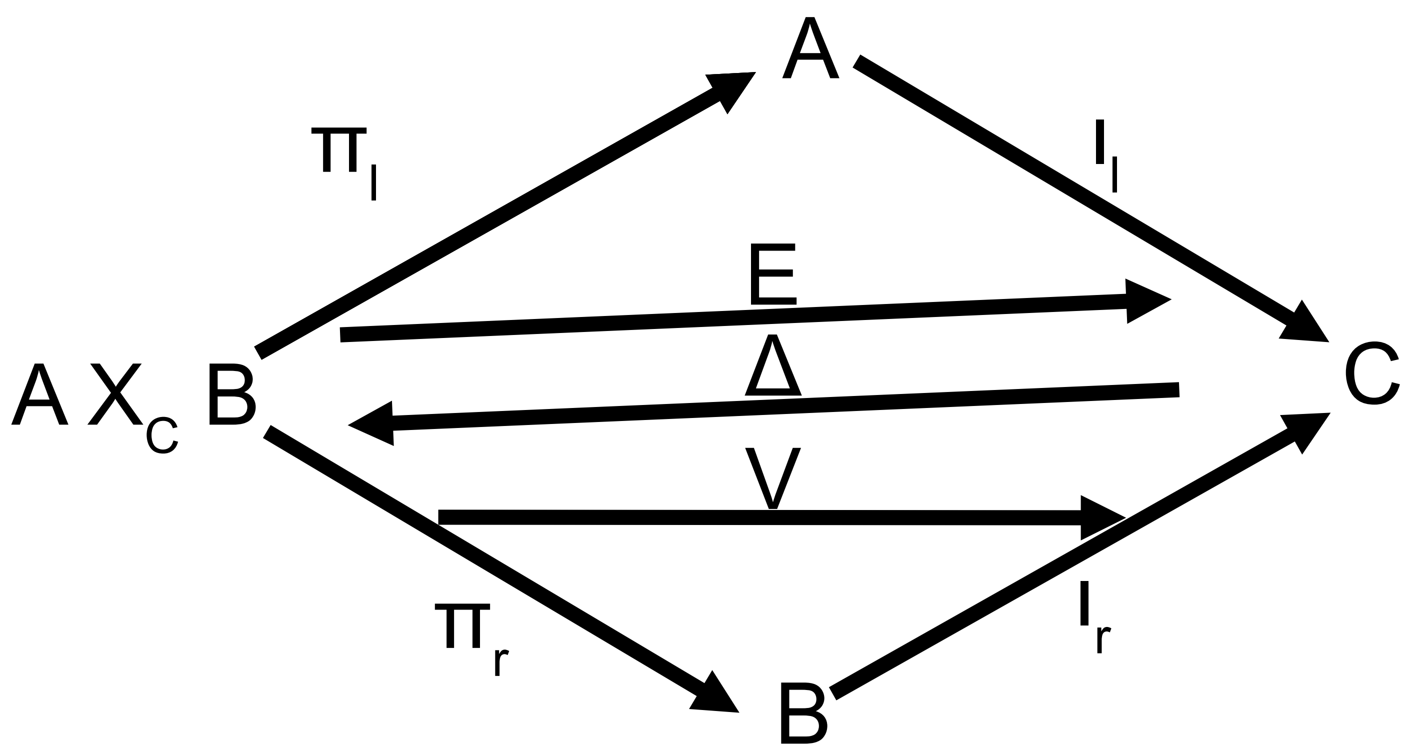
- Pullback
  - Terminal object 1
    - Exactly one arrow from every object in category to the terminal object (least upper bound)
  - Products
    - All objects A,B are related through products A X B
  - Exponentiation (connectivity)
    - hom(A x B, C) ≡ hom(A, C^B)
    - F: _ X B: Ƹ--> Ƹ; G: $\_^{B}$: Ƹ → Ƹ;   F --| G

Pullback – Cartesian closed Category



C is A+B+C

Pullback Logic



Adjointness requirements E ┤ Δ and Δ ┤ V

Build-up of Logic: E existential quantifier, V universal quantifier, Δ diagonal functor, Π projection; ı inclusion

Arguments for Topos

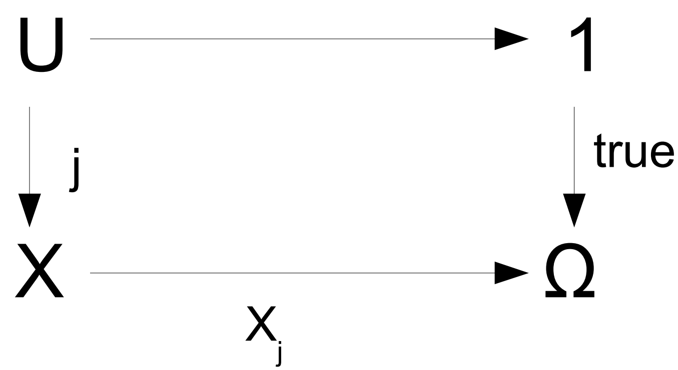
- The natural categorial structure for information systems is the topos with its:
  - cartesian product for representing relationships
  - terminal object for identity
  - subobject classifier for membership criteria
  - internal logic Heyting for query and rule processing
    - internal logic is object oriented
    - Heyting is intuitionistic, more general than Boolean

Arguments against Topos

- No readily-accessible examples of usage (applications)

Subobject Classifier

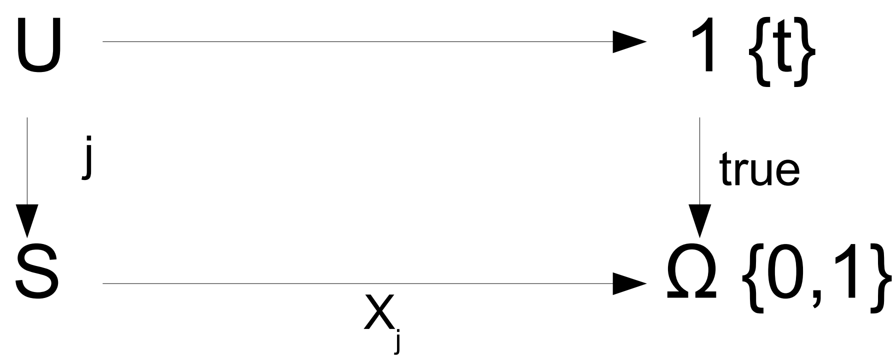
- Live within the topos
- Defined by pullback square:



- Where Ω is the subobject classifier
  - 1 is the terminal object of the topos
  - j: U → X is an arrow in C
  - X_j is the characteristic function
  - U is the limit of the pullback, X is the subobject

Subobject Classifier Example 1

- Defined by commuting pullback square:

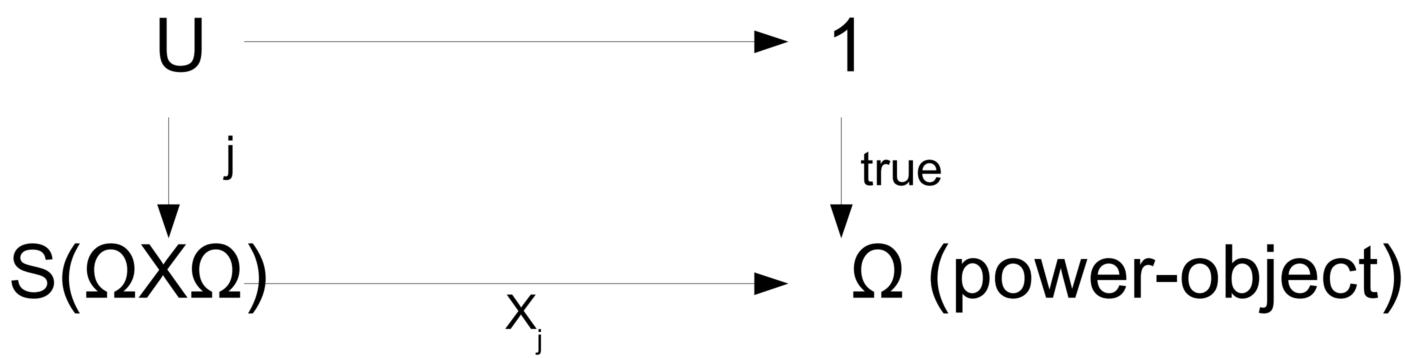


Subobject classifier is Boolean {0,1}

Characteristic function Xj defines subobject S of category represented by 1

Subobject Classifier Example 2

- Defined by commuting pullback square:



Subobject classifier is non-Boolean, a power-object of some objects

Characteristic function Xj defines subobject S of category represented by 1

S is of type AND (intersection)

Topos Logic

- Quantification, projection, product, join through pullback diagrams
- Mitchell-Bénabou Language of a Topos
  - Types are defined + variables of the types
  - Formulae are defined to build expressions
  - Predicates are constructed for membership tests
  - Logical operations include: intersection, union,
  - Internal logic is intuitionistic (Heyting)
    - Handling of negation is more sophisticated
      - e.g. 'not unhappy'

Requirements/Capabilities

Approach	Maths structure	Structuring Capability	Searching	Query symmetry	Query Closure	Transactions	Interoperability	Commercial Examples
Hierarchical	Trees	1:N	Tree traversal	No, bias to downwards direction	No, tabular display	Yes, CICS	No	IMS
Network	Graphs	N:M	Graph traversal	No, bias to initially defined paths	No, tabular display	Yes	No	IDMS
Relational	Relations of sets	N:M	Set operations	Yes	Yes	Yes	No	Oracle, DB2, Access
Functional	Functions	1:N	Function composition	No, bias to initially defined paths	No	No	No	None
Cartesian-closed	Pullbacks	N:M	Quantifiers	Yes	No	Yes, monad	Yes, natural	None
Topos	Categorical Topos	N:M	Quantifiers, internal Heyting logic	Yes	Yes	Yes, monad	Yes, natural	None

Visualisation Application

Vickers, Paul, Faith, Joe, & Rossiter, Nick,

Understanding Visualization: A Formal Approach using Category Theory and Semiotics,

IEEE Transactions On Visualization And Computer Graphics 158. IEEE Transactions On Visualization And Computer Graphics, 2012 Jun;19(6):1048-61. doi: 10.1109/TVCG.2012.294 (2013). pdf

Initial Category for Visualisation

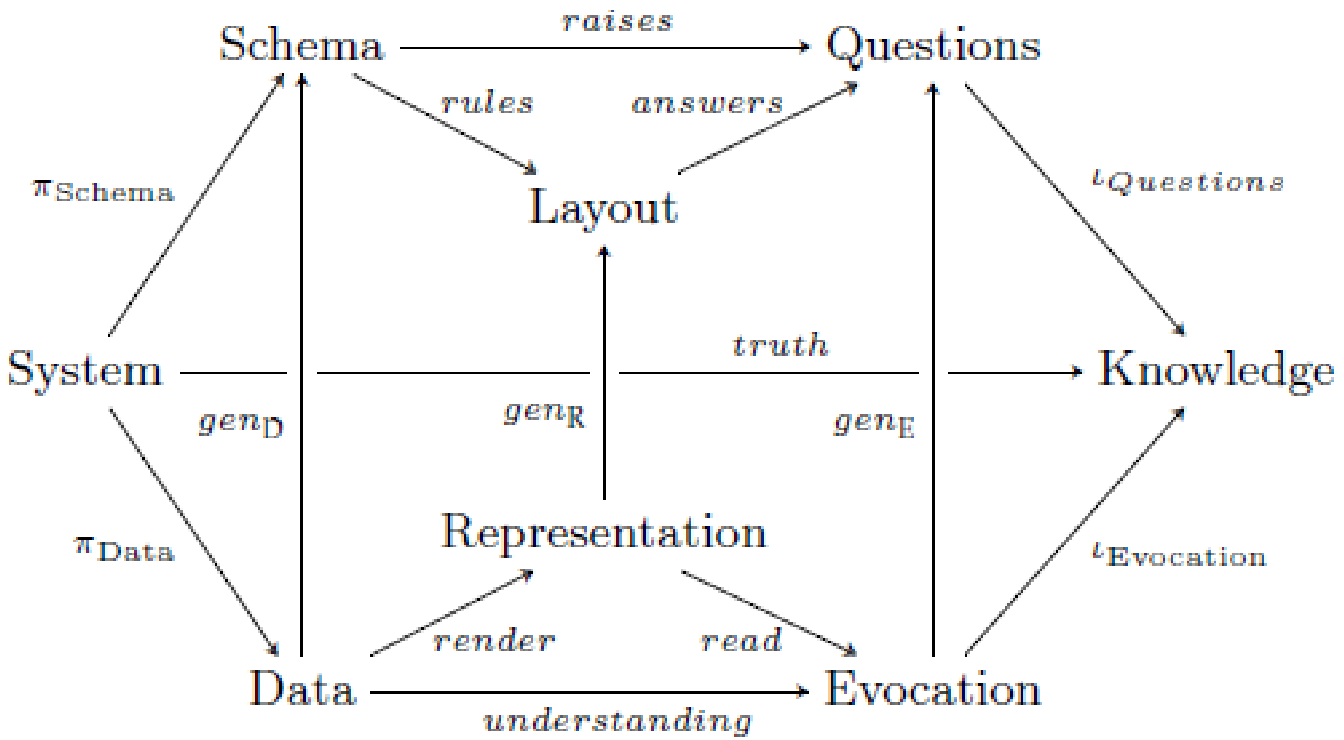


Fig. 3. The visualization process category with product System and co-product Knowledge. System has projection morphisms to Schema and Data. Questions and Knowledge have inclusion morphisms to the co-product Knowledge.

Paper 2

Under development

Comparing Visualizations: Equivalence in Perceptualization Processes

Paul Vickers, Nick Rossiter, Michael Brockway, and Joe Faith.

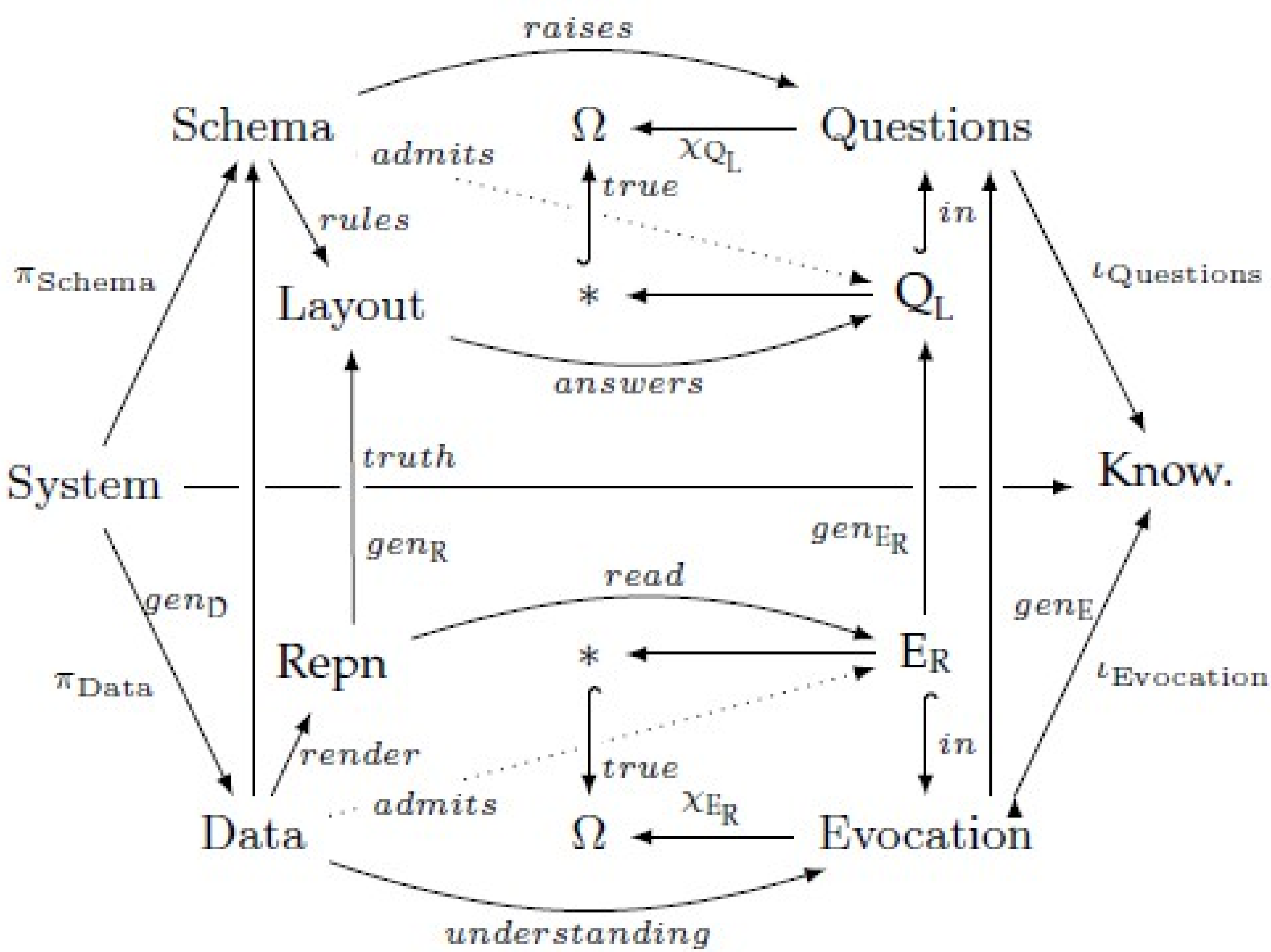


Fig. 5. The full visualization category with the subobject classifier
furnishings added.

Basically a lattice of pullbacks

Formal Topos, Paper 2

Incomplete, Qs to be added

N name, T type,
V value, E encodings
P physical representation

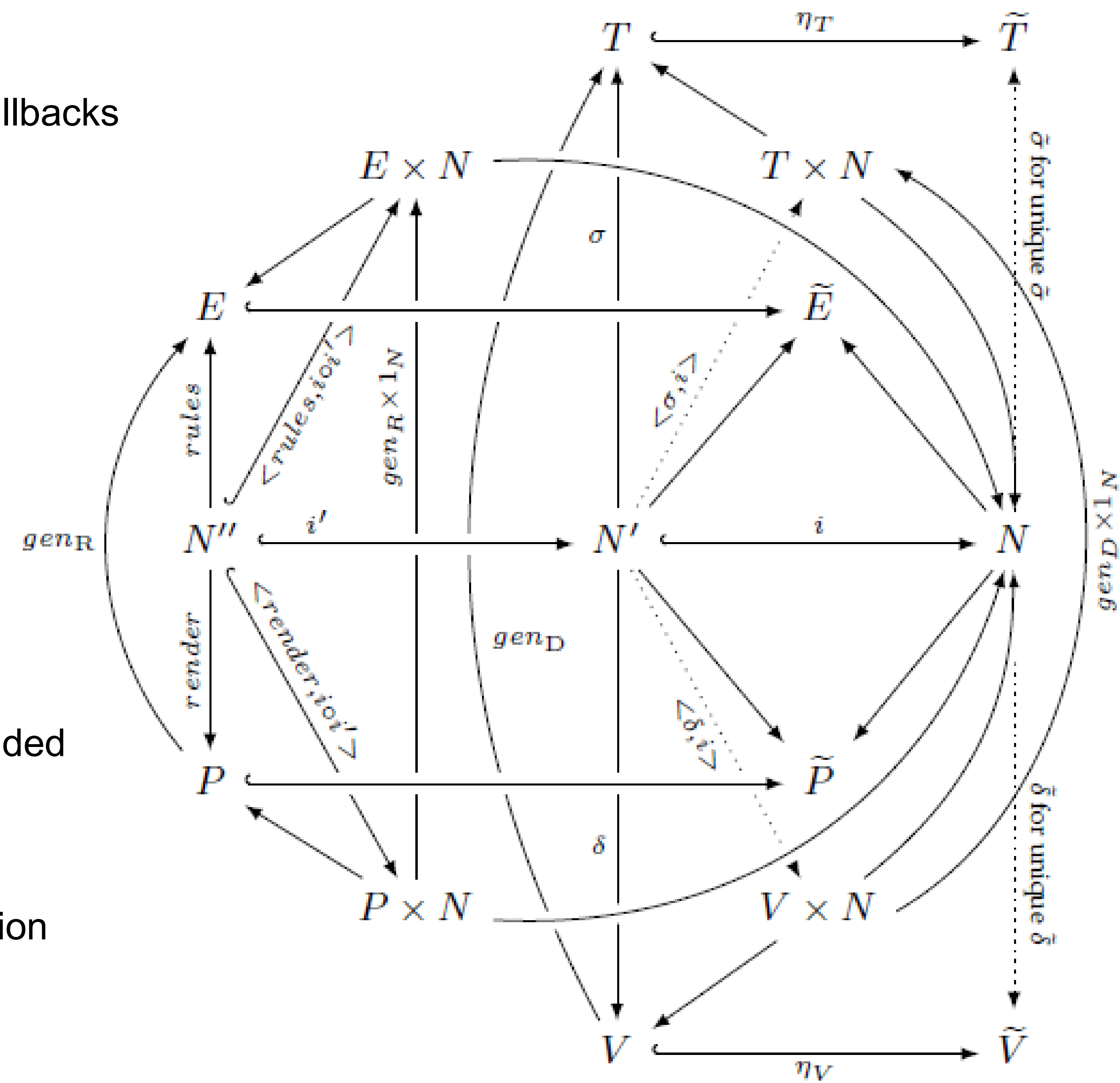


Fig. 7. Schema, Layout, Data, and Representation: The Expanded View

Final Thoughts

- Topos does look a very promising candidate for information systems
  - Ticks all the boxes for requirements
  - Well developed theoretically
- But some significant problems remain in application:
  - Need to develop an accessible way of drawing them
  - Need to develop how the internal logic will work
  - Need to make the internal logic accessible