# Towards an Ontology of Time Lapse Experiments (Cellular Genealogies).

## Report on work in progress

**Patryk Burek**[‡] , Heinrich Herre[‡], Ingo Roeder[†], Ingmar Glauche[†], Nico Scherf[†]

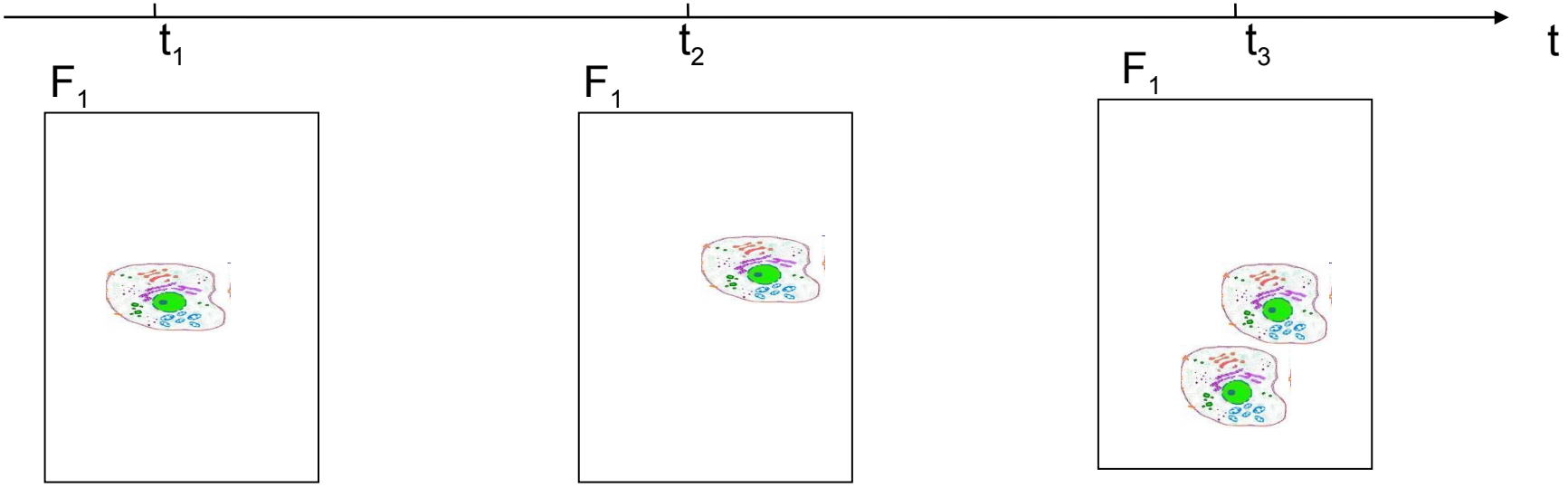
[‡] Institute for Medical Informatics, Statistics and Epidemiology (IMISE), University of Leipzig, Germany

[†] Institute for Medical Informatics and Biometry, Dresden University of Technology, Germany

# Overview

1. Preliminaries: Time Lapse Experiments, Cellular Genealogies
2. Goal and Motivation
3. Requirements
4. Architecture
5. General Formal Ontology lite (GFO-lite):
   - Main Ontological Choices
   - Selected Desing Patterns
6. Ontology for Cellular Genalogies (CGO)
7. Conclusions, Future Research

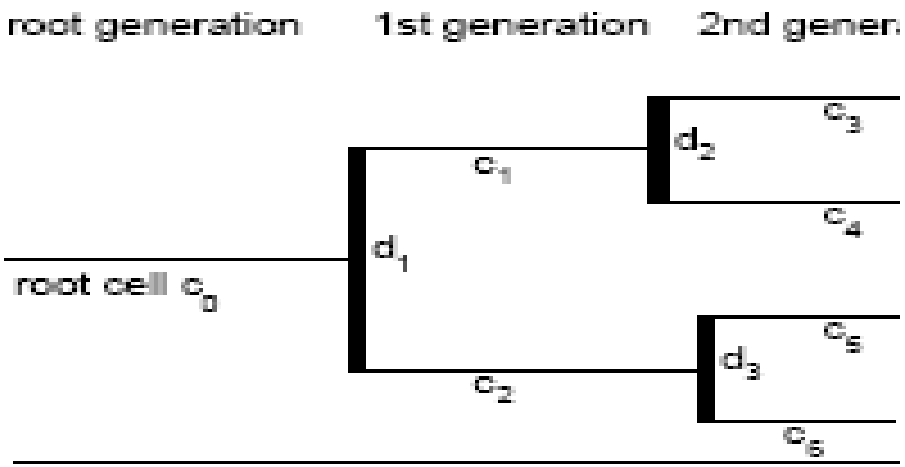
# Time Lapse Experiments and Simulations




- The application of time lapse video microscopy for the analysis of cell cultures facilitates the tracing of cells, comprising all the progeny over extended time periods up to several days.
- This includes the temporal analysis of cell specific parameters like morphology, expression of marker genes.

# Cellular Genealogies

Information on cells behavior over time (including divisions) can be comprised into a pedigree-like structure, referred to as *cellular genealogy*, in which the founder cell represents the root and the progeny is arranged in the branches.



# Goal and Motivation

Goal:

- To develop stem cell ontology and refined annotation scheme tailored for modeling and annotating results of time lapse experiments (in particular: cellular genealogies).

Motivation:

- There is no common data format for annotating, storing and exchanging data on time-lapse experiments and cellular genealogies.

# Requirements

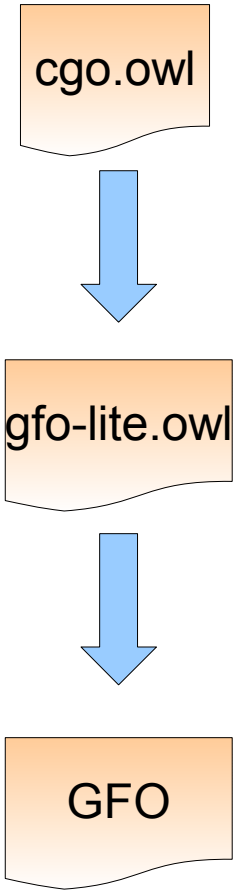
The developed ontology should be:

- *Extensible.* Time-lapse experiments and the concept of genealogies are relatively young and under dynamic development, thus the ontology and the annotation scheme must be easily extensible
- *Not ad hoc.*
- *Not implementation-biased but nevertheless formal*
- *Well-structured and founded on the solid conceptual foundations.*

  Current refactoring efforts of e.g. Gene Ontology or Cell Type Ontology demonstrate that refactoring of the deployed ontologies is a difficult and expensive enterprise. It is cheaper to build ontologies on the solid foundations from the beginning then to refactor deployed ontologies.

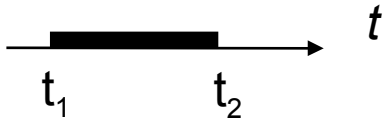
# Architecture of the provided solution

1. Ontology for Cellular Genealogies (CGO):
   - *Domain* ontology for representing cellular genealogies and annotating results of time-lapse experiments and simulations
2. General Formal Ontology *lite* (GFO lite):
   - Provides a backbone for CGO i.e. main *modeling constructs* and *design patterns* for their application.
   - *Top level* - independent of any particular domain – supports processes, objects, and functions.
   - *Independent* of any particular encoding formalism - applicable in various technical contexts (e.g. OWL, UML).
   - Makes ontological choices *explicit*, not accidental.
   - The lite version of General Formal Ontology which is a top level ontology developed by OntoMed and applied e.g. in health care process modeling

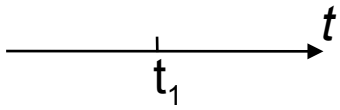


# Overview of GFO-lite: relation to time

1. *Abstracts* - entities not related to time, e.g. :
   - Mathematical objects e.g. Set
   - Property – specifications of qualities which we measure, observe or calculate, e.g. weight, color, speed
   - Property Value- volumes used in measurement, observation or calculation e.g. 10kg, green, 40$.
2. *Time-extended Entities* - entities existing through time, e.g. cup, desk, football match.




3. *Presential Entities (Presentials)* - snapshots of time-extended entities located at time points, e.g. cup at time point t1.



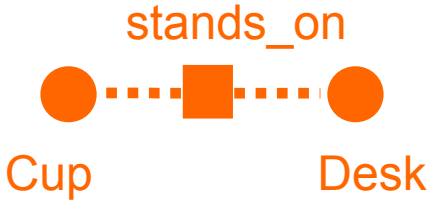
# Overview of GFO-lite: Object and Relation

*Object*:

- An entity typically perceived and distinguished from other entities as existing to some extent independently of its surrounding, e.g. desk
- It is not *of* something, thus it is independent of other entities although can be described by them.


*Relation:*

- An entity that connects other entities
- Arbitrary arity
- In contrast to an object, can not exist without its players and in this sense is dependent on them.



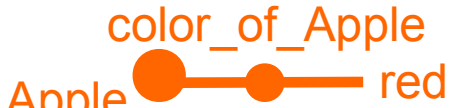
# Overview of GFO-lite: Situations and Characteristics

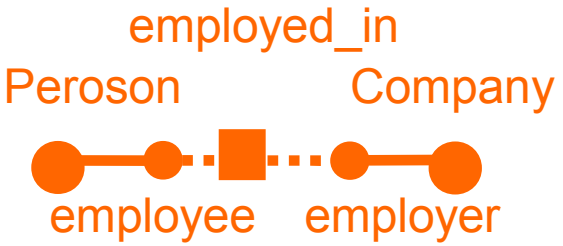
*Situation:*

- Complex Entity compressing other entities.
- Similar to an object in a sense that it can be modeled w/o references to any other entity.
- On the other hand it resembles relation in a sense that it can relate other entities.


*Characteristic:*

- entity characterizing other entities
- *Role* - is what an entity would play within the context of Relation
- *Quality* – an assignment of a Property Value to an entity.



# Matrix of main GFO-lite categories

| *Time-extended Entities* | *Presential Entities* | *Abstract Entities* | ***Graphical Symbols*** |
|---|---|---|---|
| Enduring Object | Presential Object | Abstract Object |  |
| Occurence | Presential Relation | Abstract Relation |  |
| Happening | Presential Situation | Abstract Situation |  |
| Time Extended Quality | Presential Quality | Abstract Quality |  |
| Occurence Participant | Presential Role | Abstract Role |  |
| | | Property | property |
| | | Property Value | property value |
| | | Function | |

# Selected design patterns: Relation

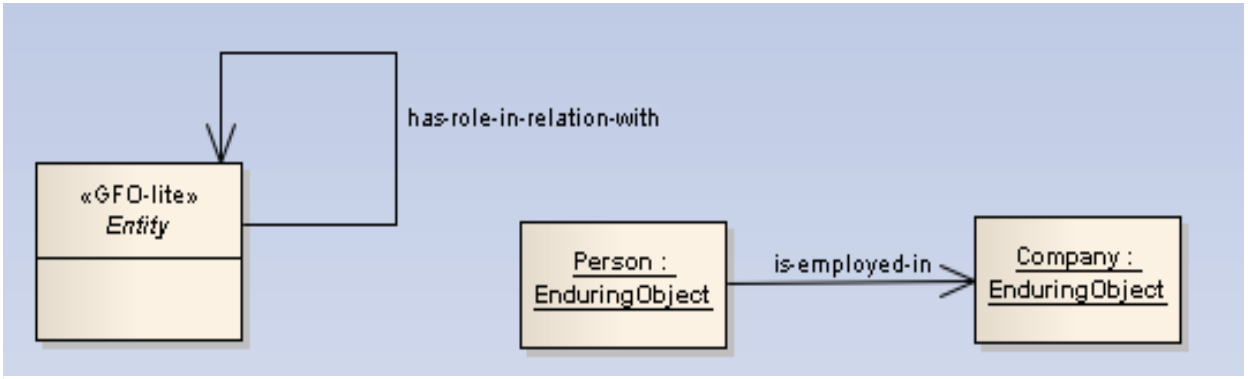
*Design Problem:* How to model relations?

*Competency Question:*

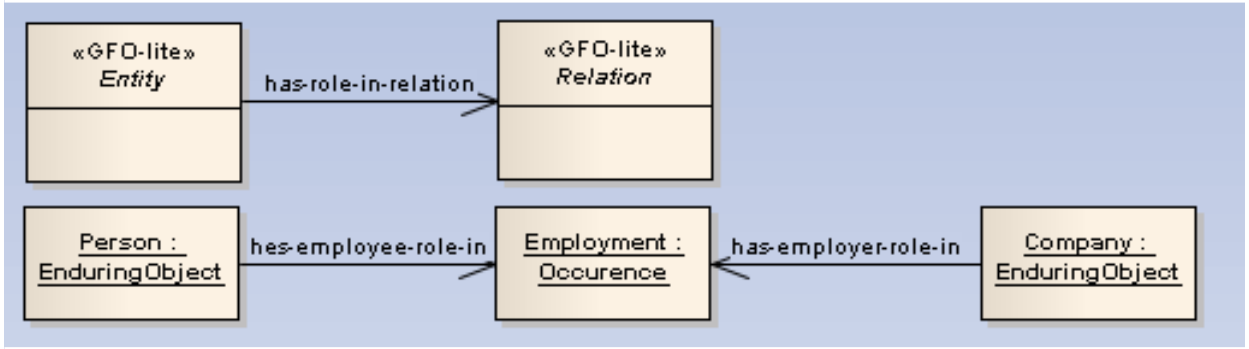
- How many players are involved in relation?
- Does a relation have its own characteristics?
- Do entities involved in relation play distinct roles in that relation?
- Do roles of the entities involved in a relation group some of their characteristics as e.g. role driver groups such characteristics as driving experience, driving license?
- In case of binary relations, is it only one-directional link between two entities or is it bi-directional link?
- Is the ontology under development considered to be developed further?

# Selected design patterns: Relation

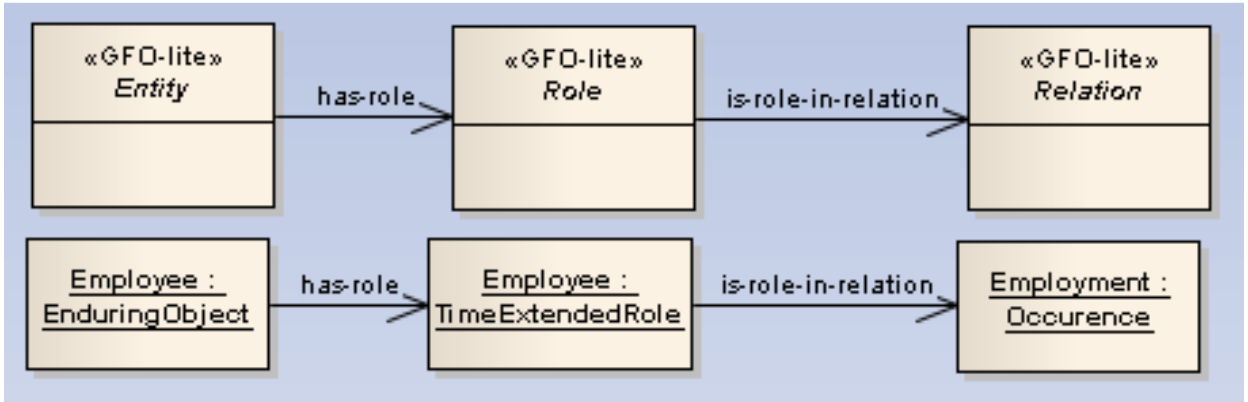
*Design Patterns:* 1)




2)



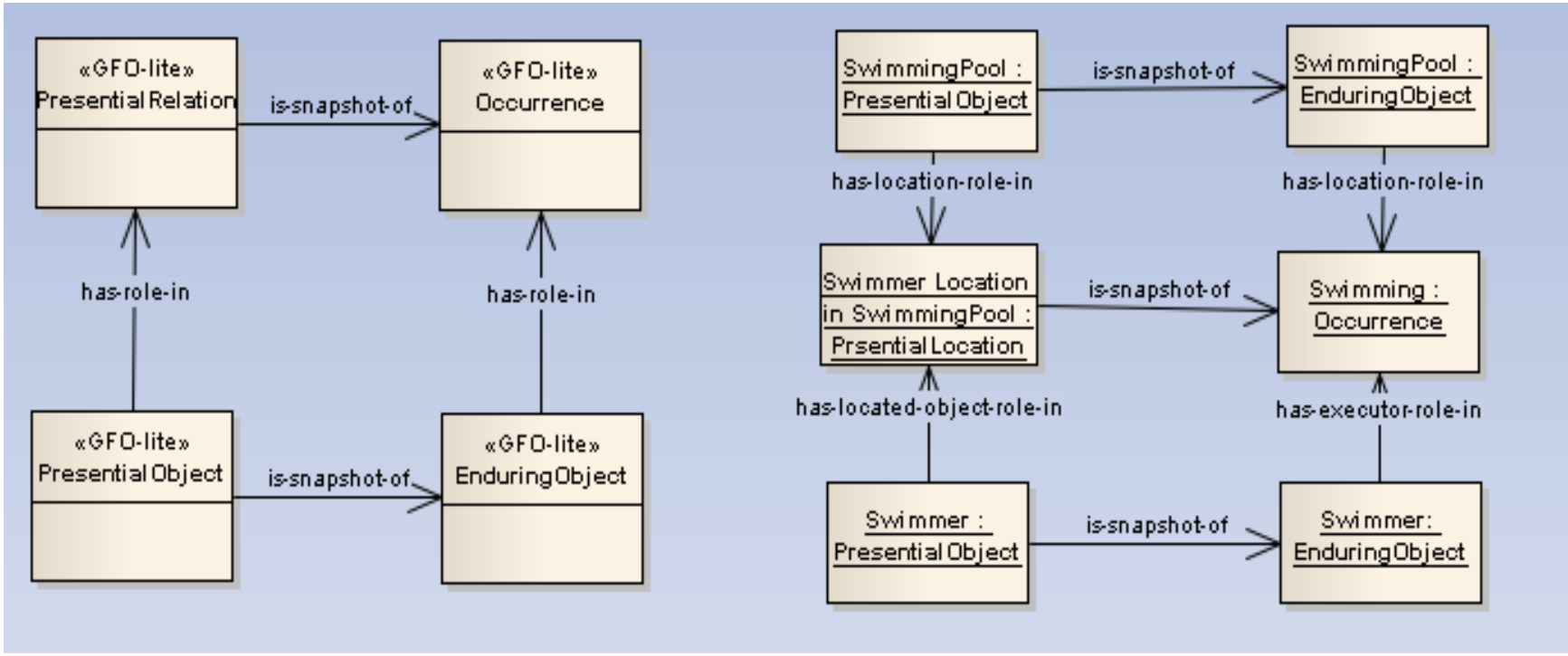

3)



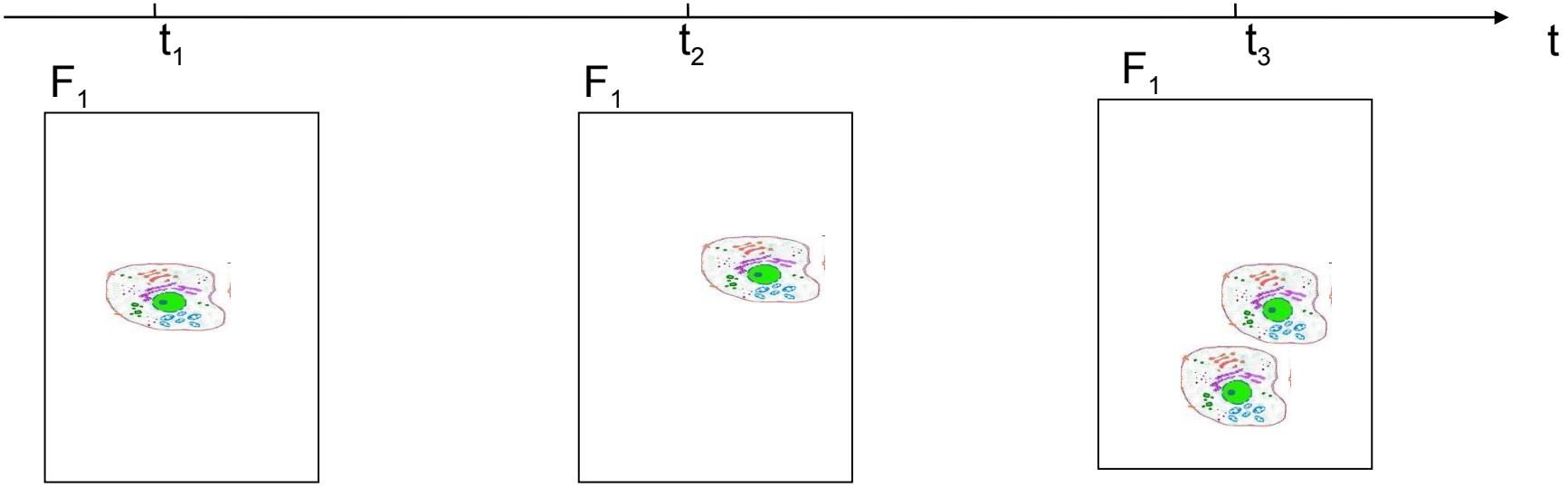
# Selected design patterns: Reified Presential

*Design Problem:* How to model interplay between enduring objects, processes / occurrences and their temporal snapshots.

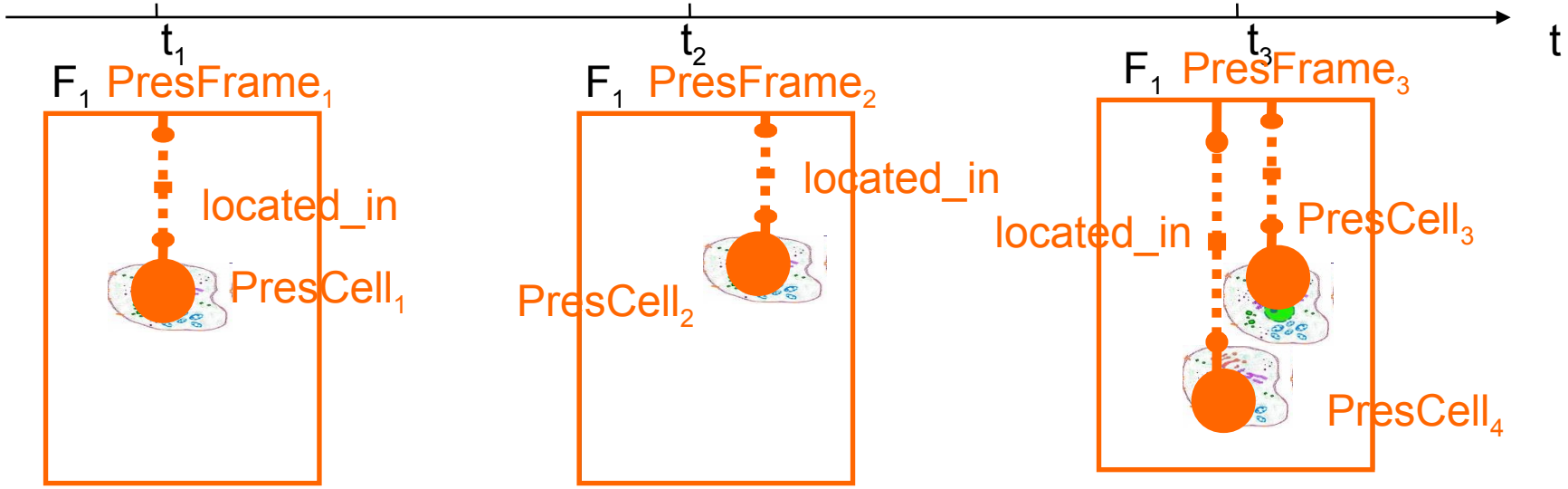
*Design Pattern:*



# Modeling with GFO-lite: Main concepts of CGO



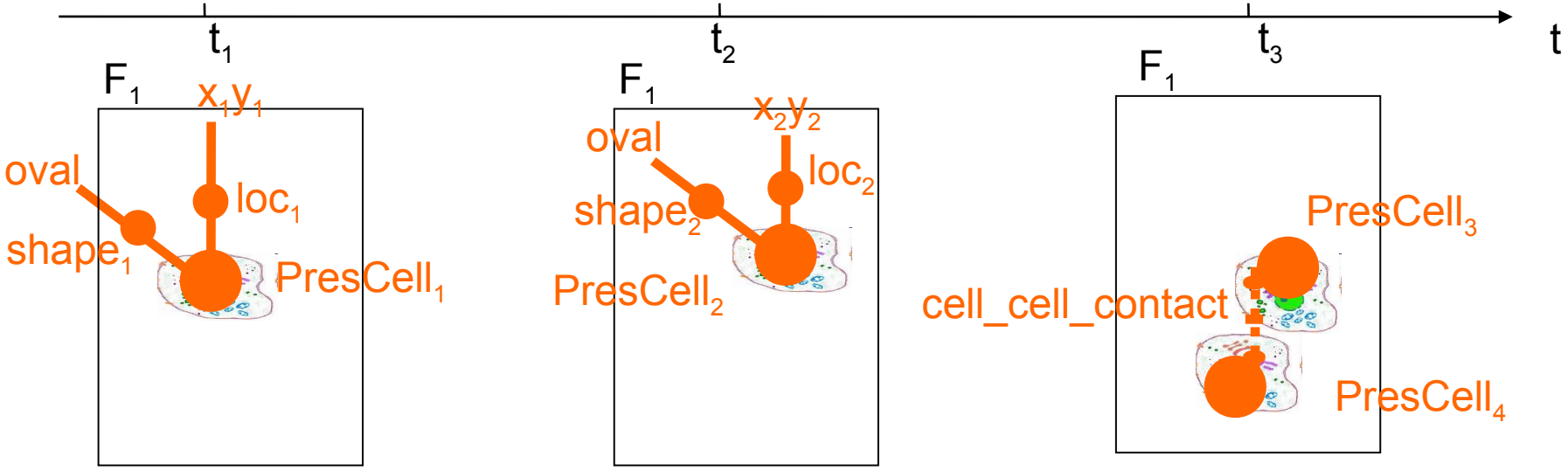
# Modeling with GFO-lite: Main concepts of CGO




Each frame is Presential Situation

Each observed cell is Presential Object located_in Frame

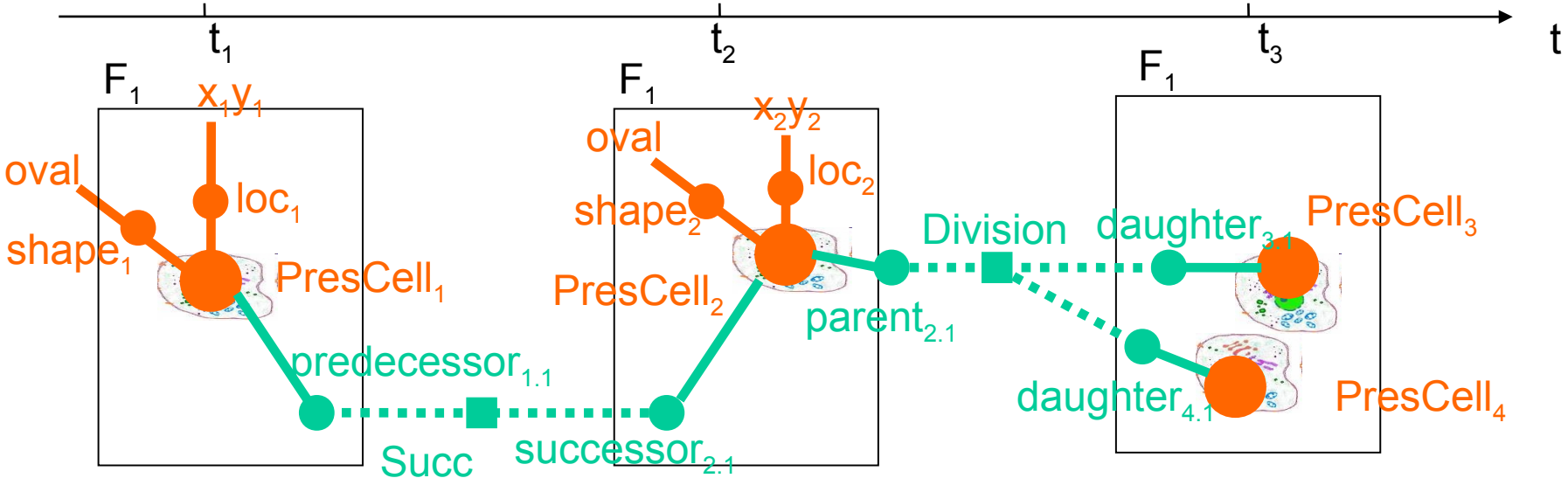
# Modeling with GFO-lite: Main concepts of GTO




Presential Cell may have Presential Qualities assigned to it e.g. *location*, *shape.*

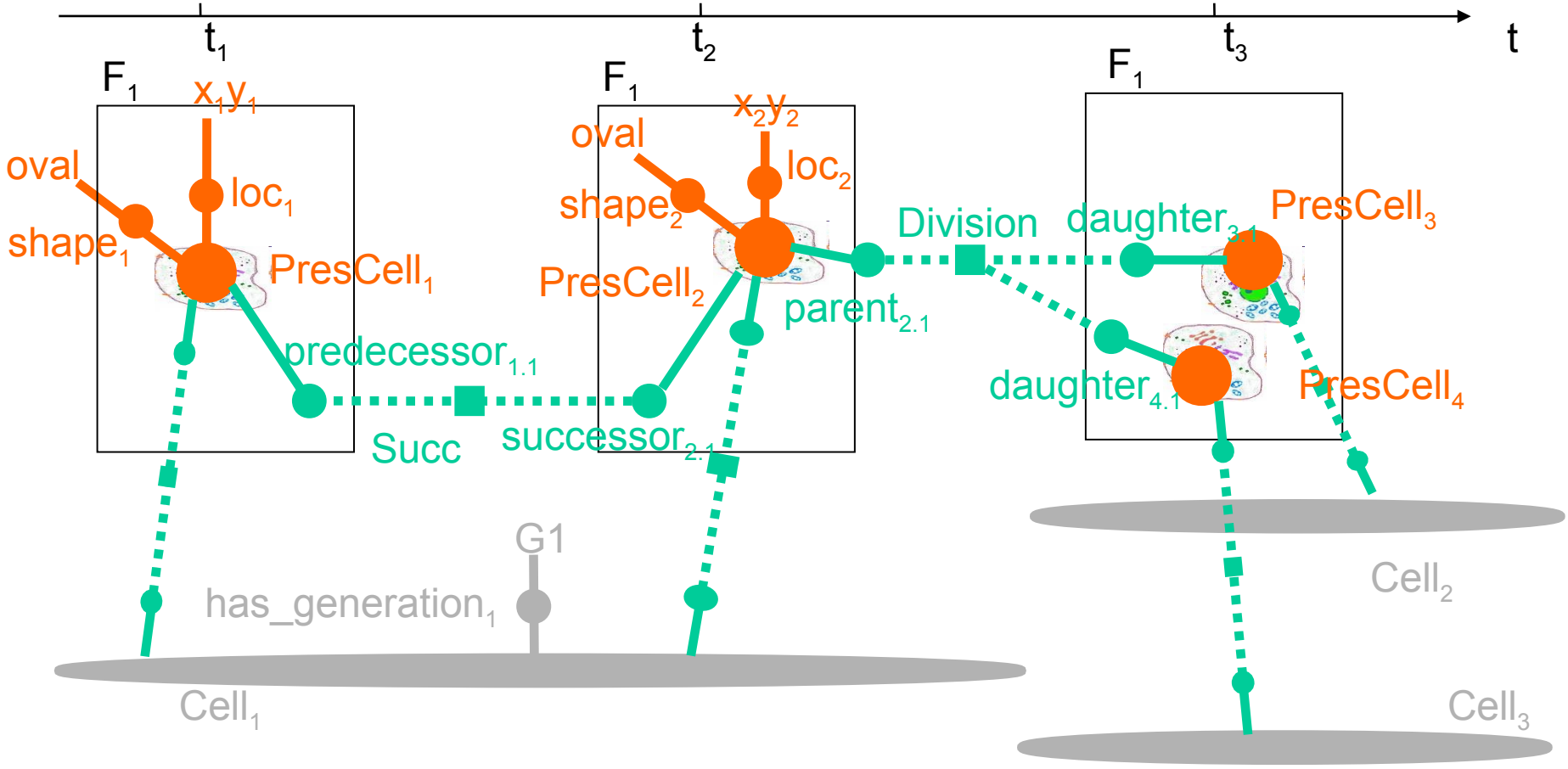
Presential Cells may be related by a Presential Relation e.g. *cell_cell_contact*, in which each cell has role *contacting_cell*.

# Modeling with GFO-lite: Main concepts of GTO




Presential Cells belonging to different Presential Frames can be related by abstract relations of *Succession* or *Division* in which they play respectively the roles of *predecessor* and *successor*, *parent* and *daughter*.

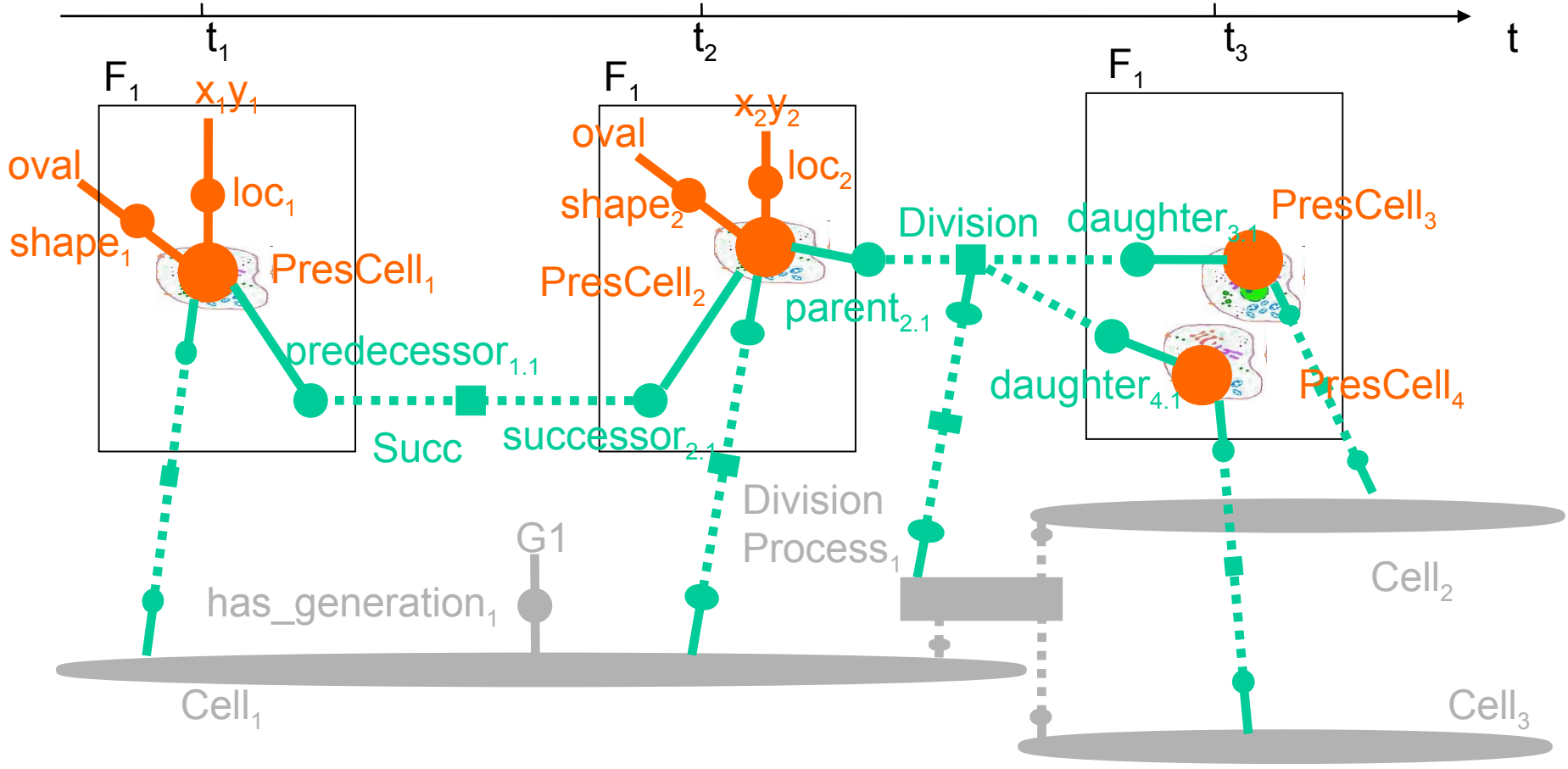
# Modeling with GFO-lite: Main concepts of GTO




*Cell* is an Object (time-extended) constructed from Presential Cells related by Succession Relation.

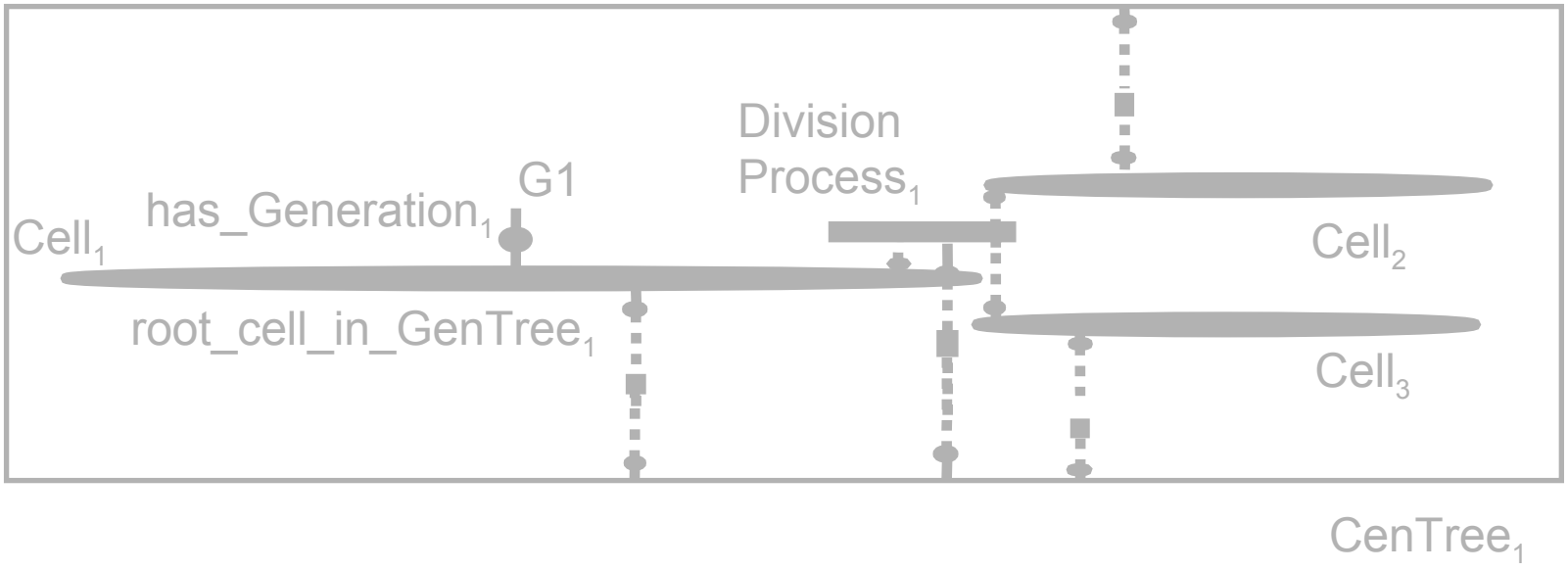
Each Cell may have Qualities assigned e.g. *has_generation*.

# Modeling with GFO-lite: Main concepts of GTO




*Division Process* is a Occurrence (Process) which is a counterpart of the presential Division Relation.

# Modeling with GFO-lite: Main concepts of GTO




*Cellular Genealogy* is a time-extended Situation (Happening) compound of one Cell playing a role of a *root_cell* in the tree and all successors and daughters of the root cell.

# Conclusions

The developed framework consist of two ontologies

1. *GTO – domain ontology:*
   - Enables describing results of time lapse experiments and simulations.
   - Handles both presential perspective on frames and cells as well as processual perspective on cellular genealogies
2. *GFO-lite - core ontology:*
   - Provides top level categories and design patterns
   - Intended to reduce time and cost of ontology construction.
   - Cross domain – can be adopted in the development of other domain ontologies
   - Provides extension mechanisms, e.g. characteristics can be added to any type of entities, n-ary relations, n>=1

# Future Research, Related Work

*Nearest future research:*

- Evaluation and extension of GFO-lite & GCO (e.g. adding the experiment context such as scientists involved, lab,..)
- Integration with existing experiment ontologies and other bio-ontologies (e.g. Cell Type Ontology, Cell Behavior)

*Related work:*

- Top level ontologies: DOLCE, Sowa's, SUMO, BFO
- Design patterns: software engineering and system modeling, W3C

Thank you!