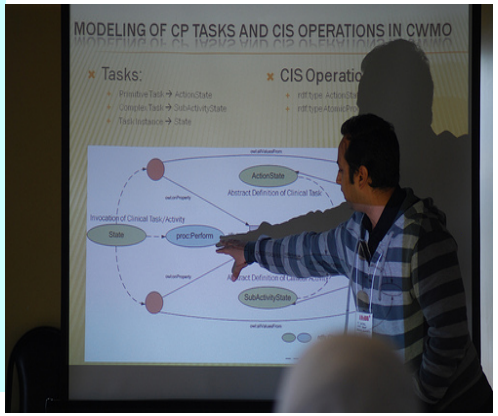
And Introduction to GeoVoCamp Processes and *Lightweight Semantic Approaches - ODPs*



Community Approach to Developing Lightweight, Modular Ontologies & Patterns

Dr. Gary Berg-Cross gbergcross@gmail.com

Spatial Ontology Community of Practice (SOCoP*)* *Socop.org*

Partially supported by NSF SOCP INTEROP Project

http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0955816

Outline of Topics

1. What is a VoCamp and how does it work?
   - Origin, Rationale and Session Goals
   - Ontology Engineering Process
     - Problems, Component and Relation Identification & Clarification
     - Conceptualization Phase
       - Systematic organization & framing with visual expression
     - Formalization
     - Illustrations
     - Tools & documentation
   - The emphasis of the steps is not to create a final, perfect ontology in a particular domain, but on creating vocabularies and then schema patterns with these that are **good enough (constrained and defined enough) for people to start using for publishing data on the Web and supporting integration**.
   - The intention is to follow a "paper first, laptops second" format, where the modelling is done initially on paper and only later committed to code.

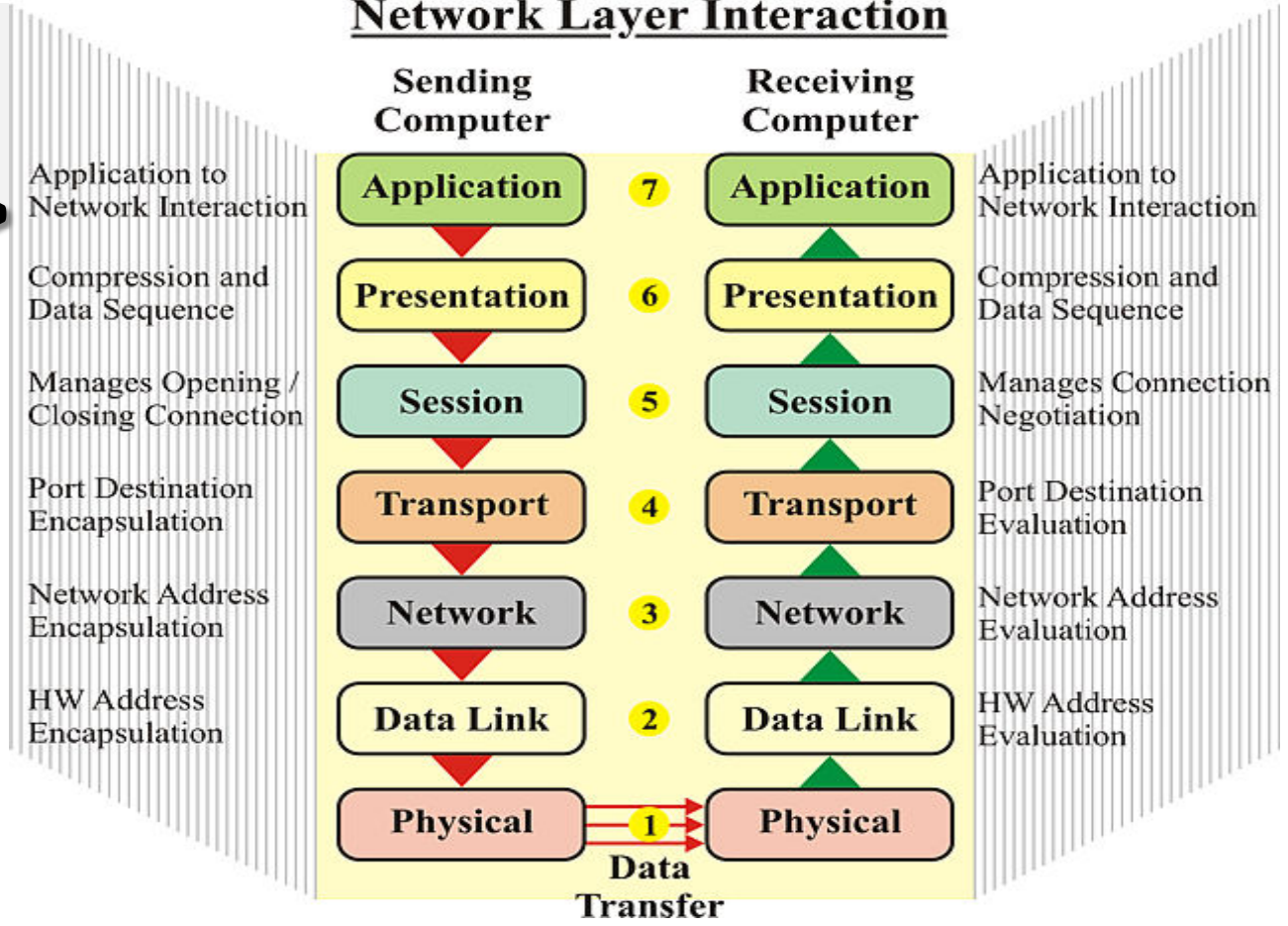
2. What is an Ontology Design Pattern (ODP) and why do this at a Vocamp?
   - Illustrations – conceptualizations, formalizations and alignment
   - Setting and Cadastral topic

Origin and Motivations for VoCamp(s): after http://vocamp.org/wiki/WhatIsVoCamp

- Growth of Semantic Web is heavily dependent on the availability of vocabularies/ontologies that can be used to publish data.
  - Some early work developed core vocabularies like FOAF that are widely used but there are many gap areas with little or no vocabulary coverage.
    - This is a barrier to progress so we need an easy, effective way to create vocabularies and if possible integrate them into schemas with appropriate constraints and relations .
- VoCamp is a series of informal events where people can spend some dedicated time creating lightweight vocabularies/ontologies for the Semantic Web & Big Data (semantic) interoperability.

Long History to SI Challenge

- Long time work on “data integration and sharing”.
  - Semantics is **FEATURED** in the Application layer of OSI
- Intensive work in the AI & knowledge engineering areas.
- More recently the Semantic Web thrust pursued the goal of robust semantic interoperability & robust exchange of data.
  - **BUT**, we need some degree of **deep knowledge** and support of reasoning to fulfill SW vision.



To many the goal of semantic interoperability remains elusive. *"i.e., the difficulty in integrating resources that were developed using different vocabularies and different perspectives on the data. To achieve semantic interoperability, systems must be able to exchange data in such a way that the **precise meaning** of the data is readily accessible and the data itself can be **translated** by any system into a form that it understands.“*

VoCamp Style is informal –an unconference
Ingredients for success include: Goals & "Group Work on Semantic Models of Interest"

Building on past workshops, VoCamps and interests:

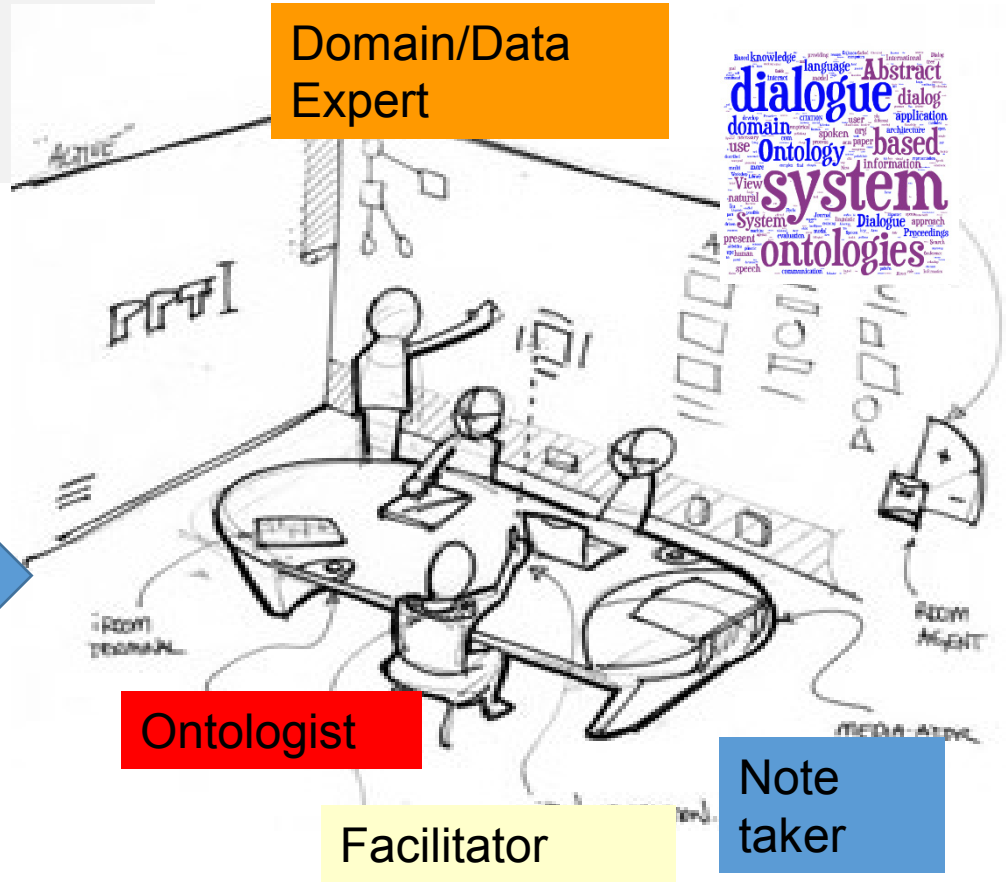
- We seek clarified agreement & reduced ambiguities/conflicts on geospatial/earth science phenomena that can be formally represented in:

1. Constrained, engineered models that represent general patterns
2. We like data-grounded work since:
   - *Much of the utility of geospatial ontologies will likely come from their ability to relate geospatial data to other kinds of information.*

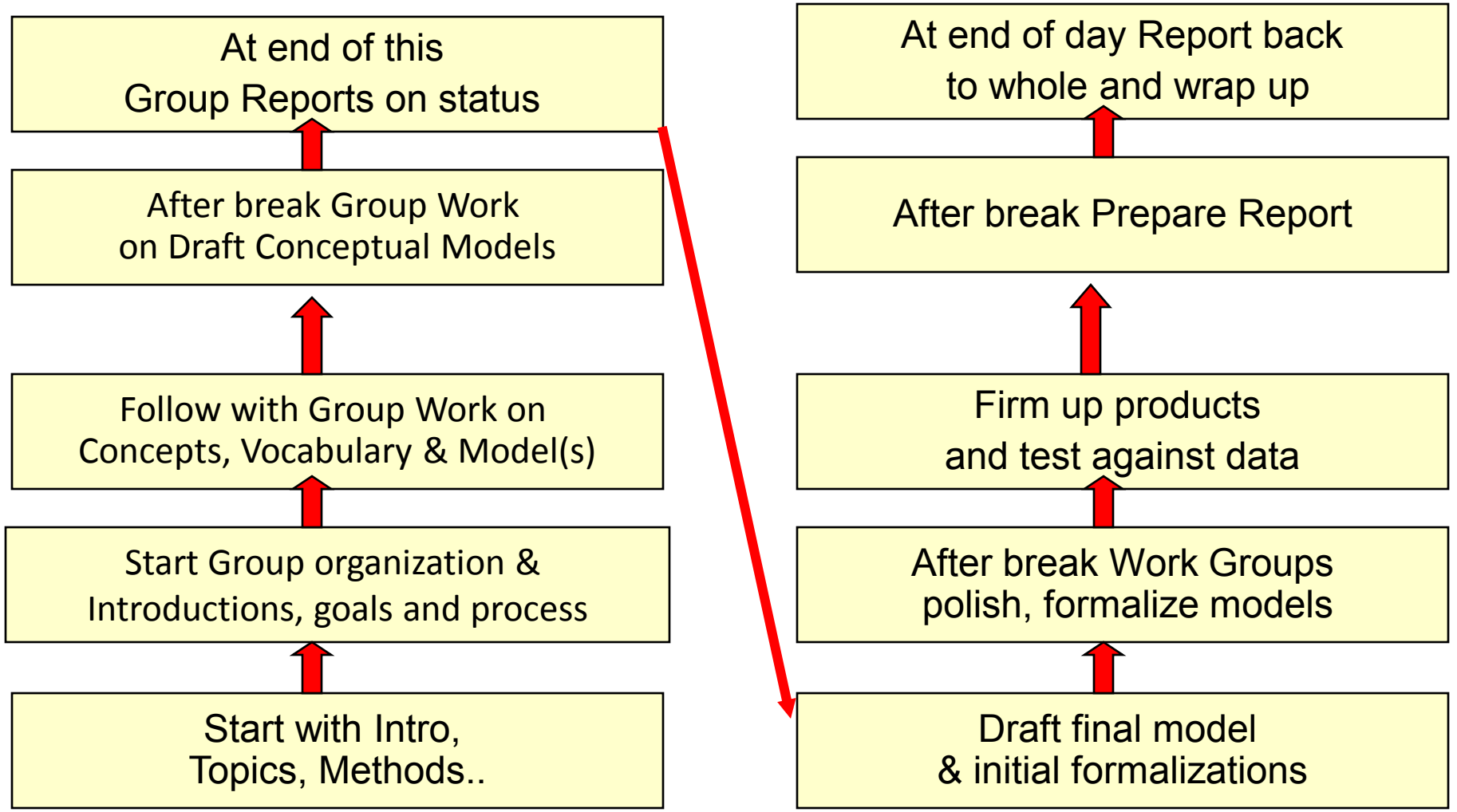
Reminder: We are not here to teach a new discipline, but to employ multiple disciplines as a team.

More on Ingredients for success:

1. <u>Practical things to Organize a VoCamp</u>
    1. Group Goals, advertise/market the meeting
    2. Pick a convenient location for core group
    3. Check timing for conflicts
    4. Recruit a critical mass of talent
2. But we have more <u>technical aspects</u>:
    1. 2-4 Workgroup Teams with topic leaders & mixed talents,
        1. People select what they want and presentations are minimal
    2. Phased Session Structured for small Socio-Tech Work Sessions
        1. From Conceptualizations to Formalizations
    3. Lightweight Methods for Ontologies and Ontology Patterns



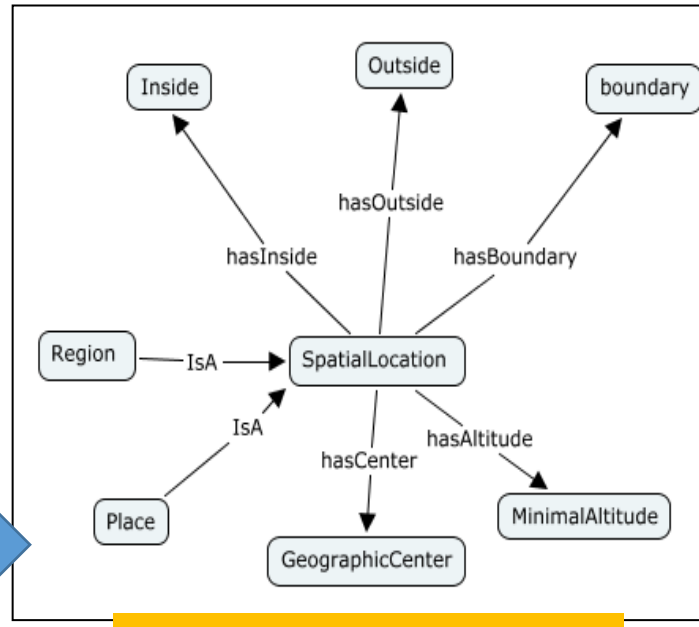
Logic of Phased Work Sessions (Usually 2 days)



More Detail on Process

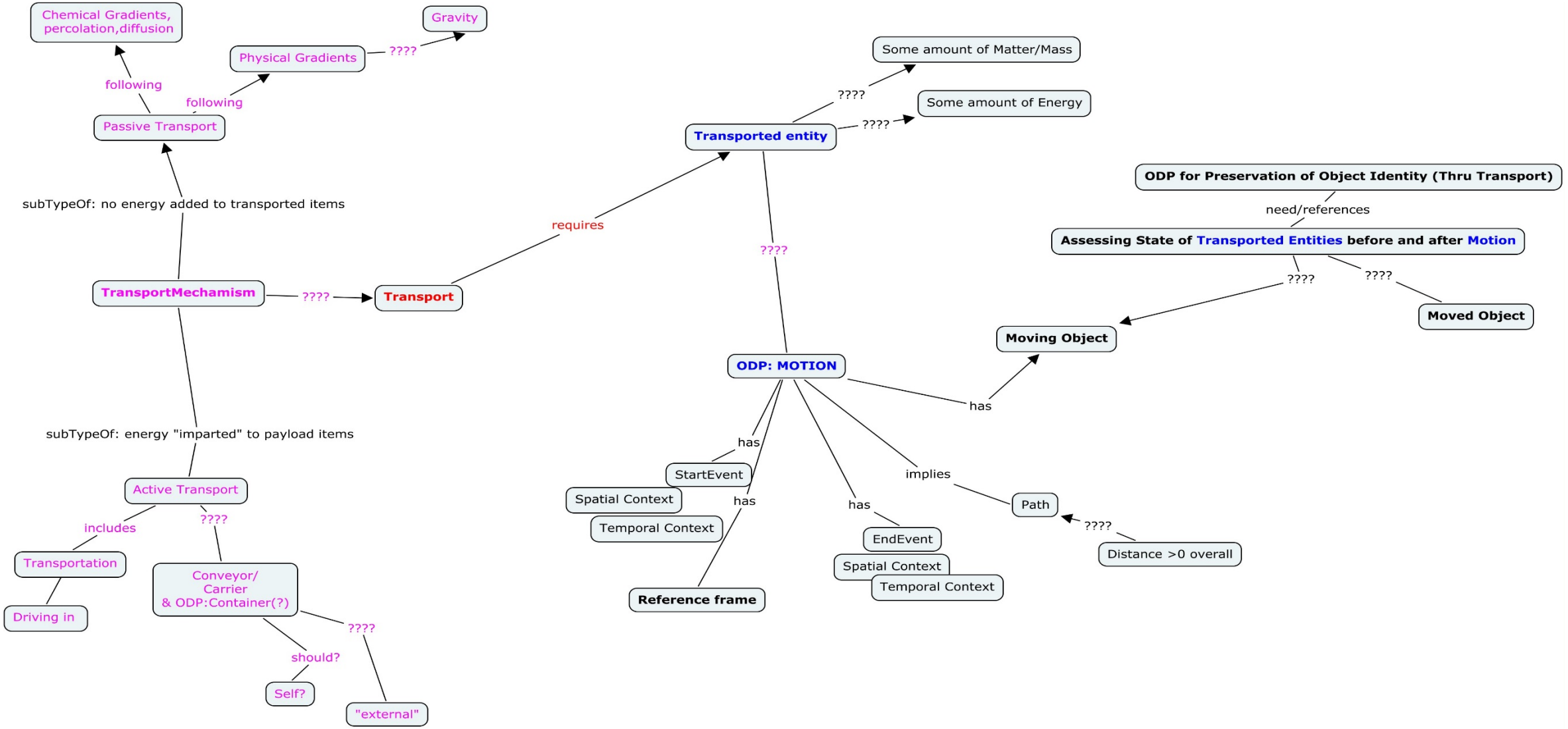
1. Start with Problems, Component and Relation Identification & Clarification (~1-2 hrs)
2. Ontology Engineering starts with Conceptualization Phase ~ (half a day)
   - Group work on vocabulary, systematic organization & framing with visual expression
   - Constrained, engineered models with commitments to support understanding, reasoning & data interoperability and/or
   - Creation of light, general ontology patterns that provide a common framework to generate ontologies that are consistent and can support interoperability.
   - Move from Conceptualizations to Formalizations
     1. Cautions and Tradeoffs
        1. No single taxonomy satisfies all needs.
     2. Ground patterns with data
     3. Document work

Repeat



Imperfect, but a step?

Don’t be Afraid if Work in Progress is Sketchy



More on Step 1: Acquire Scoped Domain Knowledge & Vocabulary Principle: Clarify & Identify Problem, Components & Relations

- Identify scope: We prefer working from scenario examples with potential data to help structure requirements, defining the purpose of the ontology and illustrate the nature of a problem topic.
  - **What are we talking about**? What do you mean when you use these words..
    - **Streams as objects (not processes?)**…..
  - What is the **scope**? Not how streams flood, or pools of water,…
  - What is the **purpose** of this modeling? What questions are addressed?
  - What data is relevant….
- We should **leverage** existing work but not slavishly
  - leverage thoughts & experiences from other groups that are collaborating on ontologies
  - reference or include supporting vocabularies/ontologies,
- Terminologies can be a **starting point**, but the path should be to the **concepts** behind what the terms mean to domain people and be relatable to "data."
  - Controlled vocabularies and other terminologies are necessary lexical resources to refer to concepts
    - Linguistics or conceptual analysis
    - Handle multiple meaning and similarities

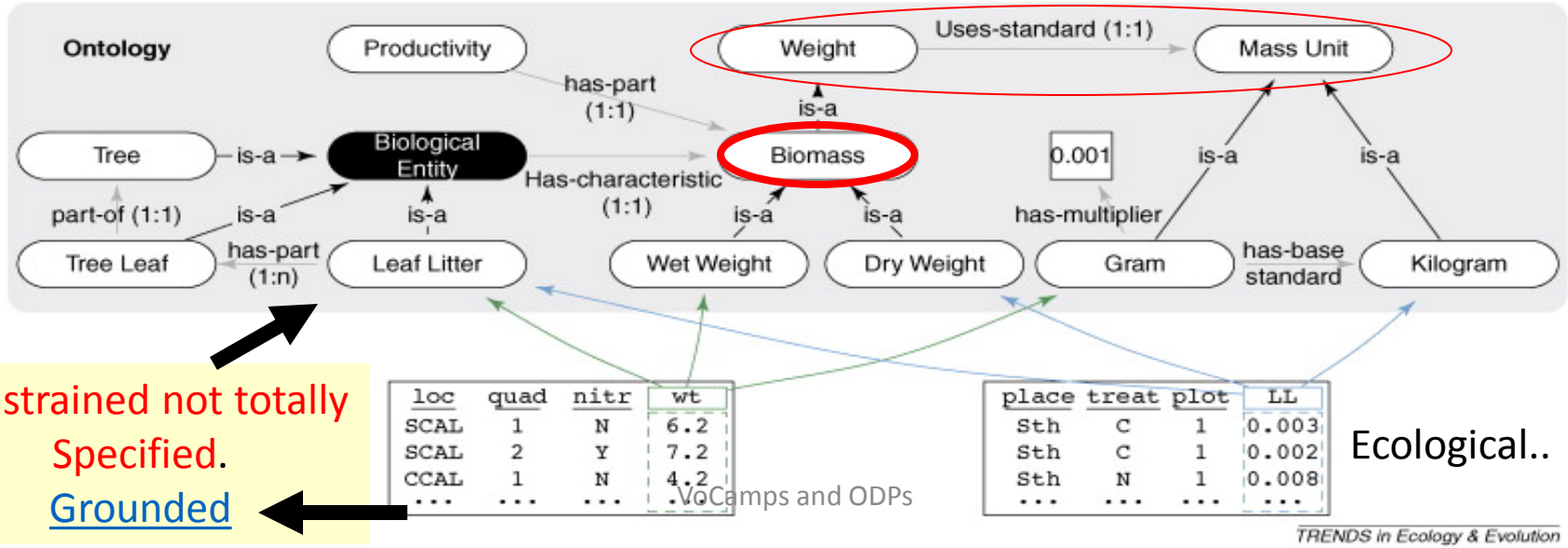
Illustrated somewhat with Setting and Cadastral examples.

Lightweight Methods & Products

- Choose lightweight approaches grounded by scenarios and application needs.
  - Low hanging fruit **leverages initial vocabularies** and existing **conceptual models** to ensure that a semantics-driven infrastructure is available for use in **early stages of work**
  - Reduced entry barrier for domain scientists to contribute data

Simple parts/patterns & direct relations to data

Triple like parts



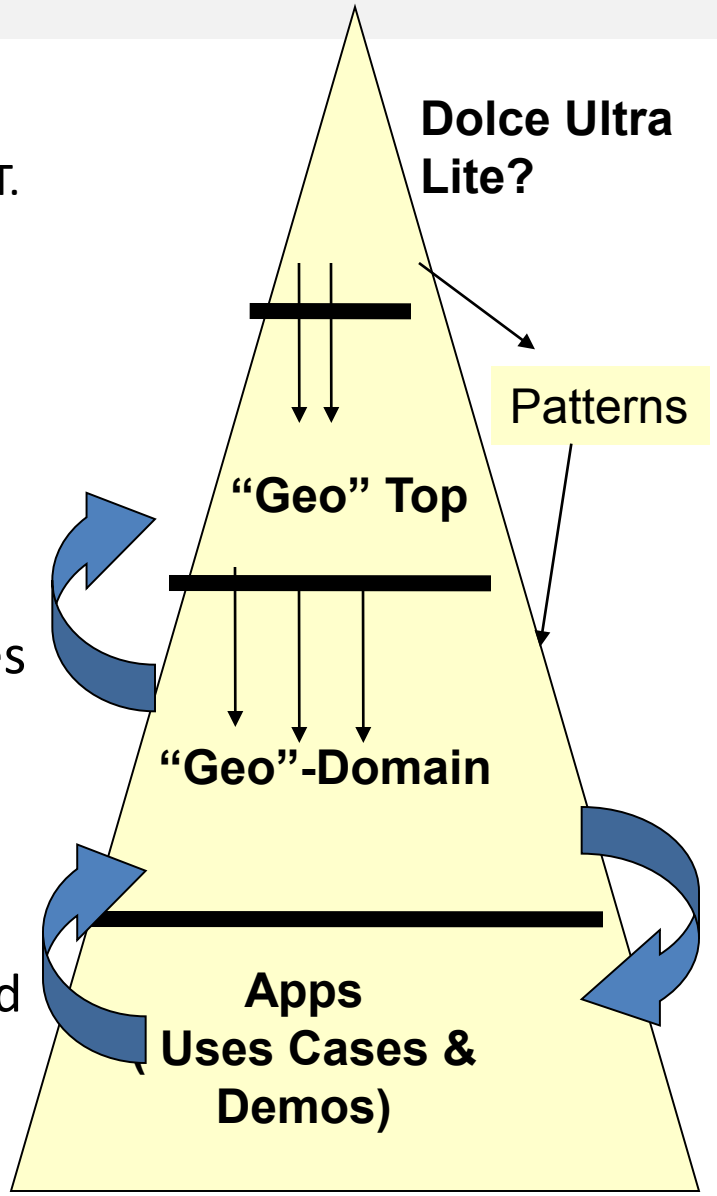
Constrained not totally Specified.
Grounded

Ecological..

Ontologies and Ontology Design Patterns (ODPs)– what and why?

Many levels & types of ontologies

- Something like DOLCE is quite complex as are domain ontologies like SWEET.
- **Problem**
- It is hard to reuse only the “useful pieces" of a comprehensive (foundational) ontology (hard to find), and
  - the cost of reuse may be higher than developing a scoped ontology for particular purpose from scratch
- **Solution Approach**
  - Use small, well engineered, modular starter set ontologies with
    - explicit documentation of design rationales,
    - Minimal Constraints for generalizability and
    - best reengineering practices Formalize these after domain expert agreement
  - **Reuse** – The hope is that these can be built out and combined with other ODP modules to address more complex topics.



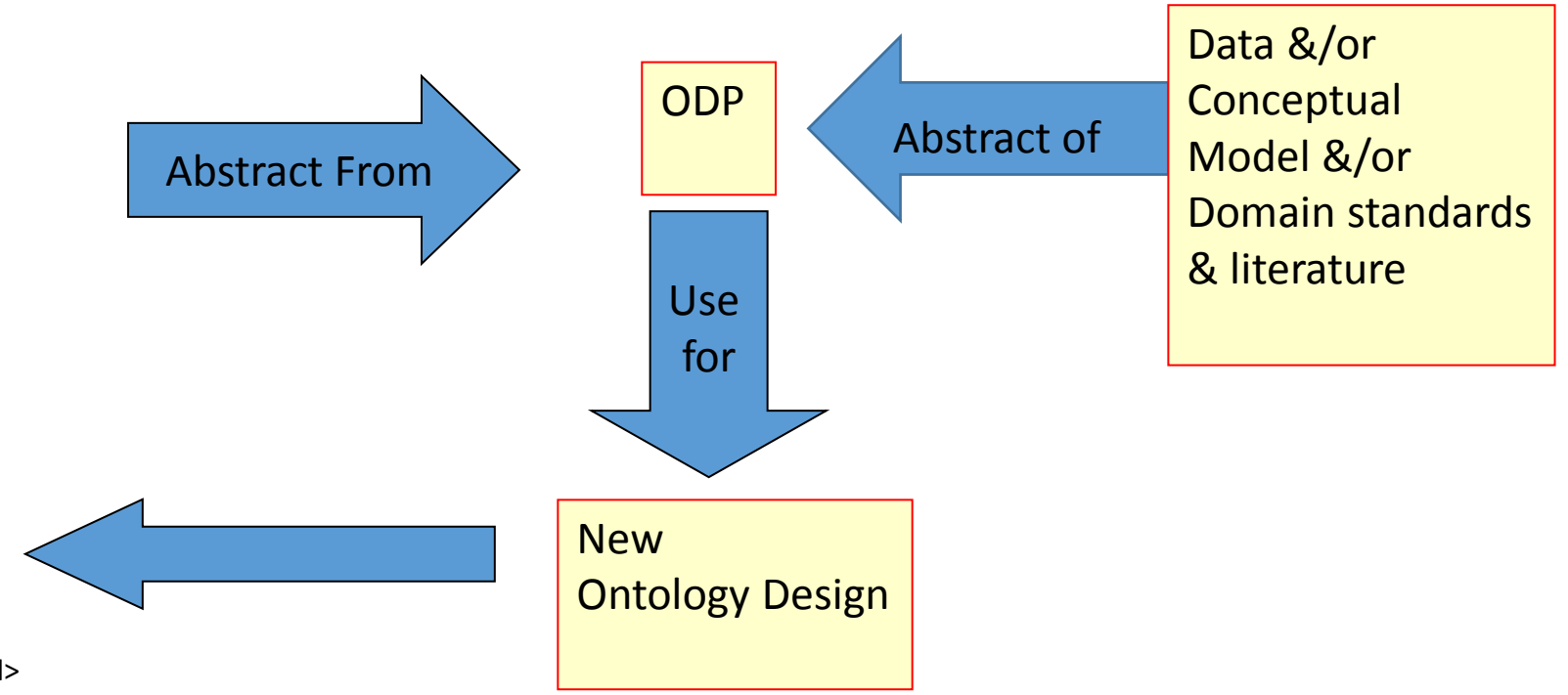
More on ODP Rationale

- "For solving semantic problems, it may be more productive to agree on minimal requirements imposed on .. Notion(s)"
  - Werner Kuhn (**Semantic Engineering, 2009)**
- ODPs, like OWL, are tools for ontologies
  1. They are more easily understandable with good explicit documentation for design rationales
  2. Robust, expressive, manageable and flexible (illustrated later)
  3. Can be used to build on modularly for reoccurring problems needing representation
  4. Capture best practices
  5. Should help bridging/integrating ontologies
  - These serve as an initial constraining network of "concepts" with vocabulary which people may build on/from for various purposes.

Illustration of Ontology-ODP Relations

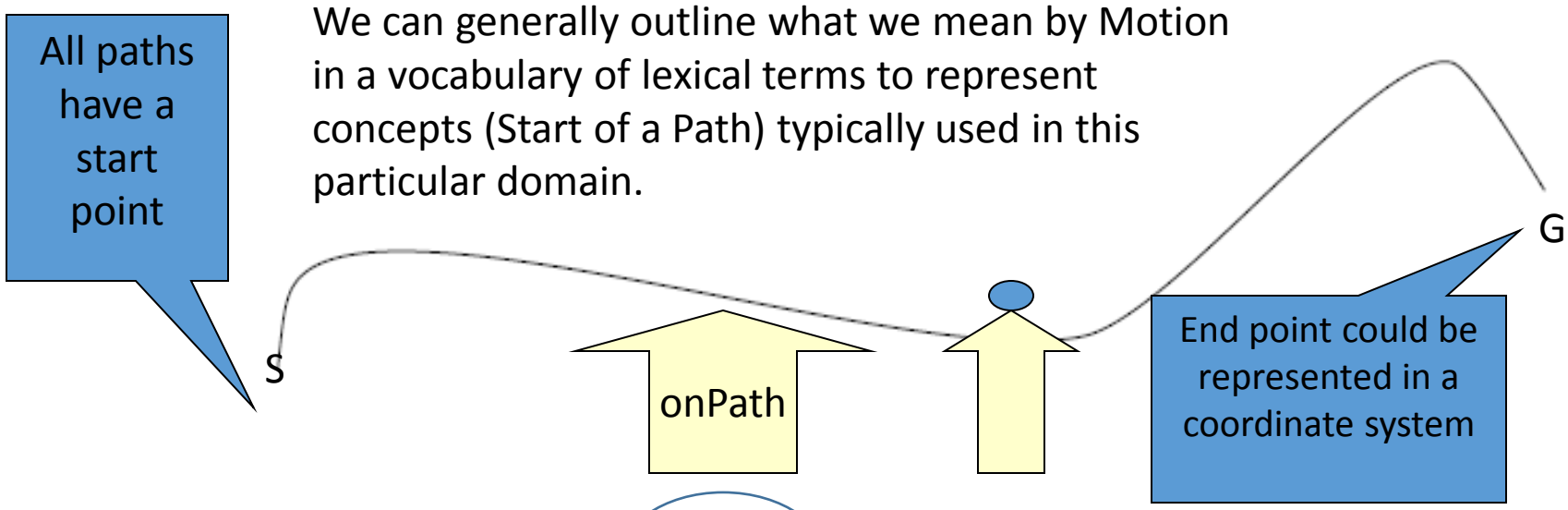
Small DUL Portion

- <owl:Class rdf:ID="SocialObjectAttribute">
- <rdfs:label xml:lang="en">Social attribute</rdfs:label>
- <rdfs:subClassOf>
- <owl:Restriction>
- <owl:onProperty>
- <owl:ObjectProperty rdf:ID="isRegionFor"/>
- </owl:onProperty>
- <owl:allValuesFrom>
- <owl:Class rdf:ID="SocialObject"/>
- </owl:allValuesFrom>
- </owl:Restriction>
- </rdfs:subClassOf>
- <rdfs:label xml:lang="it">Caratteristica sociale</rdfs:label>
- <rdfs:subClassOf>
- <owl:Class rdf:ID="Region"/>
- </rdfs:subClassOf>
- <rdfs:comment>Any Region in a dimensional space that is used to represent some characteristic of a SocialObject, e.g. judgment values, social scalars, statistical attributes over a collection of entities, etc.</rdfs:comment>
- </owl:Class>
- <owl:Class rdf:ID="WorkflowExecution">…



"Unfriendly logical structures, some large, hardly comprehensible Ontologies" (Aldo Gangemi)

Conceptual Pattern Illustration- A Schema for Motion:



- participants: path, moving object, start, goal

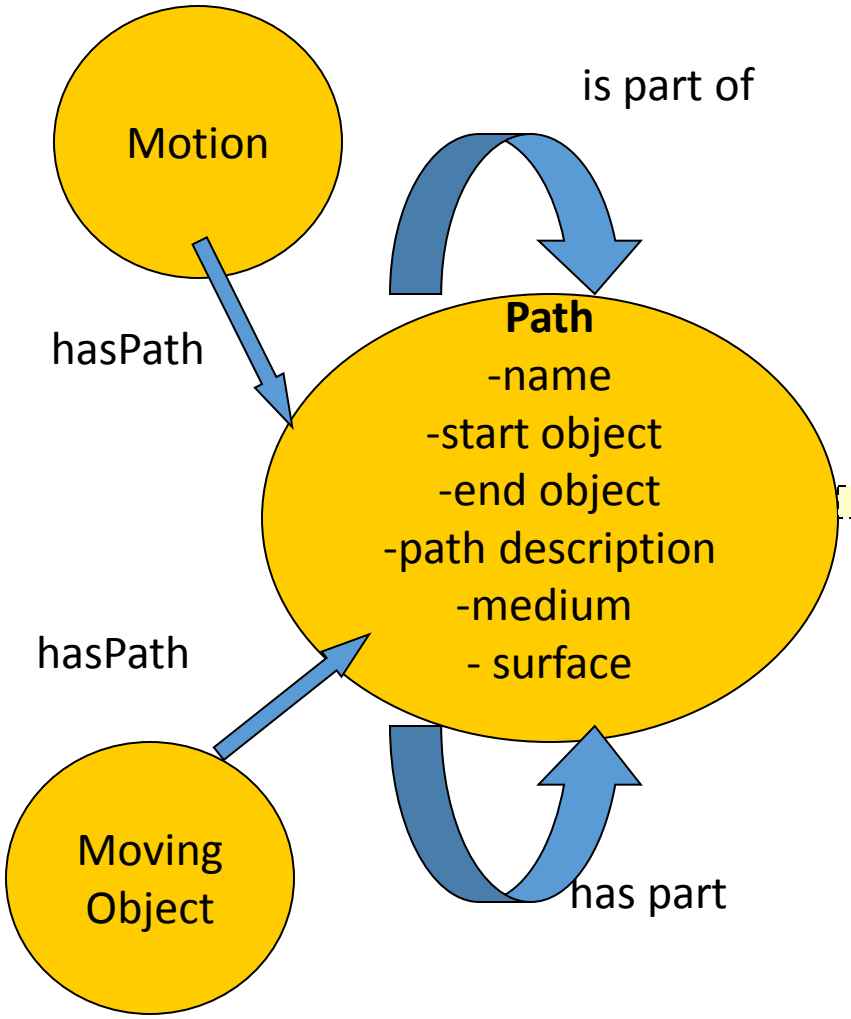
We remain general in the pattern since this is a cognitive activity & the concept has flexible semantics depending on human intentions and perspectives. The pattern can generate alternate descriptions conforming to alternate interpretations.

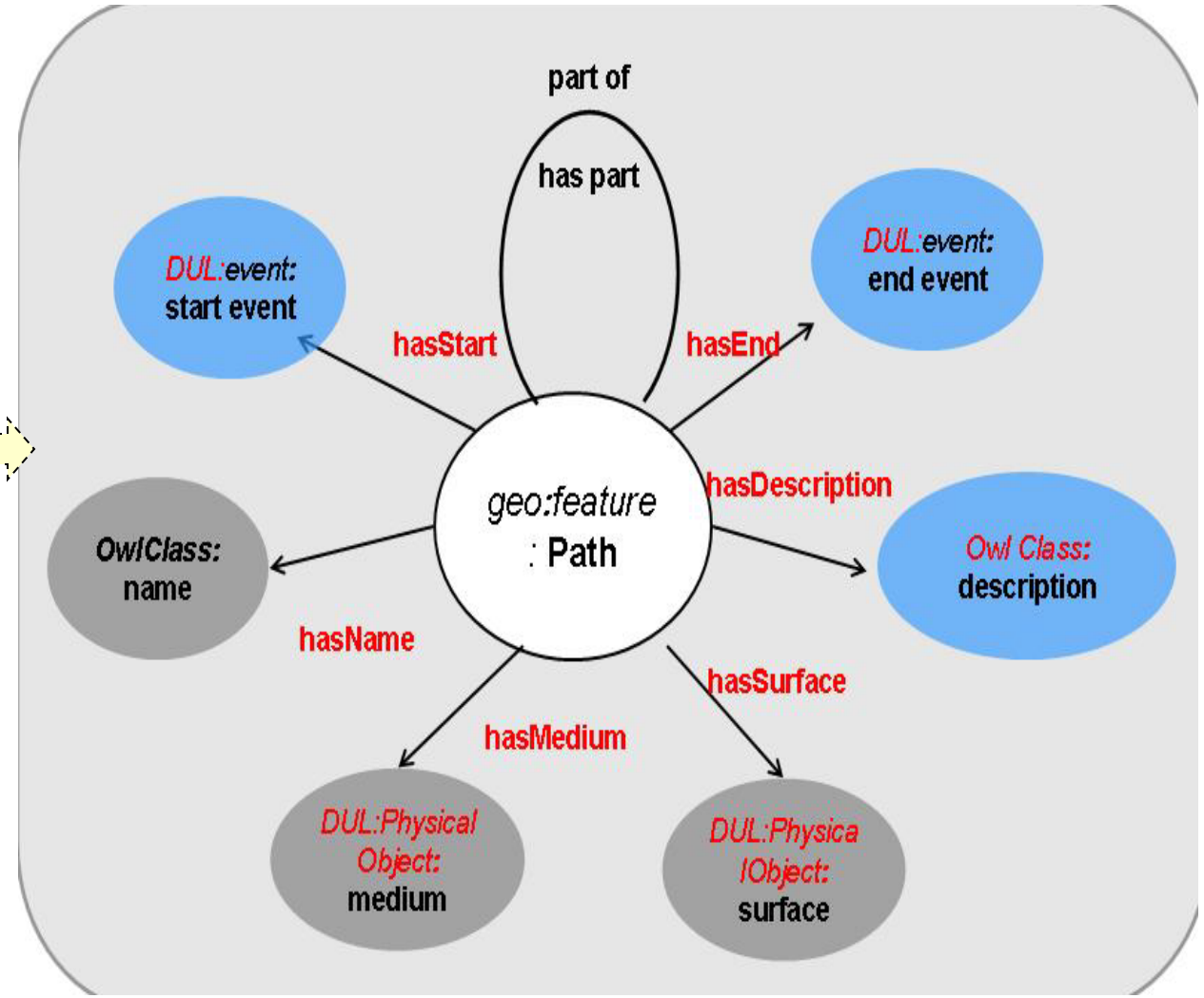
For details, see:

Kuhn, W., 2007. An Image-Schematic Account of Spatial Categories. *Spatial Information Theory, 8th International Conference, COSIT 2007*. Melbourne, Australia: Springer Lecture Notes in Computer Science 4736: 152-168

Just OWL Classes
Motion is an OWL:Class

Geo-VoCamp Pattern – Path from an info perspective





Light constraints by relations and what is related

Formally Committing to our Conceptualization with Axiomatized Semantics

1. Formal –can be represented/put into a form amenable to automated processing [**formal language**]
   - Ontologies formalize concepts with axioms defined on such concept vocabularies
   1. Sufficiently axiomatized – include detailed constraining descriptions, such as transitivity, as axioms (not just text descriptions)
2. Rigorous – stands up to rational analysis
   1. Distinguish which concepts have instances
      1. Named classes can (potentially) have instances

Motion / Path RDF/ontology in Turtle (TTL): (Terse RDF Triple Language – uses a . , ;])

Namespace prefixes

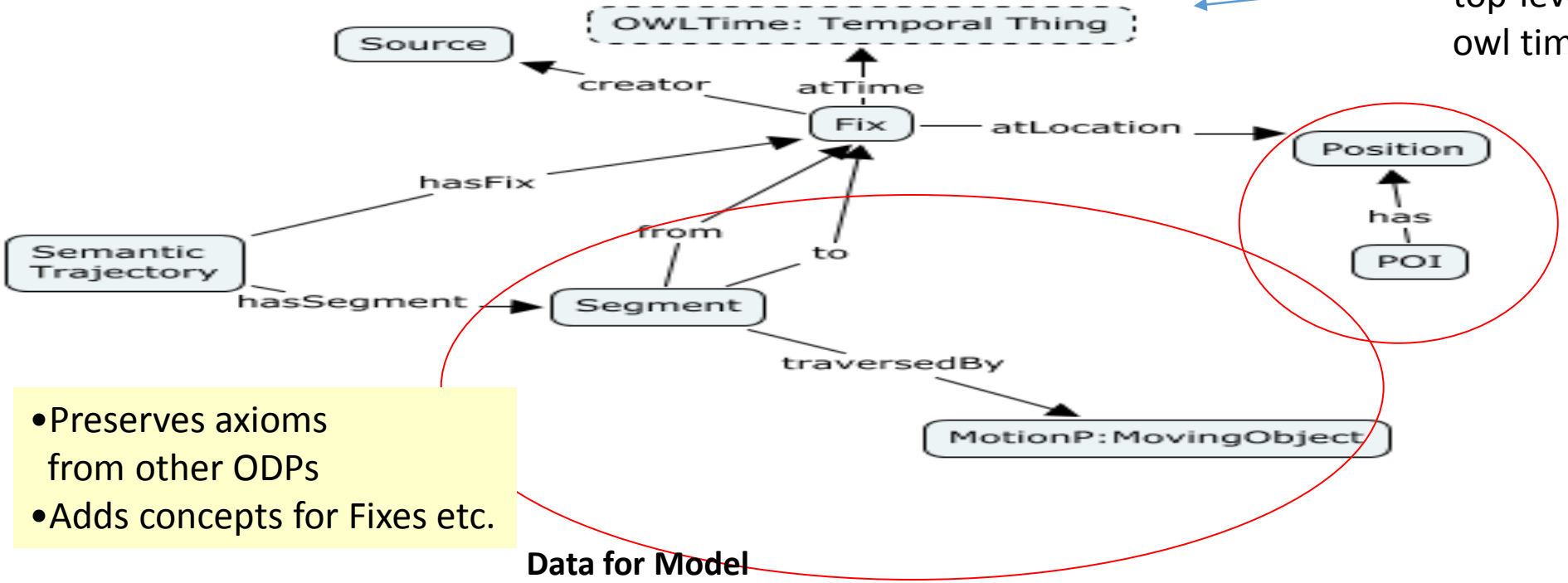
```
@prefix rdf:  <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix xsd:  <http://www.w3.org/2001/XMLSchema#> .
@prefix geo: <http://www.opengis.net/def/geosparql/> .
@prefix sf: <http://www.opengis.net/def/sf/> .
@prefix gml: <http://www.opengis.net/def/gml/> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix spw:
    <http://www.w3.org/2001/sw/BestPractices/OEP/SimplePartWhole/part.owl> .
@prefix event: <> .   @prefix : <http://vocamp.org/ontology/movement/spatial/> .
    (Default prefix)

# Ontology description    :Ontology a owl:Ontology;
    owl:versionInfo "0.5";
    rdfs:comment “A geospatial instantation of the generic Movement & Path
    pattern created at GeoVoCampSB2012.";
    dc:title "Spatial Motions and Paths" .
```

A restriction class should have exactly one triple linking the restriction to a particular property, using **owl:onProperty**.

```
# Motion class
rdfs:comment "A motion is an event in which some entity moves through
    space" ;    rdfs:subClassOf [
                a owl:Restriction;
                owl:onProperty :startEvent;
                owl:allValuesFrom event:Event
    ];
:Motion a owl:Class;
    rdfs:subClassOf event:Event;
    rdfs:label "Motion";
rdfs:subClassOf [
                a owl:Restriction;
                owl:onProperty :endEvent;
                owl:allValuesFrom event:Event
    ];
    rdfs:subClassOf [
                a owl:Restriction;
                owl:onProperty :path;
                owl:allValuesFrom :Path
    ];
    rdfs:subClassOf [
                a owl:Restriction;
                owl:onProperty spw:hasPart;
                owl:allValuesFrom :Motion
    ] ;
    rdfs:subClassOf [
                a owl:Restriction;
                owl:onProperty :startEvent;
```

Composing New ODP from Old: Semantic Trajectory

Option of Aligning to top-level Concepts to owl time



- Preserves axioms from other ODPs
- Adds concepts for Fixes etc.

Data for Model

:mikestrip a :SemanticTrajectory; :hasSegment [a :Segment; :from :fix1; // mikeshome:to :fix2;// rest stop :traversedBy :fordFocus], [a :Segment; :from :fix1; // rest stop :to :fix2],// WrightStateU :traversedBy :fordFocus], [a :Segment;:from :fix1; // WrightStateUniversity:to:fix2],//.. :fixn].:mike a foaf:Person:mikesFordFocus a motion:MovingObject.:garminEtrexVistaC a:Source.geo: Geometry rdfs:subClassOf :Position.:mikesFordFocus a motion:MovingObject]:motion1 a……

Virtues of the ST Pattern as an ODP

Answers Competency questions (an example of bird trajectories)

- "Show the birds which stop at x and y" (spatiotemporal data)
- "Show the birds which move at a ground speed of at least 0.4 m/s“ (trajectory attribute data)
- "Show the trajectories which cross national parks" (geographic information)
- "Show the trajectory which cross national parks" (geographic information)

- Expressive.
  - It captures spatial-temporal properties, geographic knowledge, domain knowledge, as well as the relations among them.
- Simple.
  - Only a minimal number of classes and relations are introduced to make the pattern easy to understand, reuse, and extend.
- Flexible. (Illustrated on next page)
  - It can be applied to data in a range of domains, and can be integrated with existing ontologies.
- Scalable.
  - Depending on the required granularity of the application, the pattern can annotate trajectories at different scales.

In Formalized in Axiom 1, an x is constrained to have a timestamp and a position associated.

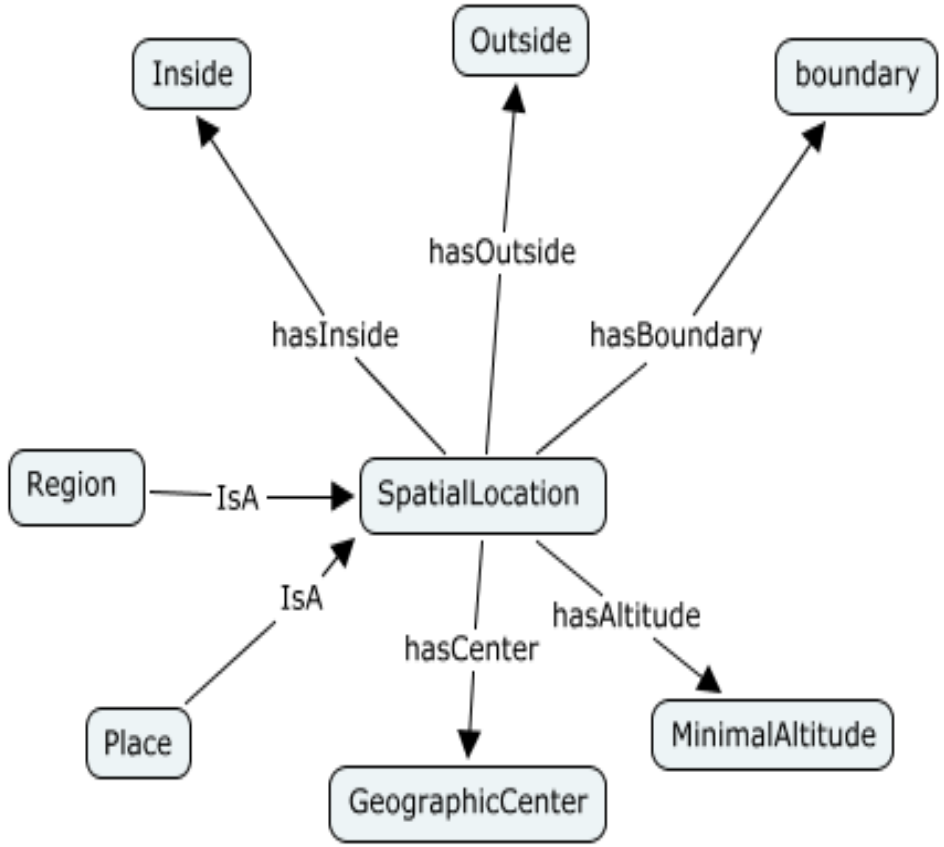
- By Axiom1, an x is enforced to have a timestamp and a position associated. An x always belongs to a traiectorv (also enforced bv Axiom 1 below).

$$Fix \sqsubseteq \exists atTime.OWL\text{-}Time{:}Temporal\ \ Thing \sqcap \exists hasLocation.Position \sqcap \exists hasFix^{-}.SemanticTrajectory \quad (1)$$

- **Flexibility** - The number of Fixes in a trajectory depends on the requirements of the particular application.
  - It can be as coarse as containing only the important trajectory points (e.g., check-ins on location based social networks LBSN), but can also be as ne as including every point recorded as a function of the sampling rate of the location-tracking device.
- This design makes the pattern flexible allowing users to model trajectories under different scales. Real-world examples of a x may include a stop of a migration flock in a wet land area, an important intersection a particular vehicle has passed by, or a restaurant.

Tools

- It is often useful to start with hand/board drawings to accommodate conversational flow.
- PowerPoint graphics can be used to tidy things up for presentation.
- Better yet is a modeling tool like CMAP with support for model constructs and automatic translation into OWL/TTL etc.



Illustrations of Conceptual Modeling & ODPs

Settings and Cadastral Concepts

What Do We Understand about State Street in Madison?

State Street is a straight pedestrian road which connects the State capital building to UW. Lined with over 300 shops, most of which are local start ups or franchises, State Street is able to cater to the student population with a mixture of restaurants, bars, clothing stores, coffee shops, theatres, book stores, liquor stores, & misc. consumer goods. Over a 1/2 mile stretch..

Do we refer to it a pedestrian mall or a shopping street?

Is it a road object, but with motor traffic restrictions?

(100-600 blocks of State Street are closed to vehicle traffic, with the exception of buses, bicycles, and authorized vehicles.)

Or a public place? A supplement to the University & students?

Or a non-identifiable part of the city surface?

All such referenced meanings are usually outside a computer

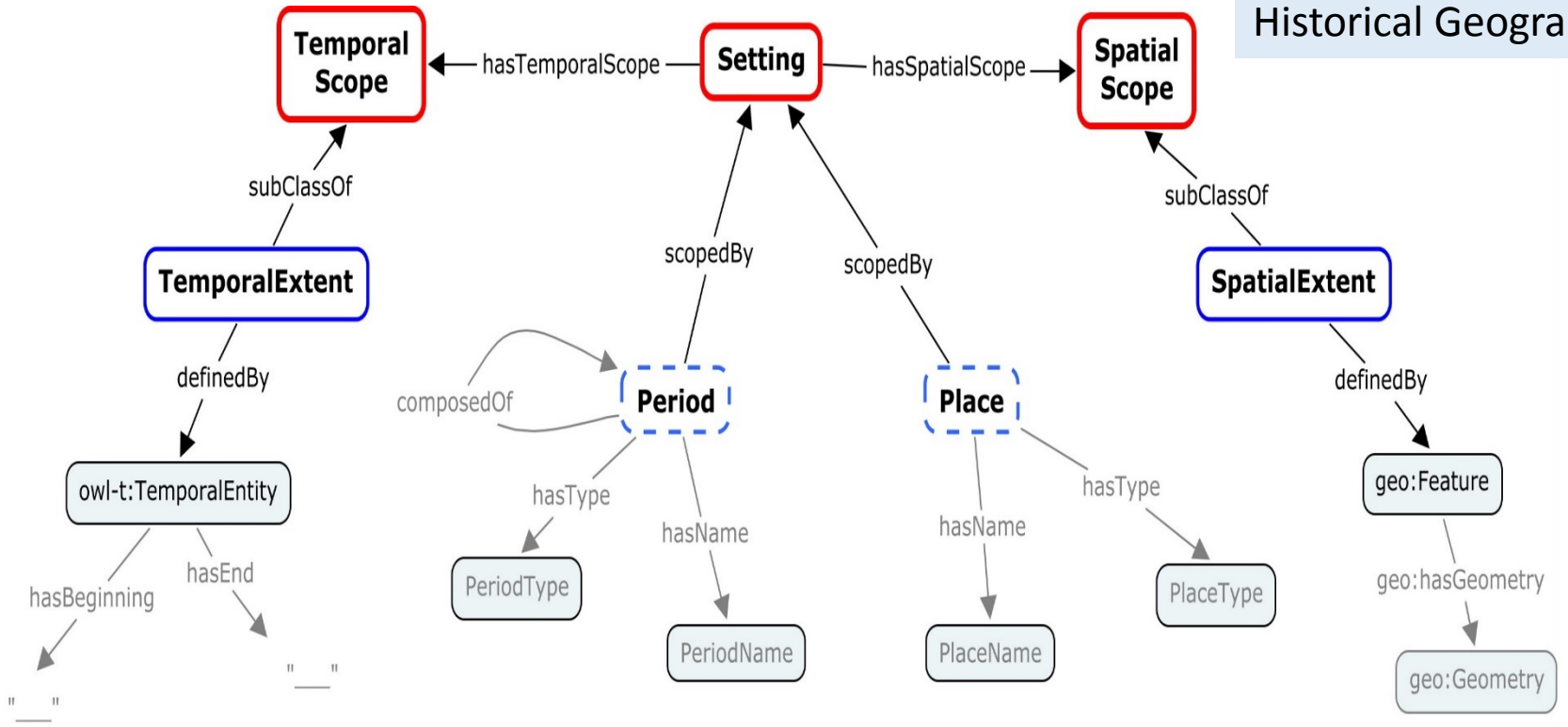
Food

Theatre

OpenStreetMap -

OpenStreetMap Humanitarian layer -

ODP Example & its Evolution – Phase 1 - Setting: Something kind-of temporal (e.g. the sixties, the 19th century) or something kind-of spatial (e.g. France)



Historical Geography

Place, Period, and Setting for Linked Data Gazetteers

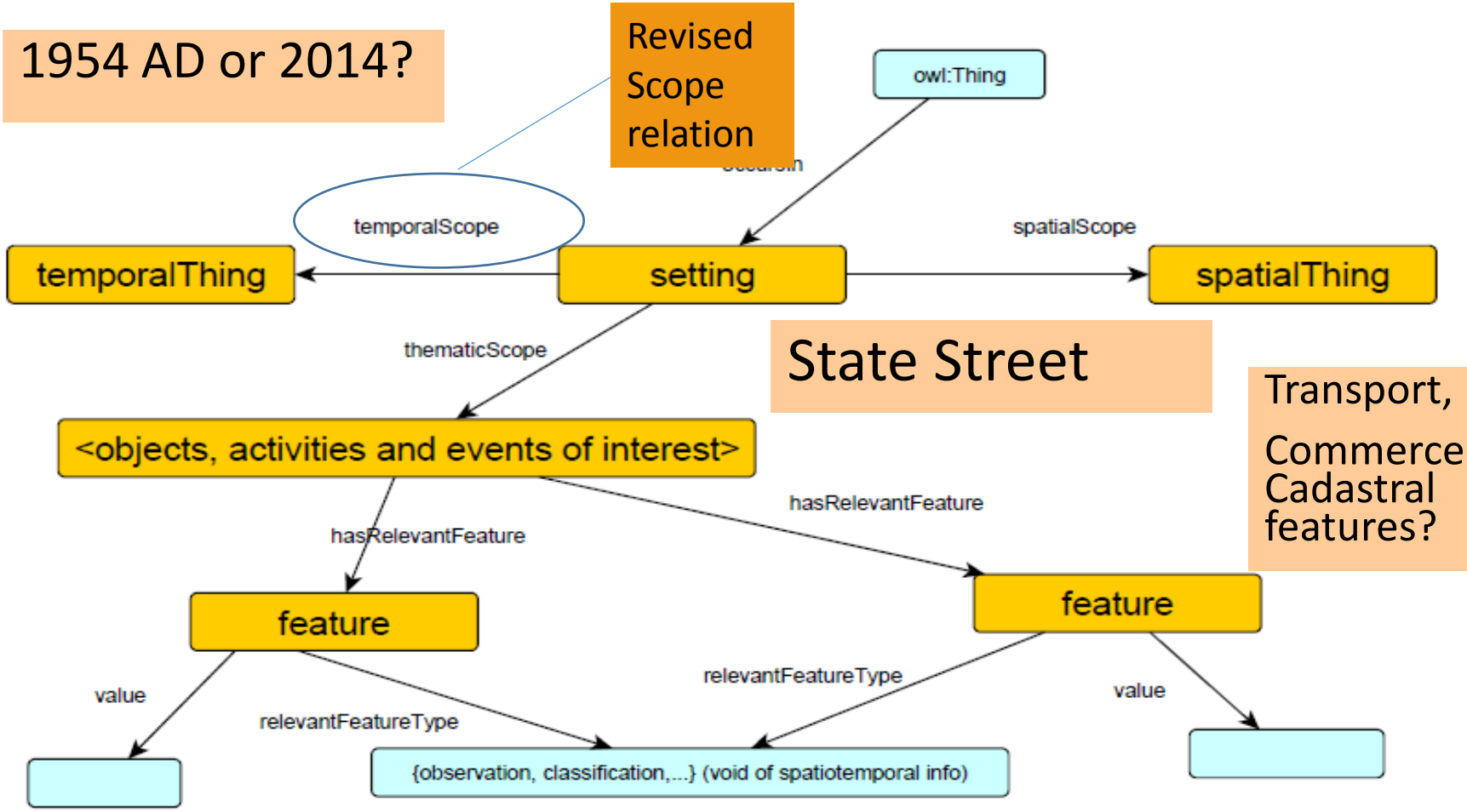
Karl Grossner,
Krzysztof Janowicz,
And Carsten Keßler

Definition of a setting (could be a Cadastral Setting).
A setting is a geospatial temporal region within which objects, activities and events occur.
Our settings of interest are all the settings in which the objects, activities, and events of interest occur.
Based on Worboys & Hornsby, (2004).”From objects to events: GEM, the Geospatial event model.”

Revised "Setting ODP" (UCSB GeoVoCamp 2014)

1954 AD or 2014?

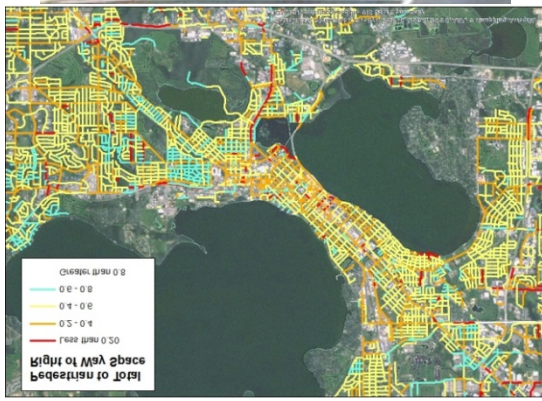
Revised Scope relation



State Street

Transport,

Commerce or Cadastral features?



Mark Schildhauer, Gary Berg-Cross, Charles Vardeman, Pascal Hitzler, Helen Couclelis, Francis Harvey, George Planansky, Ben Adams, Andrea Ballatore, Krzysztof Janowicz, Dave Kolas

Cadastral “Layers” & Fabrics: In GIS (like ArcGIS) a cadastral layer delineates real property parcels.

The **cadastral parcel** is one of the core elements of the European SDI (ESDI) as part of the upcoming implementation of the INSPIRE-directive (Infrastructure for Spatial Information in the European Community).

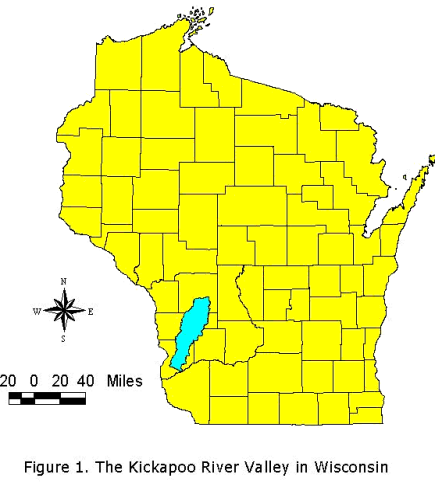
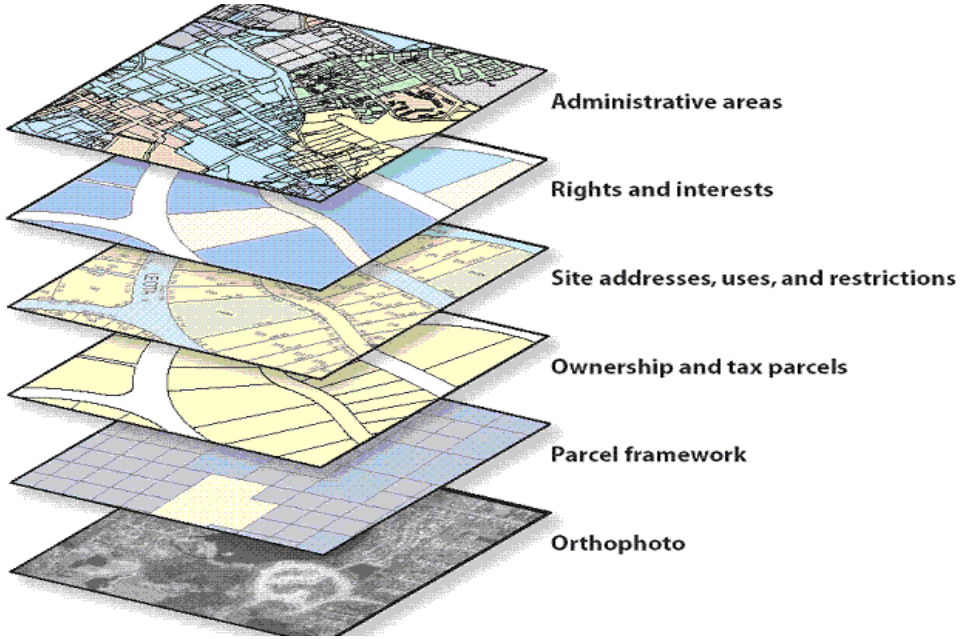
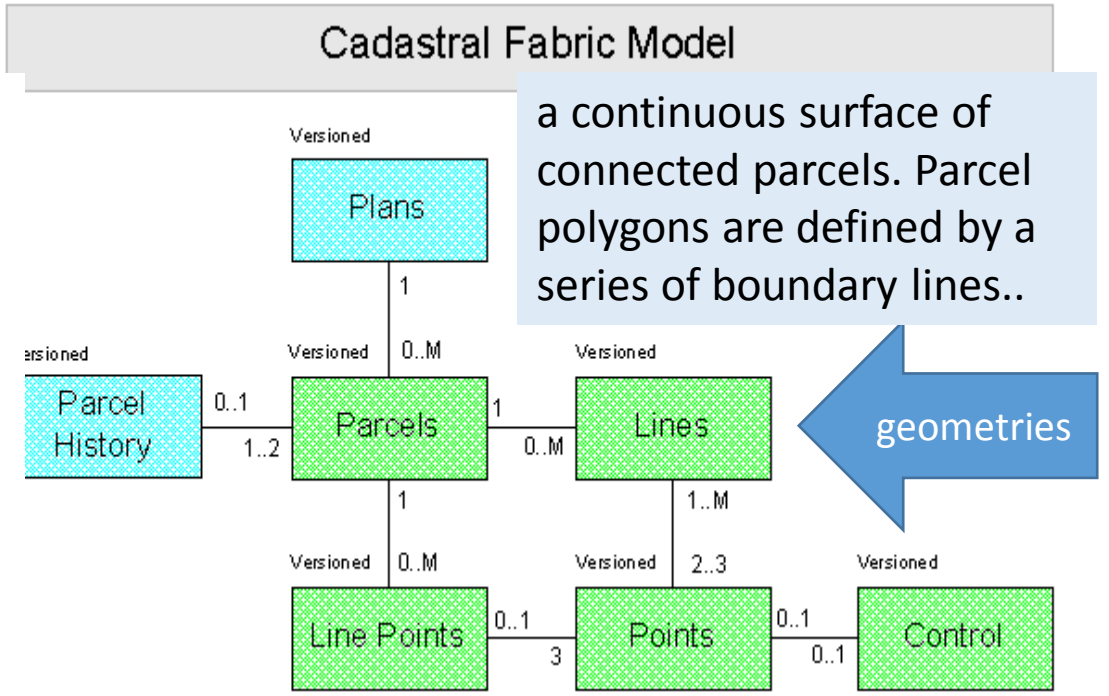


Figure 1. The Kickapoo River Valley in Wisconsin





We may treat this domain as a special type of identifying parts and their relations which involve legal, administrative, political and ownership ideas constructed from legal descriptions contained in deeds, survey maps plats, and may be shown on tax maps.

Broaching the Big Topic of Cadastral Parceling

- But parcels have geometries and boundaries which have ODPs we can leverage.

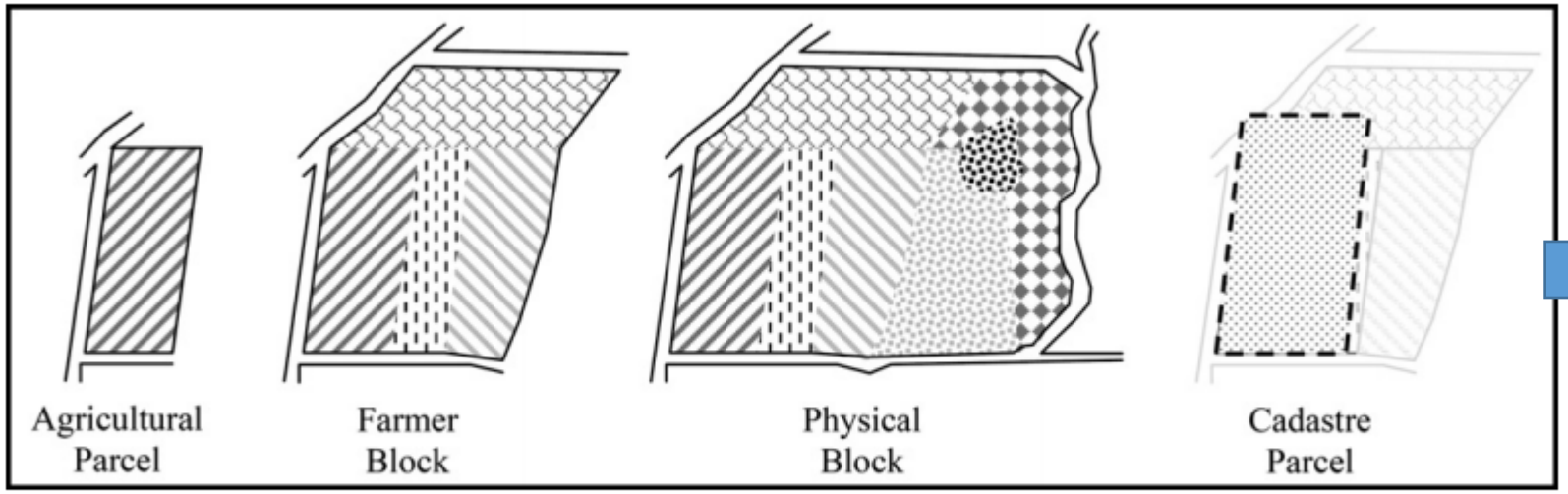


Fig. 2. Different reference parcel types used for IACS/LPIS applications and their relation (adopted and elaborated after JRC, 2001).

These are our Cadastral Objects with part & geo-relations (inside, touches, has boundary, north of….)

From Data model for the collaboration between land administration systems & agricultural land parcel identification systems Halil Ibrahim Inana et al, Journal of Environmental Management Vol. 91, Issue 12, Dec. 2010, Ppg 2440–2454

Start on Core Ideas and Vocabulary

- We can identify the units (registered Objects) & parts, but is there agreement on names?
  - blocks have parcels and/or sub-blocks
  - “A Parcel is a single cadastral unit – a Registered Object, which is the spatial extent of the past, present, and future rights and interests in real property. “
- Parcels sub-types - Agricultural, commercial, educational, tax….

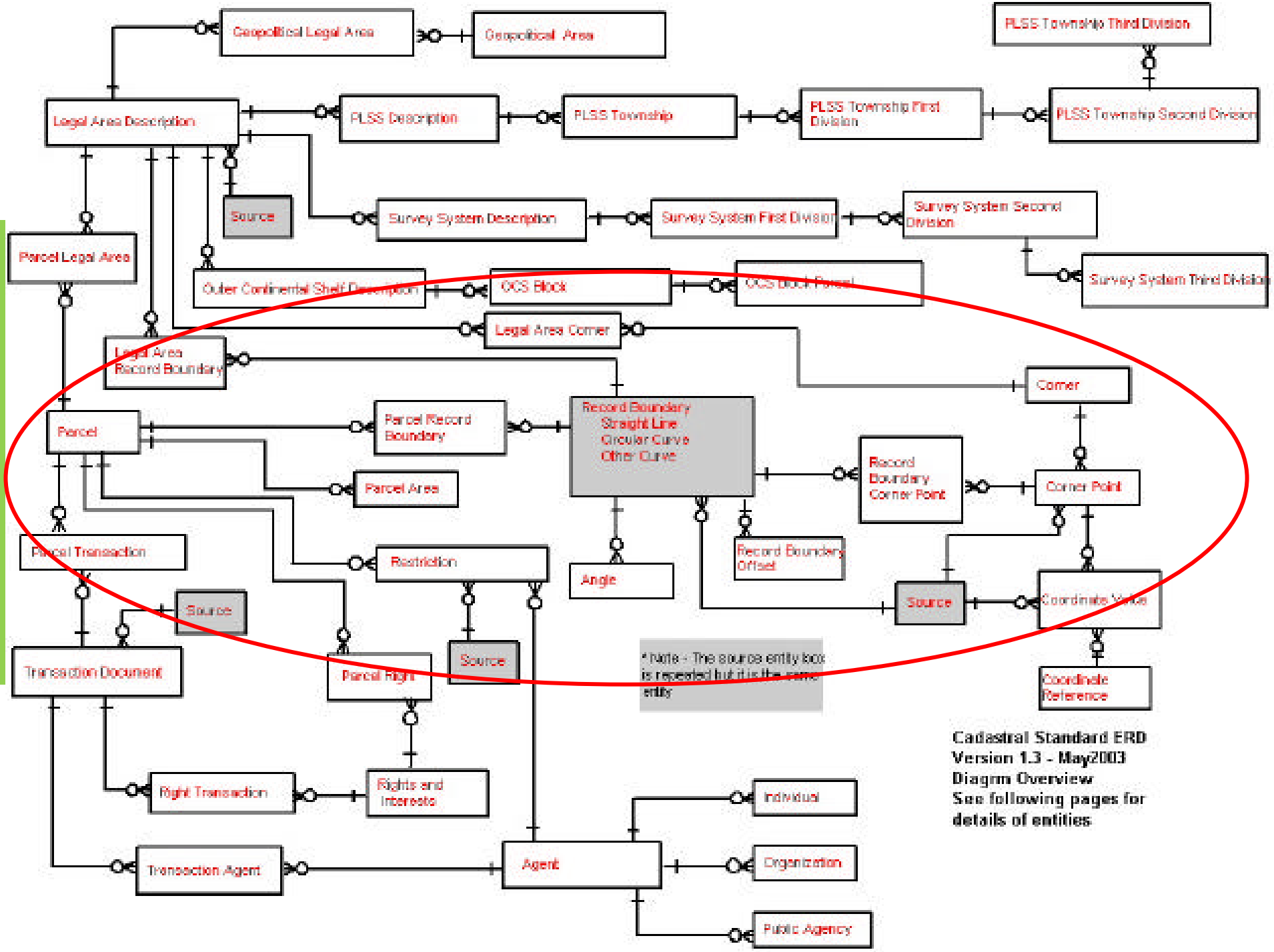
Core concepts

- Parcel characteristics/attributes such as Parcel area, but also Rights and Person associated with it.
- Basic Geometries - A parcel in the parcel fabric is composed of point features, line features, and polygon features.
  - The parcel polygon is defined by a sequence of individual lines, which form a closed loop to define a polygon.
- Additional Geometries comprised of fields considered necessary to uniquely define and reflect the spatial orientation and location of a parcel
  - – **Radius**, **Central Angle, Direction of Curve, Degree of Curve**..
- Boundaries are very important
  - Structures on parcels might be deferred till a subsequent VoCamp.

Maybe we just start from a standard and model these.

- Something like:
  - Cadastral Data Content Standard for the National Spatial Data Infrastructure - May 2003 – Version 1.3

Cadastral Data Content Standard for the National Spatial Data Infrastructure - May 2003 – Version 1.3



Composition and Part “Relations”

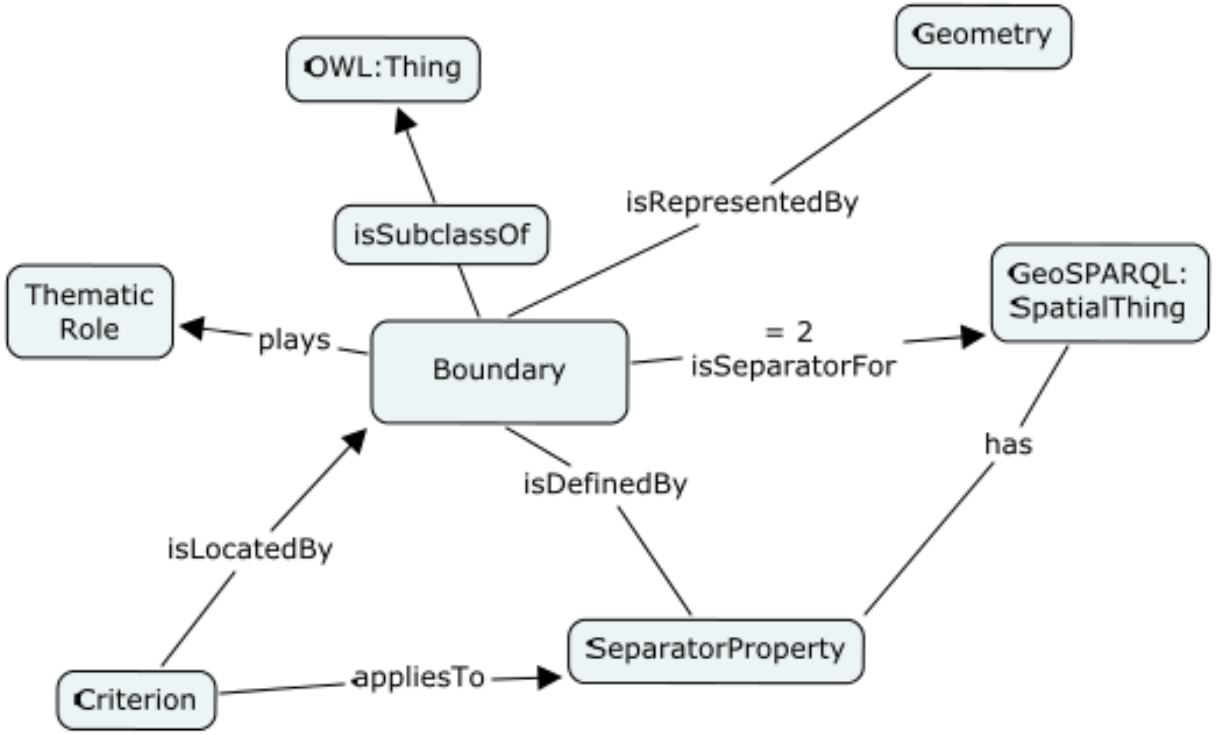
Just as relationships exist between parcels, lines, and points in the cadastral fabric, relationships exist between parcel features and other cadastral fabric objects such as the relationship between parcels and plans.

The list summarizes below the relationships between parcel features themselves and between parcel features such as address or assessed value and other objects:

http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=Modeling_a_parcel_in_the_cadastral_fabric&anchor=Line%20points

- A parcel polygon is related to many lines. (What do we mean by “related”?)
- A parcel line has two endpoints.
- A parcel line is related to only one parcel, resulting in two lines representing common parcel boundaries.
- A parcel point is related to one control point.
- A parcel point is related to one or more lines.
- A line point is related to one or more parcel polygons.
- A line point is related to one parcel line.
- A parcel point is related to one line point.
- A parcel polygon can have many line points.
- A parcel polygon is related to one plan.
- A parcel point can have one or more adjustment vectors (from a parcel fabric adjustment).
- A parcel polygon has one accuracy category (used by the parcel fabric adjustment).
- A parcel line has one accuracy category (used by the parcel fabric adjustment)……

Reuse of Existing Patterns



We might use this for a Cadastral Role.

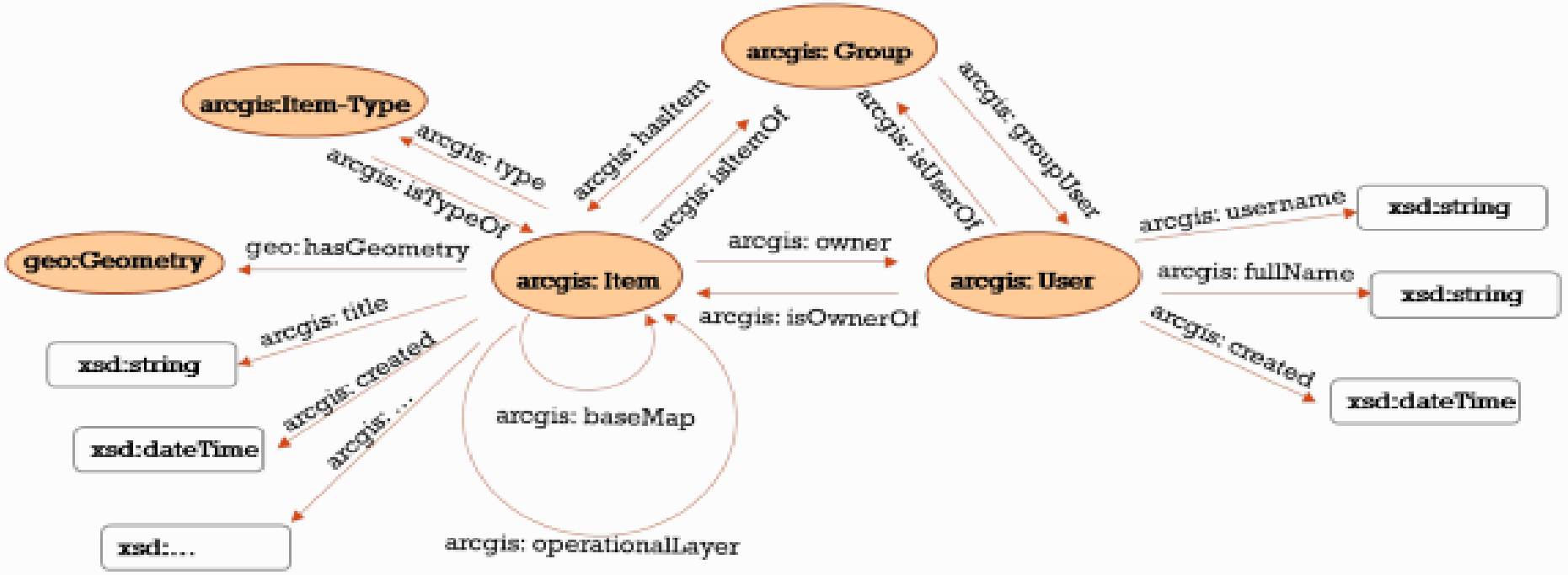
A criteria is specified to separate 2 parcels.

Cadastral Setting

- Parcel identity involves time as well as space.
- A “piece” of land has a different parcel type over time and it may have a mix of parcel types.
- A parcel may be represented at different degrees of accuracy (point, lines)
- A parcel’s setting may change due to the process of a “parcel transaction” in which there is a change of title etc….
- Is the Setting pattern useful for starting on these issues?

Backup

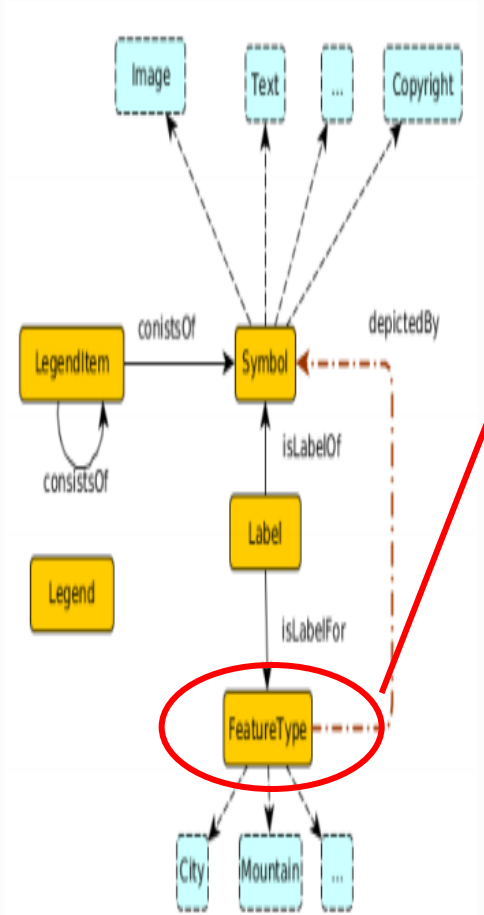
A TINY ONTOLOGY FOR ESRI'S ARCGIS ONLINE



This fragment of the **ontology** developed for ArcGIS Online defines the relations (e.g., isOwnerOf) between items (e.g., map services), users, and user groups.

Making Commitments to an ODP for Maps - Legends

Fragment of a Map Legend Ontology Design Pattern



Ontological commitments

- Should Geographic Feature Types be **classes or instances**?
- Do we want to explicitly define the **depictedBy** relation
- Is stating that a **Legend** consists of LegendItems redundant?
- ...

Fragment of the Map Legend Ontology

$$\mathbb{N}_C = \{LegendItem, Symbol, Label, FeatureType\} \quad (1)$$

$$\mathbb{N}_R = \{consistsOf, isLabelFor, isLabelOf, depictedBy\} \quad (2)$$

$$\top \sqsubseteq \neg\exists N.\top \quad (3)$$

$$LegendItem \sqsubseteq \exists consistsOf.Symbol \sqcup \exists consistsOf.LegendItem \quad (4)$$

$$Label \sqsubseteq \exists SymbolizedBy.Symbol \sqcap \forall SymbolizedBy.Symbol \quad (5)$$

$$\top \sqsubseteq \leq 1\, isLabelFor.\top \quad (6)$$

$$\top \sqsubseteq \leq 1\, isLabelOf.\top \quad (7)$$

$$\top \sqsubseteq \leq 1\, SymbolizedBy.\top \quad (8)$$

$$Label \sqsubseteq \exists isLabelFor.FeatureType \quad (9)$$

$$Label \sqcap Symbol \sqsubseteq \bot \quad \text{(also for Symbol, Label, FeatureType, LegendItem)} \quad (10)$$

$$isLabelOf^- \circ isLabelFor \sqsubseteq depictedBy^- \quad (11)$$

$$\neg\exists consistsOf^- \sqsubseteq Legend \quad (12)$$

$$\ldots \quad (13)$$

Parcel Types ?

- We can start defining parcel types.