Bringing semantic and versioning to Data (Meta) Models

David Languignon
Franck Le Petit
Nicolas Moreau

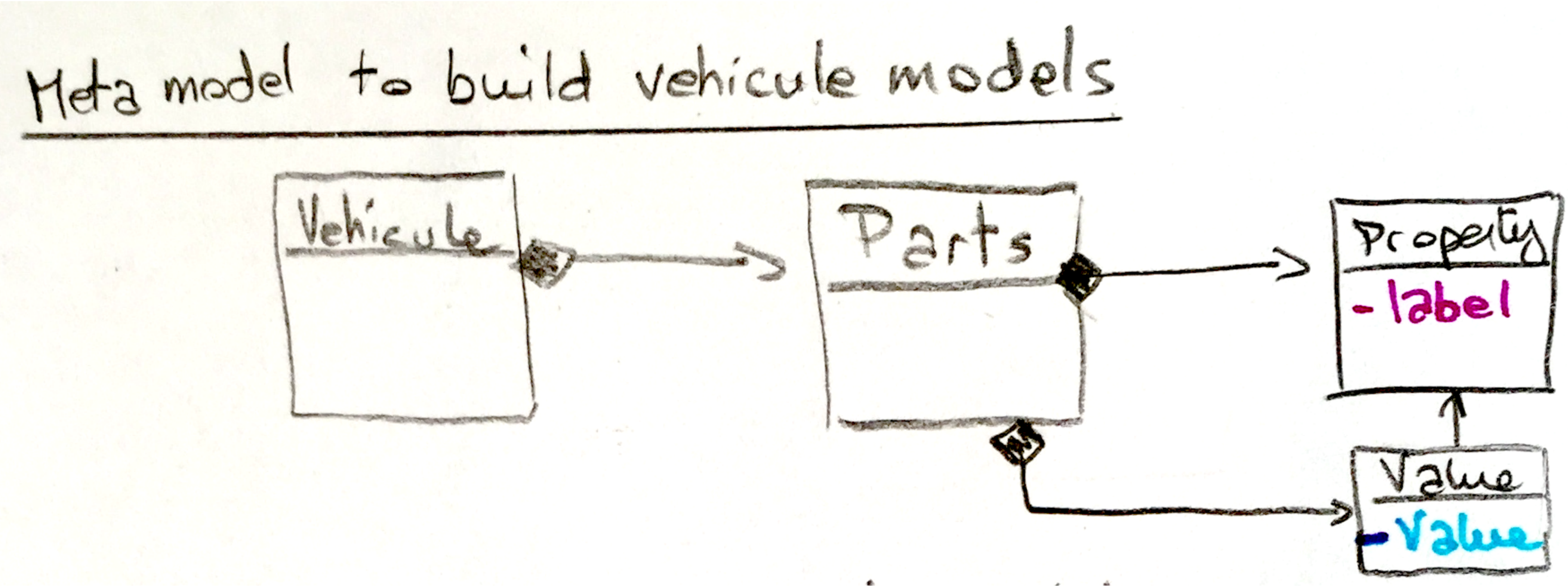
October 21 2016
Trieste IVOA interop - Italy

Data Meta Model

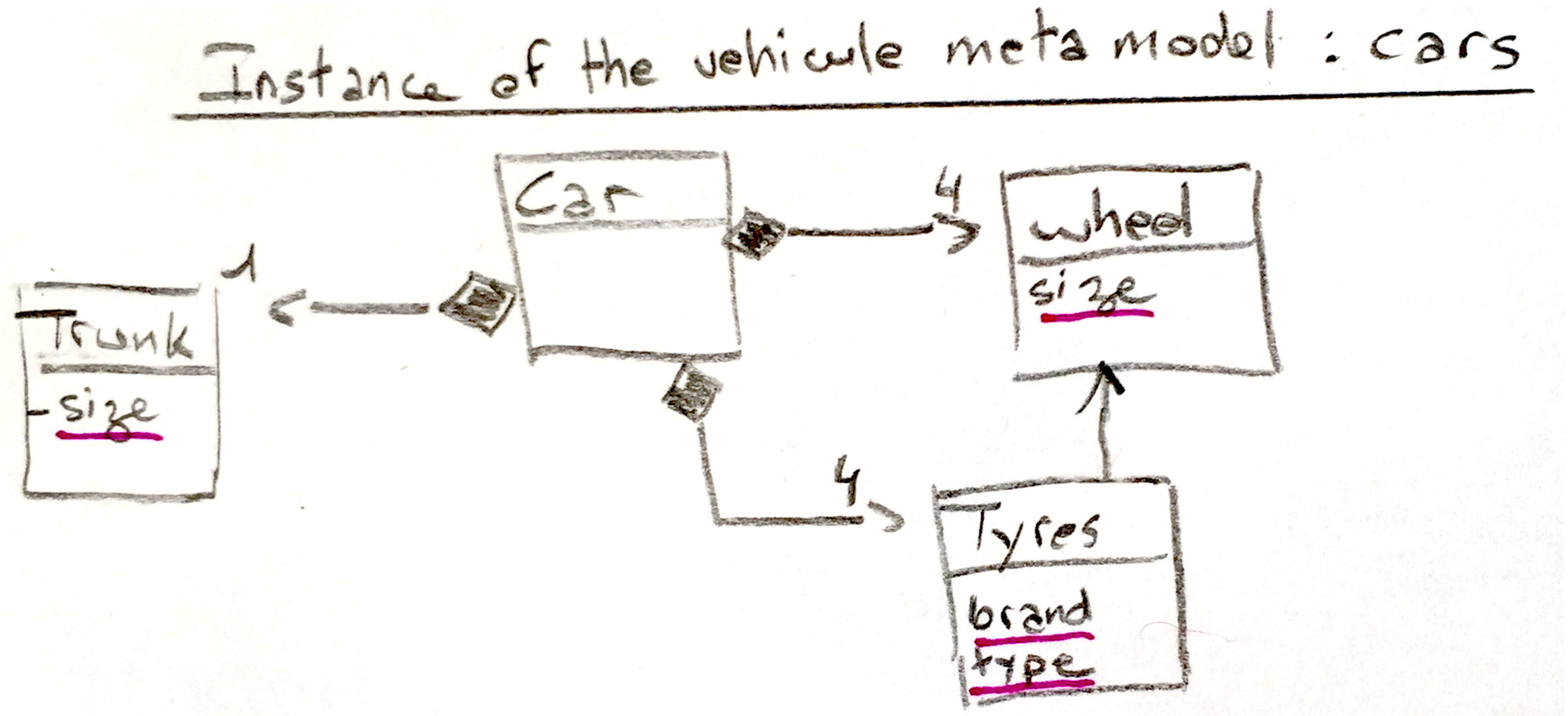
- A model to build data models
- Classes that once instanciated form a data model

- SimDM is a Data Meta Model
  - Because 1 specific Data Model must be defined per Code, and there are a lot of Codes to be modeled…

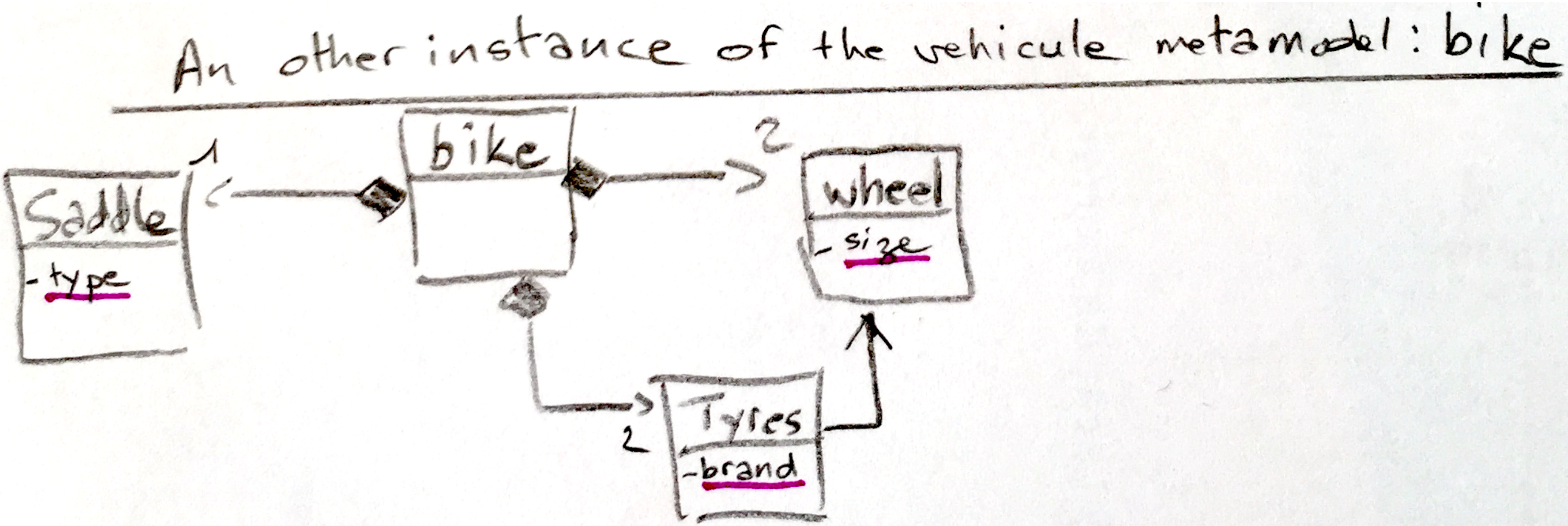
Let‘s build models of vehicles



Let‘s build a model of cars



Let‘s build a model of bikes



What do we really know about a car?

- That it has a component named wheel
- That the wheel component has a property named size

So we know what wheel and size are

- But we don‘t really know what wheel and size mean
  - at least for the person who named them this way

So we know what wheel and size are

- But we don‘t really know what wheel and size mean
  - at least for the person who named them this way

The Data **Meta** Model
defines this

A classical Data ~~**Meta**~~ Model
defines this.

So we know what wheel and size are

- But we don‘t really know what wheel and size mean
  - at least for the person who named them this way

The Data **Meta** Model
defines this

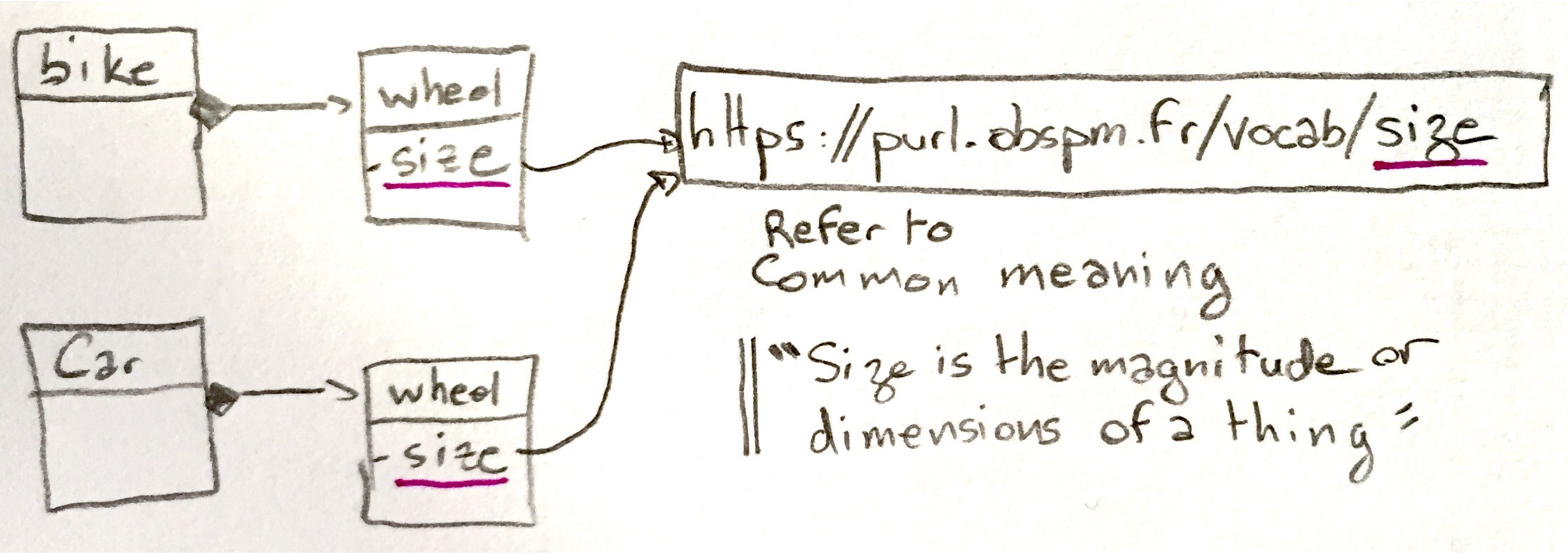
A classical Data ~~**Meta**~~ Model
defines this.

But the Data Meta Model can‘t.
Because we have an infinite number of
classical Data Model for 1 given
Data Meta Model

What to do ?

- A solution is to join a vocabulary with each Data Model instance.
  - But need to have standard/shared vocabularies

Why a standard vocabulary ?



Not so easy

- How to develop / use a vocabulary ?
  - It‘s common to have small specific vocabularies inside a team/code community.
  - Moderation/curation
  - Hosting

What we do at Paris Observatory

- SimDAL service with SimDM xml serializations
- We developed, maintain and use the standard (IVOA?) vocabulary for theory

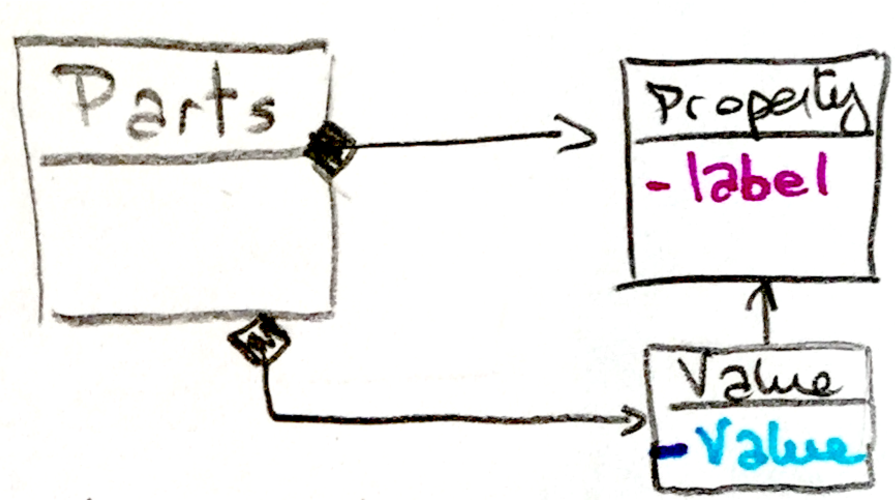
  Term example : http://purl.obspm.fr/vocab/Algorithms/GaussSeidel

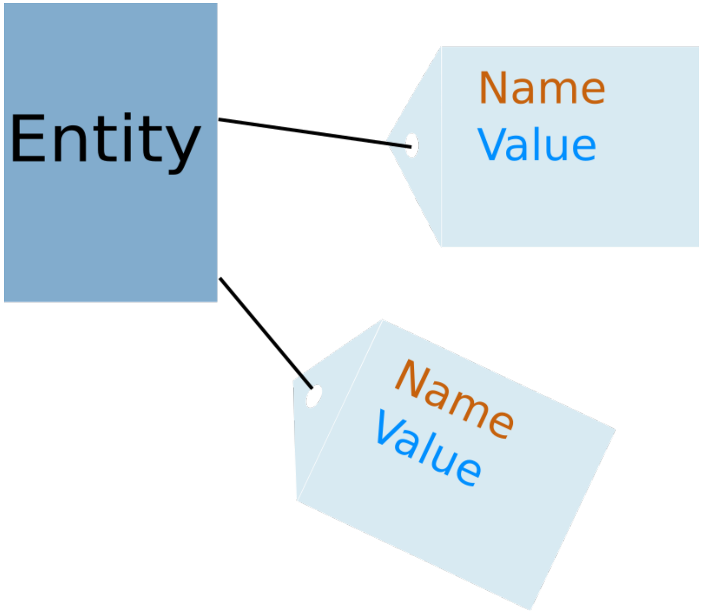
  - The vocabulary has been being developped over years by several astronomers (Paris, Strasbourg, Madrid...)
  - Public access with full featured UI and API (see N.Moreau talks)

http://votheory.obspm.fr/new_skos_service/

What we do at Paris Observatory

- We use entity tagging as implementation of data meta models

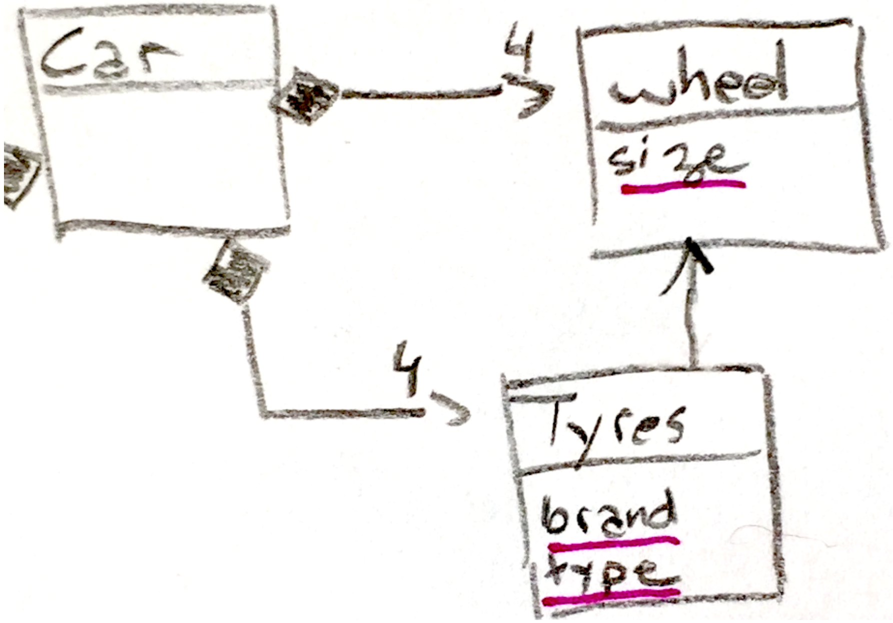


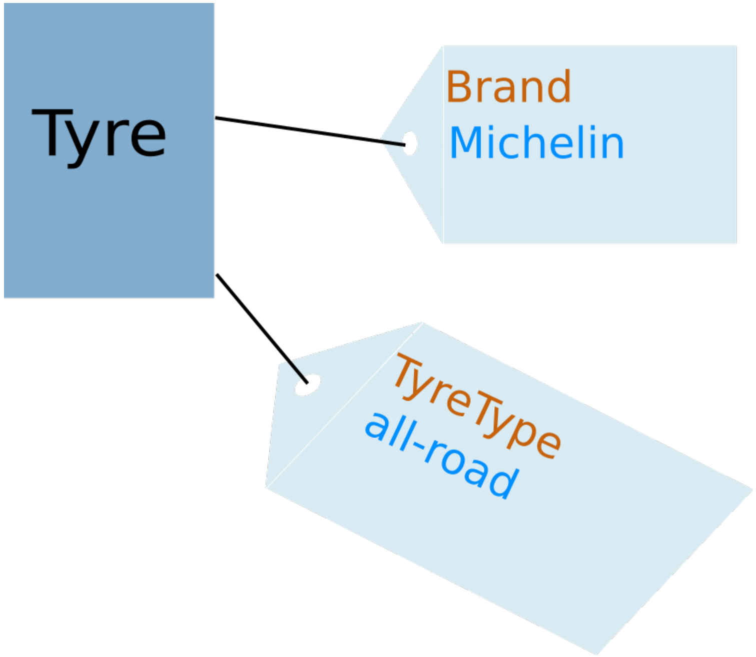


<u>Mapping</u>

Part	→	Entity
Property	→	Tag
Property.label	→	Tag.name
Value.value	→	Tag.value

Entity tagging: enable semantics





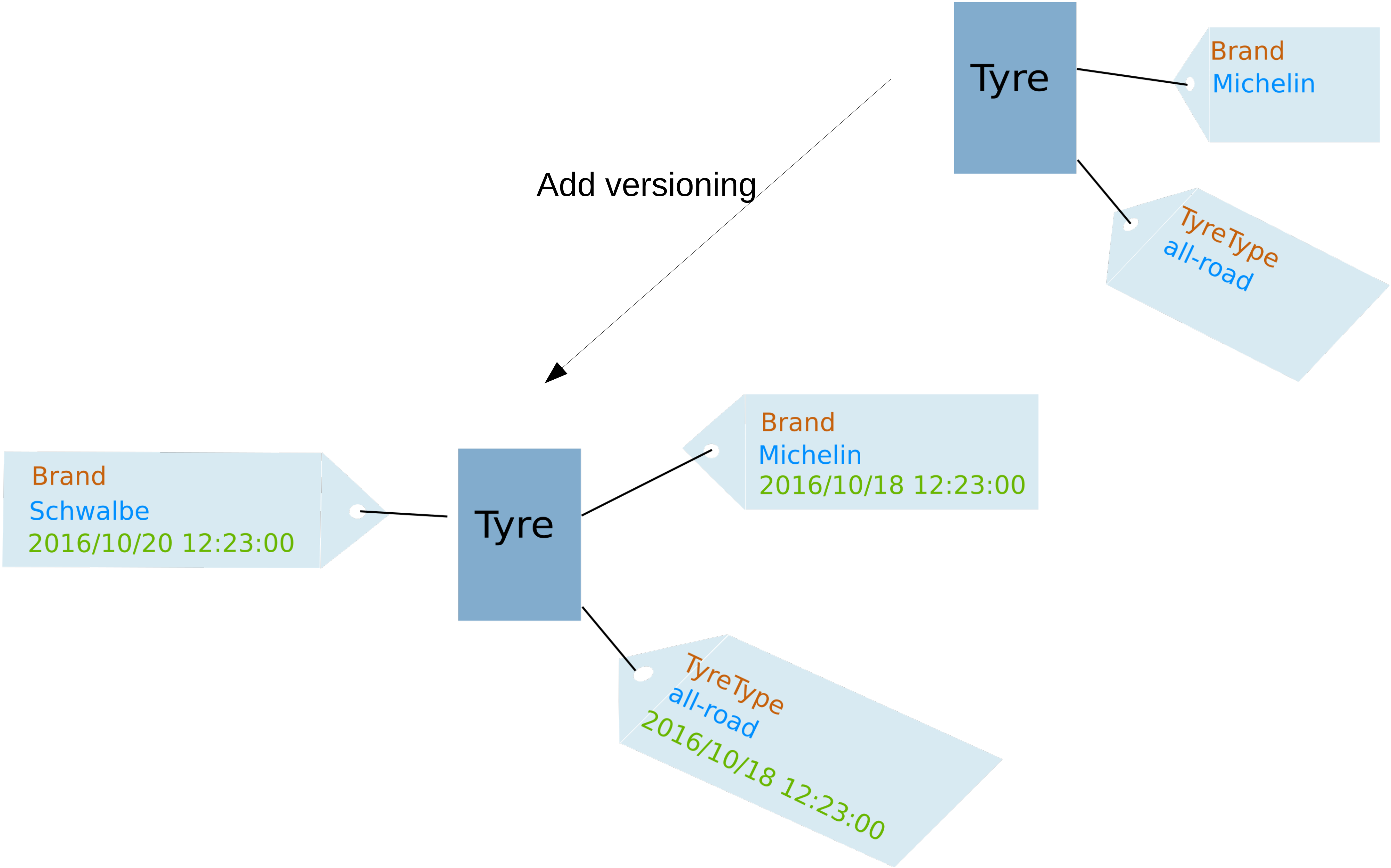
Tag names have namespace

https://purl.obspm.fr/vocab

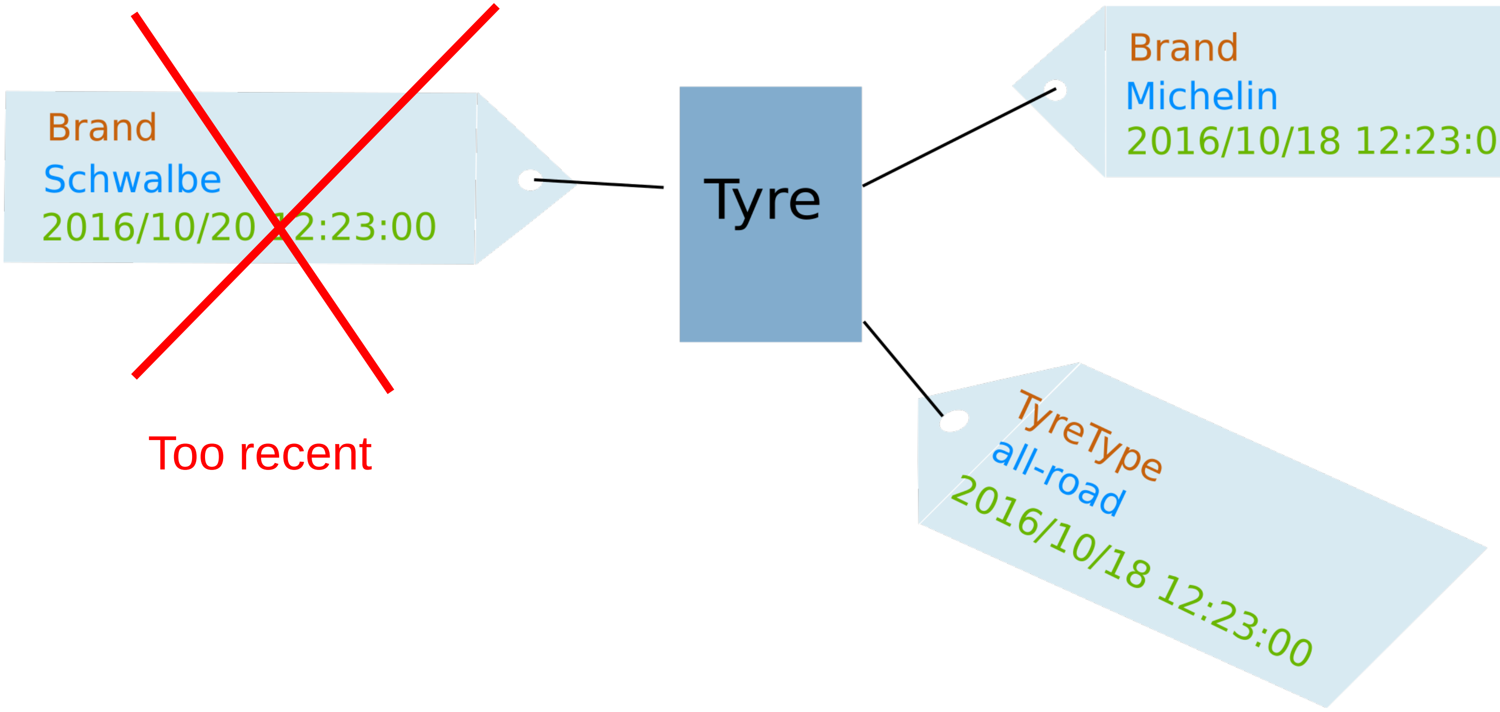
How hard would it be to add versioning ?

- Because we may update the tags attached to some entities over time, in our reference database.
  - What about the papers referencing a version of the database before an update ?

Versioning is as easy as adding a timestamp to tags



Entity tagging: enable versioning



Too recent

Database version 2016/10/19

Take away

- Semantic is mandatory for Data Meta Models to be usable
  - Standard vocabularies are required and not easy to define policies to set / manage it.
- Data Meta Models can be friendly implemented as Entity tagging (or EAV, Prolog facts (yes, 1972!))
  - Easy to plug semantic through vocabulary links
  - Versioning for (almost) free