



HAL
open science

An Ontology-based Algorithm for Managing the Evolution of Multi-Level Territorial Partitions

Camille Bernard, Christine Plumejeaud-Perreau, Marlène Villanova-Oliver,
Jérôme Gensel, Hy Dao

► **To cite this version:**

Camille Bernard, Christine Plumejeaud-Perreau, Marlène Villanova-Oliver, Jérôme Gensel, Hy Dao. An Ontology-based Algorithm for Managing the Evolution of Multi-Level Territorial Partitions. 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL 2018), Nov 2018, Seattle, WA, United States. , 2018, 10.1145/3274895.3274944 . hal-01929649

HAL Id: hal-01929649

<https://hal.science/hal-01929649v1>

Submitted on 21 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright



Camille Bernard
Univ. Grenoble Alpes, France
camille.bernard@imag.fr

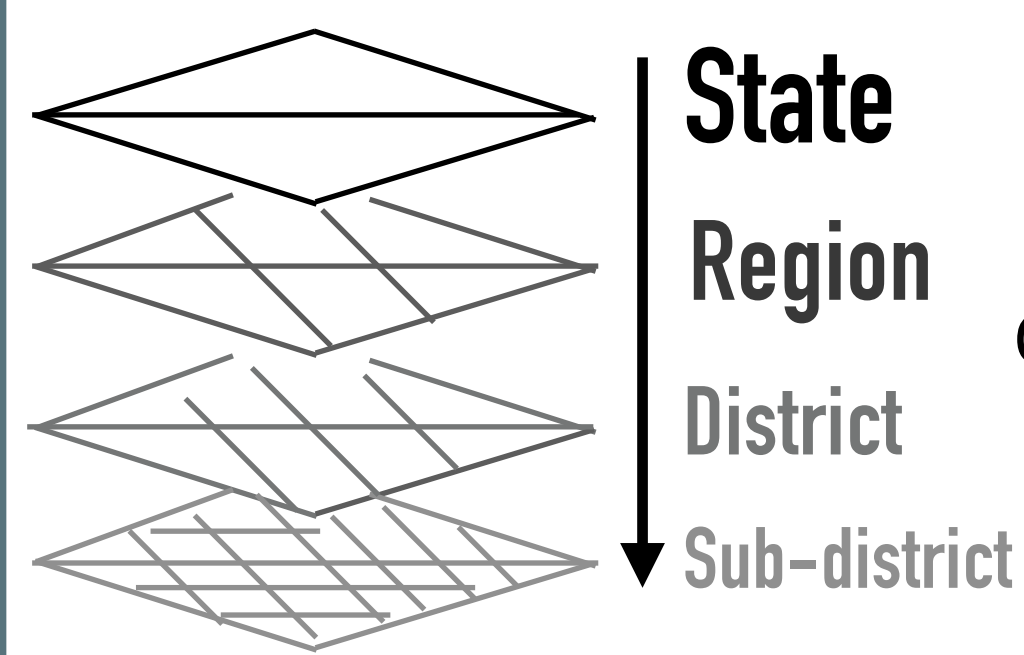
Christine Plumejeaud-Perreau
Littoral ENvironnement et Sociétés (LIENSs), France
christine.plumejeaud-perreau@univ-lr.fr

Marlène Villanova-Oliver
Univ. Grenoble Alpes, France
marlene.villanova-oliver@imag.fr

Jérôme Gensel
Univ. Grenoble Alpes, France
jerome.gensel@imag.fr

Hy Dao
Univ. of Geneva, Switzerland
hy.dao@unige.ch

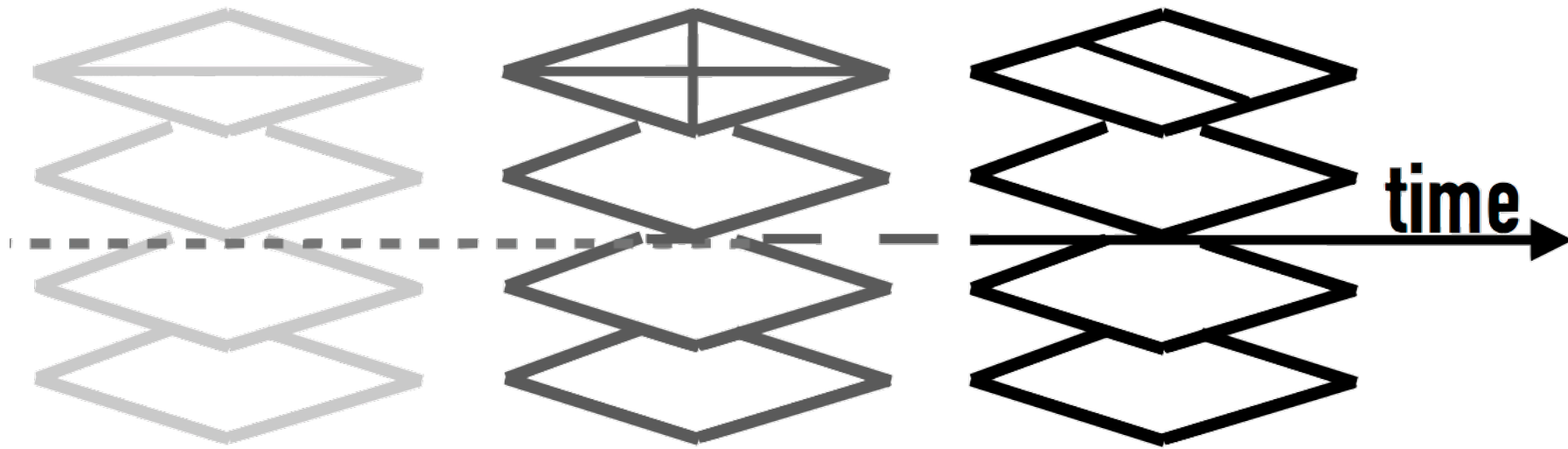
CONTEXT



State
Region
District
Sub-district

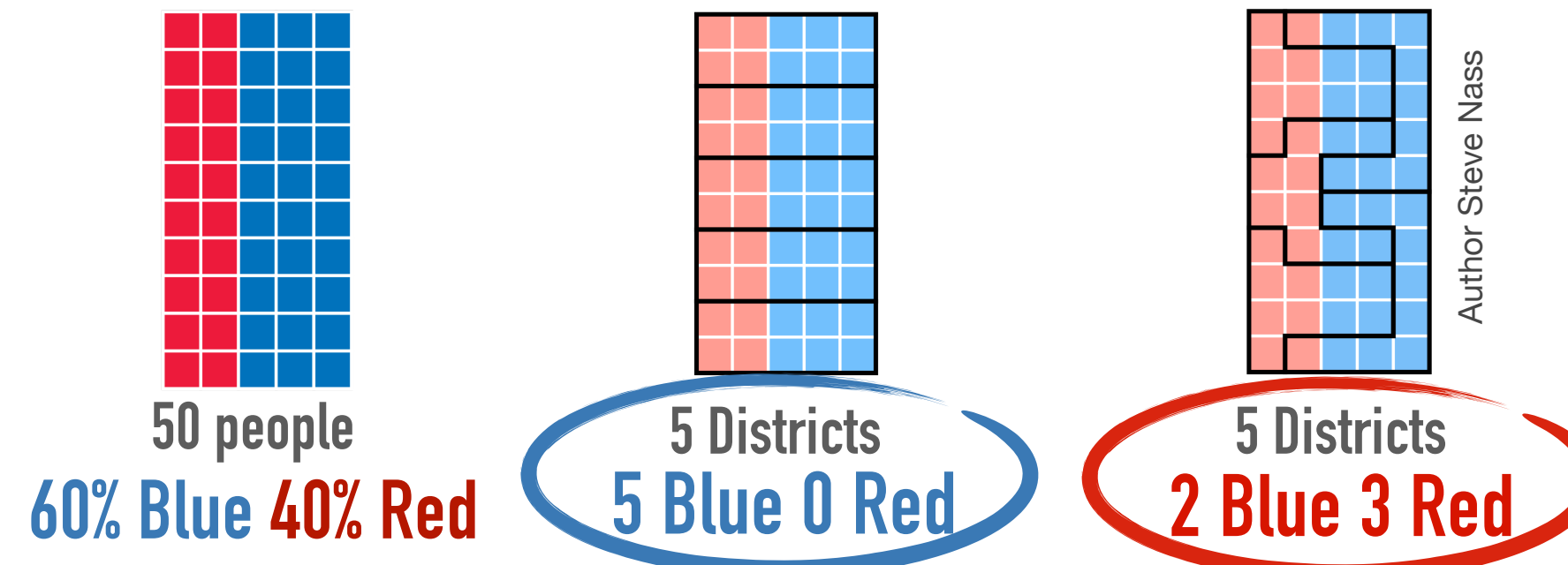
Territorial Statistical Nomenclatures (TSN) are hierarchies of territorial divisions (e.g., district, sub-district levels) defined with the purpose of collecting and presenting statistical data.

TSN change over time, for political or administrative reasons



MODIFIABLE AREAL UNIT PROBLEM (MAUP)

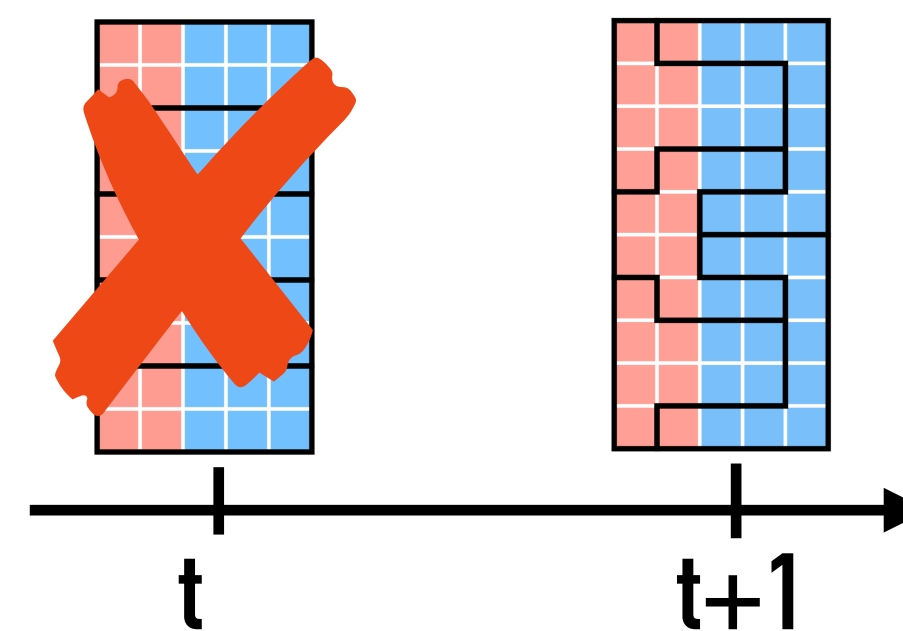
Influence of the partition on data [Openshaw & Taylor]



MAUP + TIME

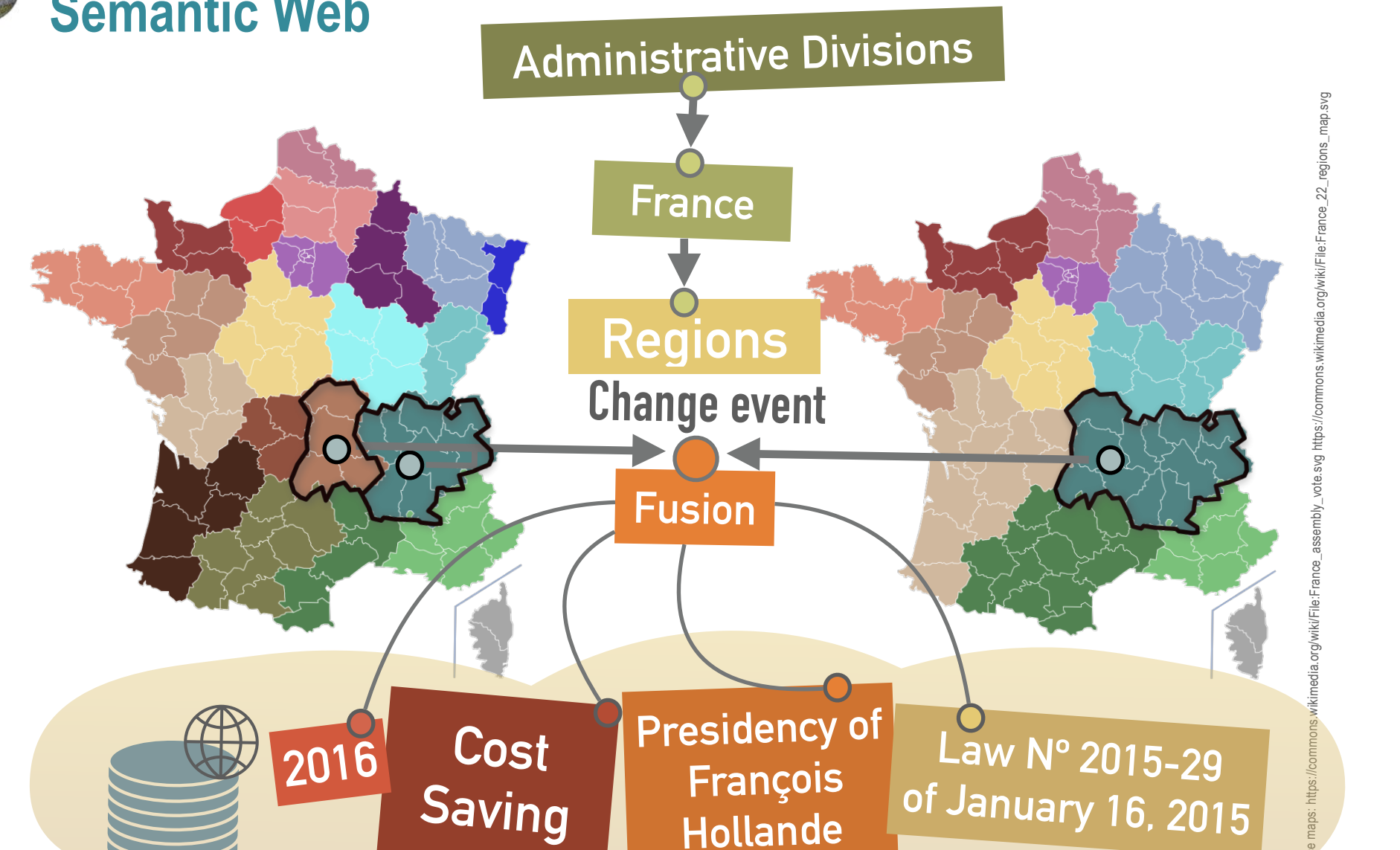
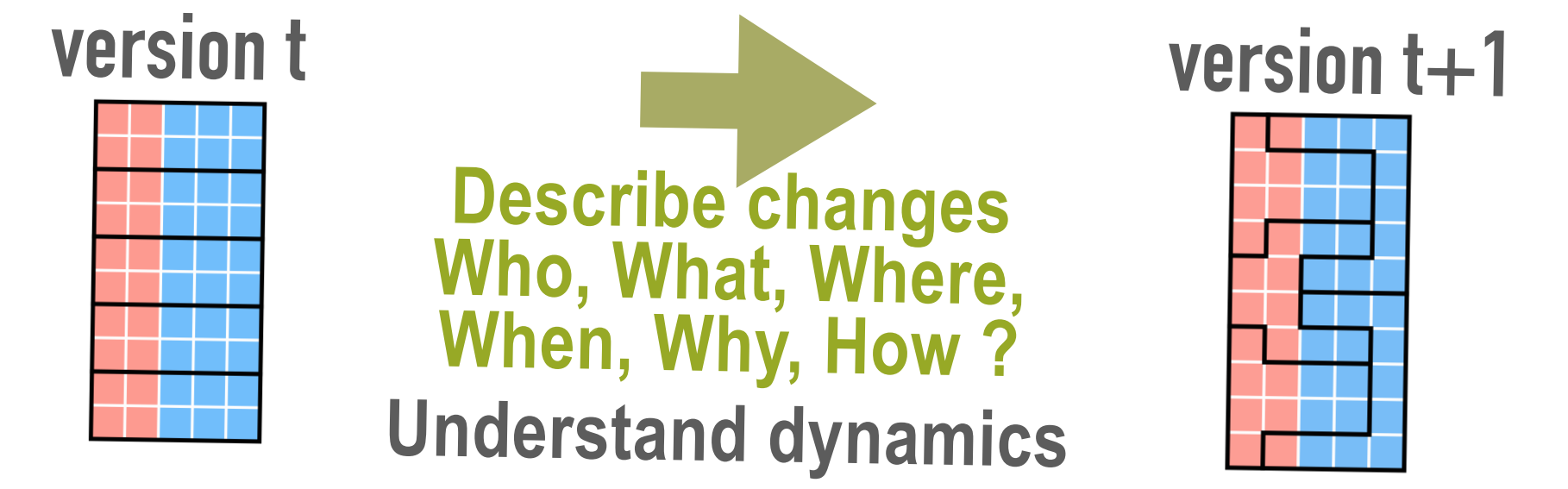
Data are not comparable due to potential changes of the territorial units over time

Classic solution:
Access to statistical data in the latest territorial division version only



OUR APPROACH

Modifiable Areal Unit Potential [Grasland and Madelin]

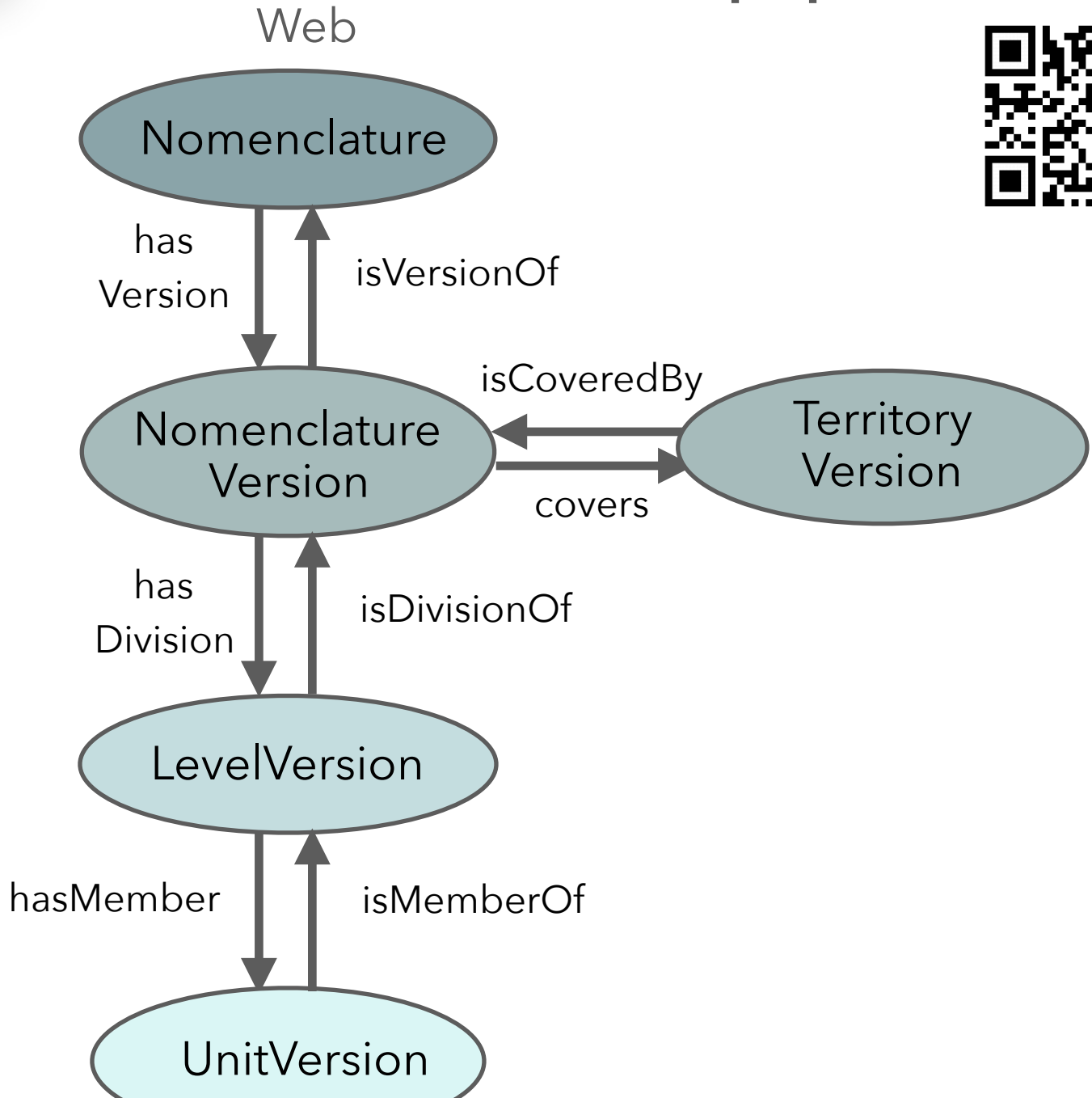


PROPOSITIONS

1 TSN ONTOLOGY

A generic model for Territorial Statistical Nomenclature in the LOD Web

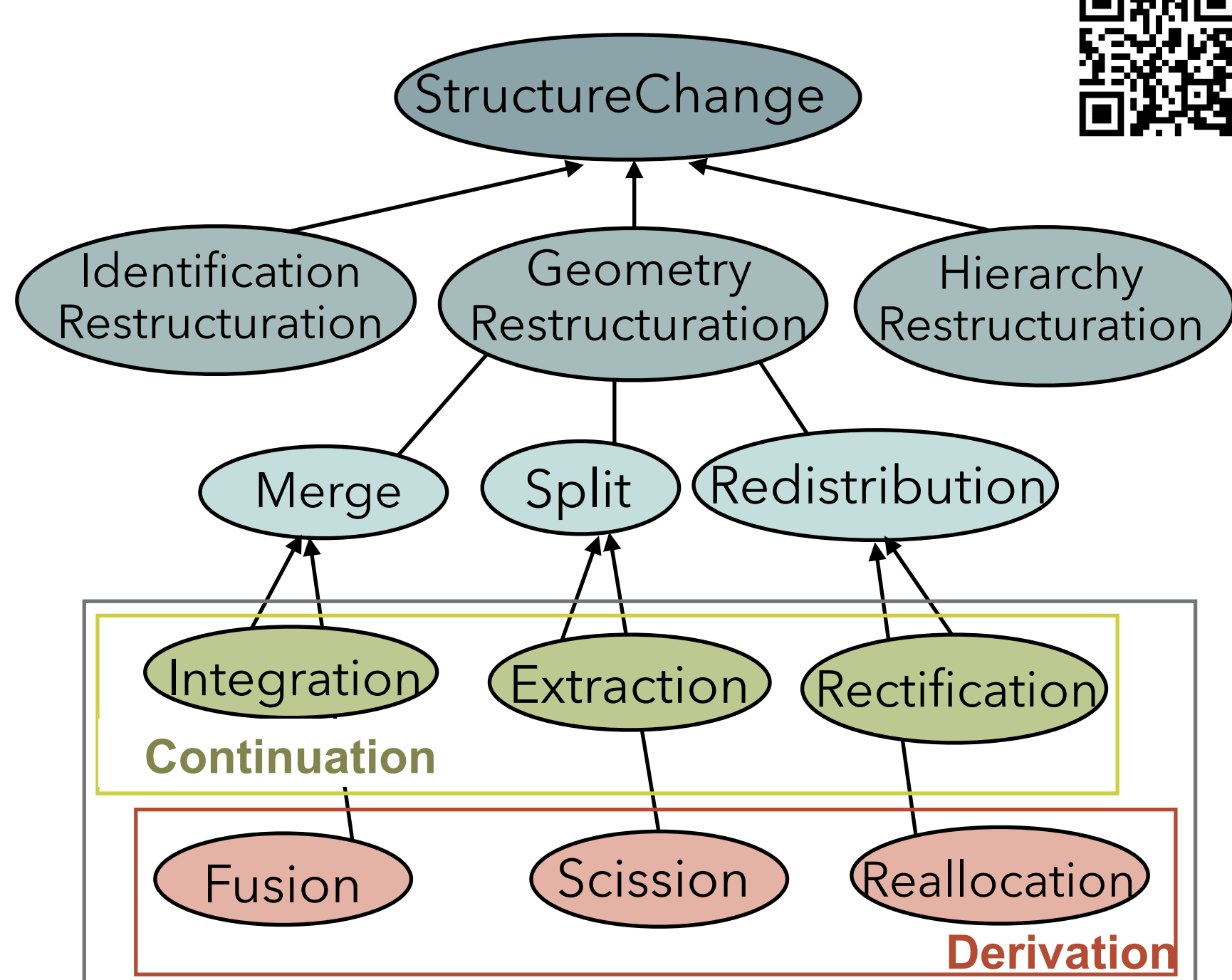
<http://purl.org/net/tsn#>



2 TSN-CHANGE ONTOLOGY

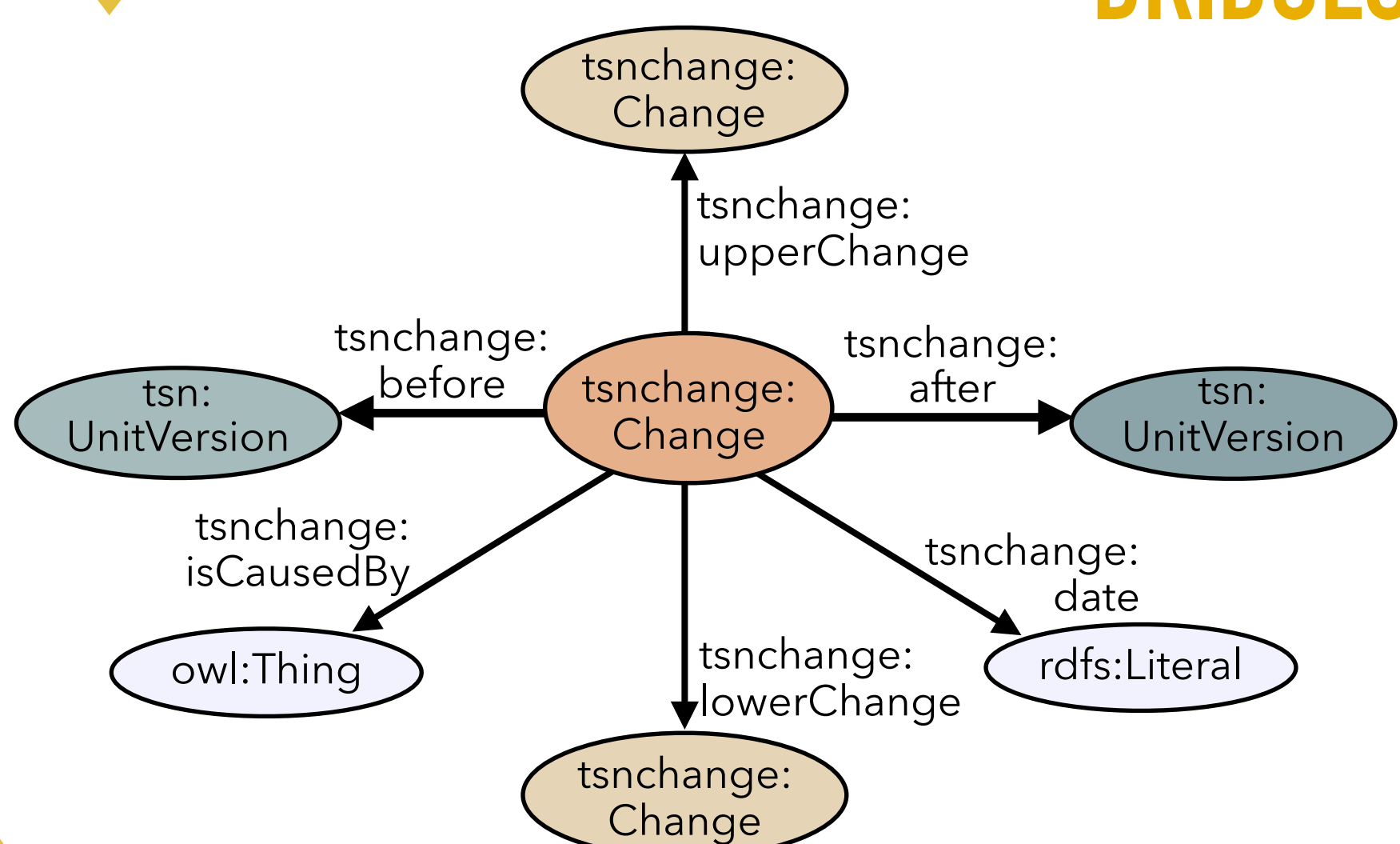
A set of tags to describe changes in TSN over time, in the LOD Web

<http://purl.org/net/tsnchange#>



3 MULTI-LEVEL CHANGE BRIDGES

A graph structure to navigate through the TSN versions + observe the propagation of a change event through the levels



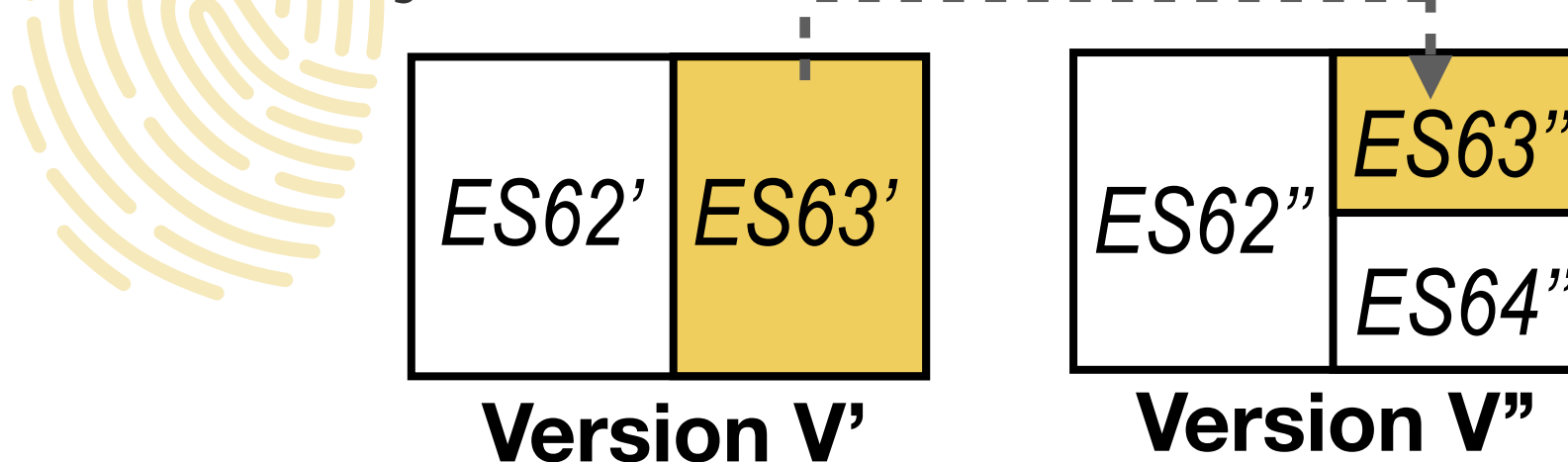
4 Generic algorithm for automatic matching and changes detection between two TSN's versions [Plumejeaud, 2010]

Input Two lists V' and V'' of units u' and u'' , all units being at one Level l .

Output A RDF graph describing Matching and Changes on Units

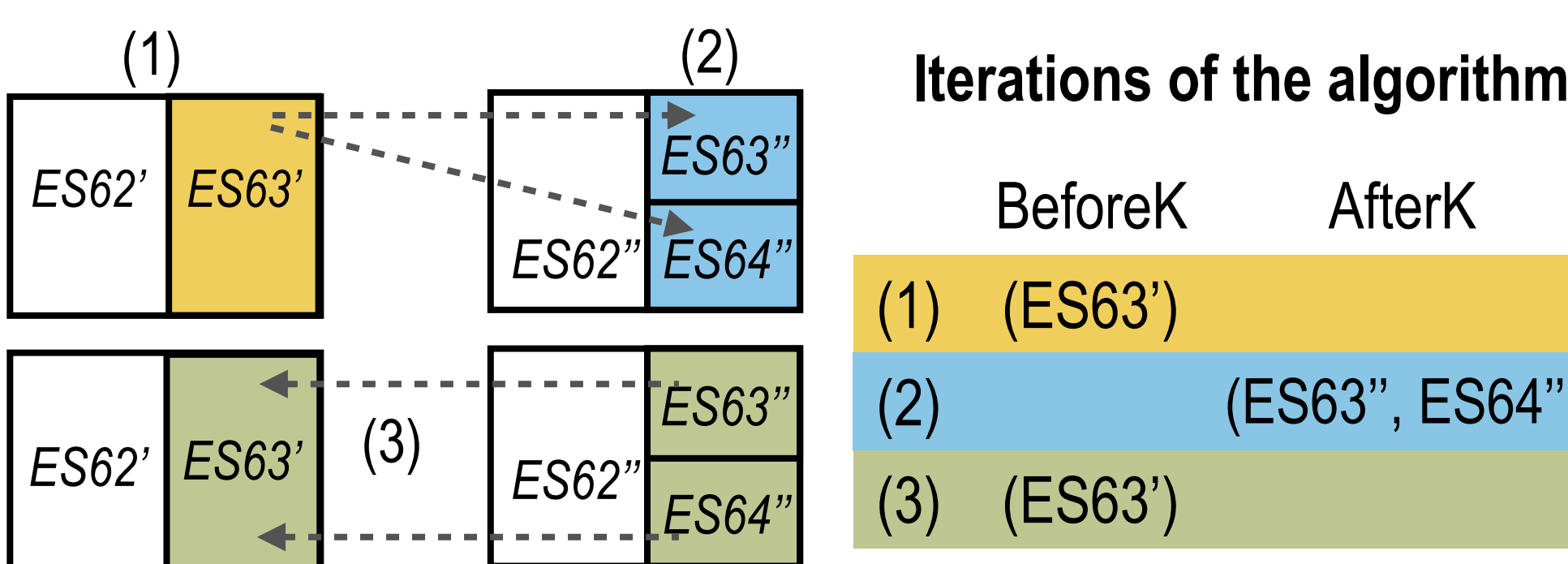
4.1 Matchings of territorial units

Which attributes may vary, in which proportion, before the entity loses its identity?



4.2 Identifying clusters of changes

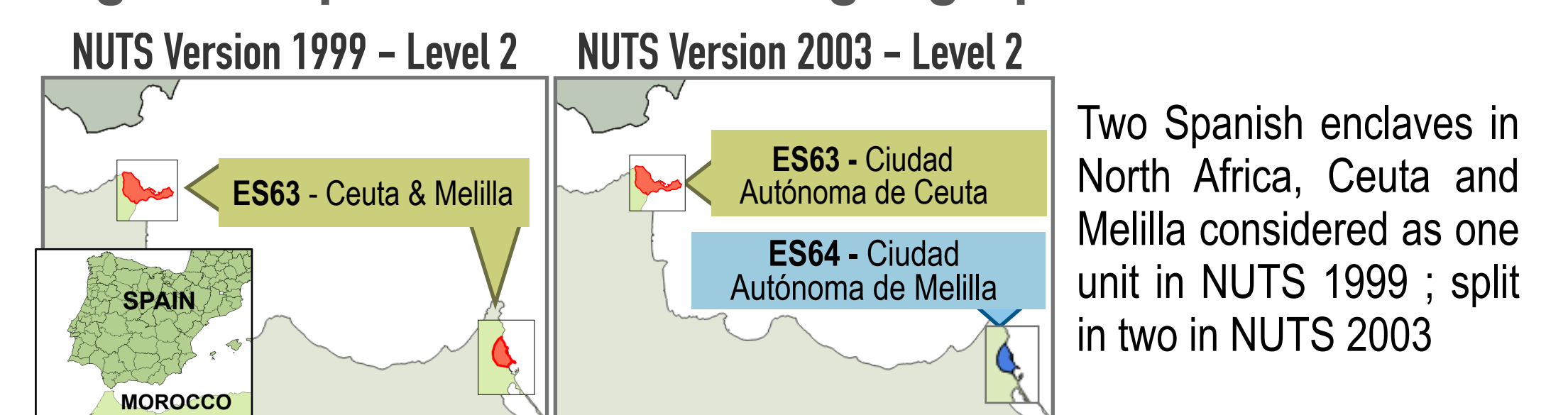
Builds through iterations two sets of units $BeforeK$ and $AfterK$. Adds successively to the sets, units that intersect ones in the other version. Stops when one set ($BeforeK$ or $AfterK$) remains identical between two iterations.



Uses relations between units in order to link changes through the TSN levels

ALGORITHM FOR CHANGE DETECTION

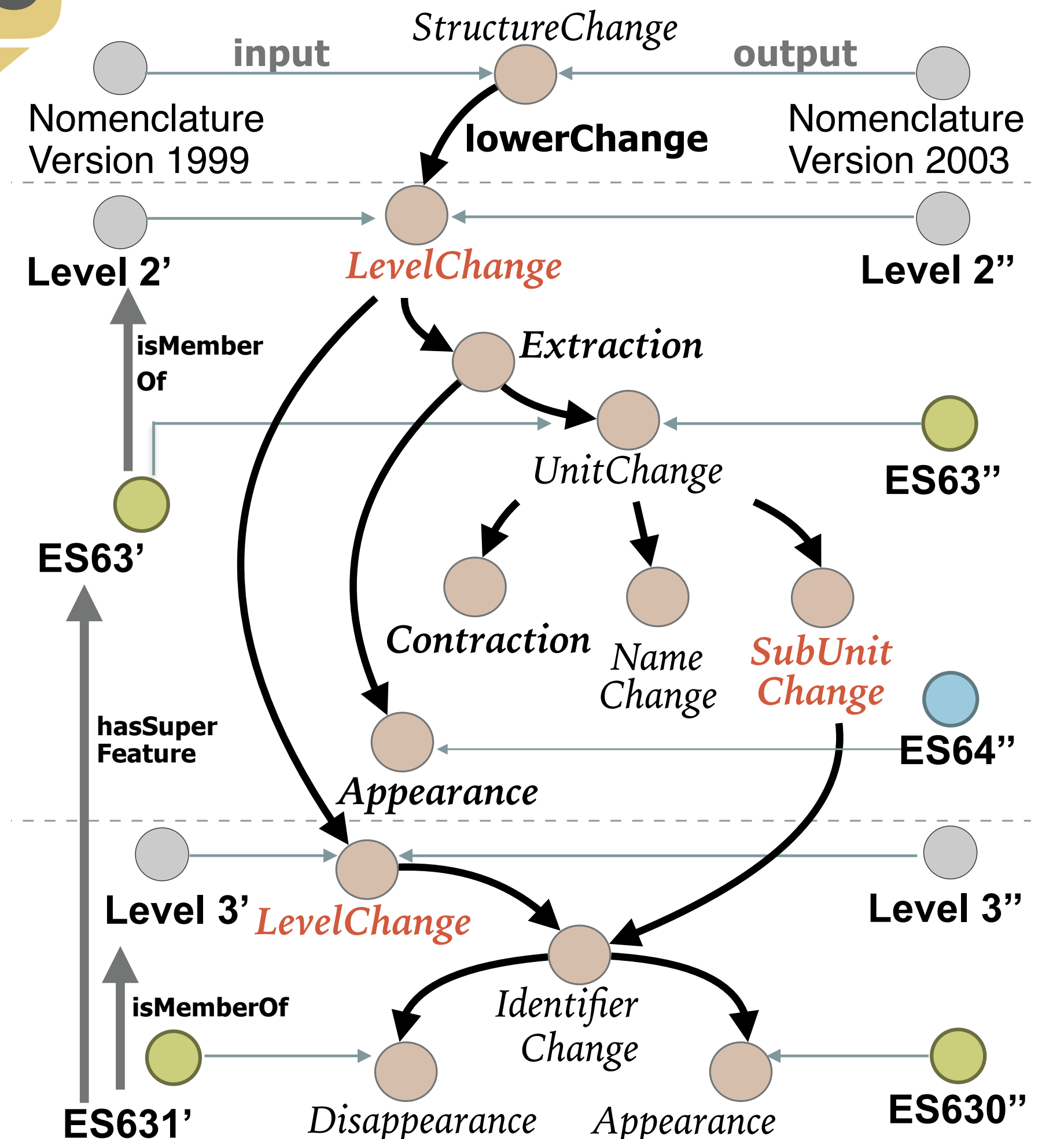
Change example in the EuroStat geographic nomenclature



Solution - configurable identity test in 3 steps:

1. An expert defines which attributes hold the identity of the units **ID? Name? Geom? Super Unit?**
2. For each attribute: a specific **Comparison Operator** is defined to measure distance between the two versions:
 $|u'.attribute - u''.attribute|$
3. Coefficients α_k assigned to attributes = define in which **proportion** attributes can vary before unit loses its identity. E.g:
 $F(\alpha_1, identifier), (\alpha_2, geom), (\alpha_3, name)$ where $\alpha_1 = 1/3, \alpha_2 = 1/3, \alpha_3 = 1/3$
 $1/3 * (|u'.identifier - u''.identifier|) +$
 $1/3 * (|u'.geom - u''.geom|) + 1/6 * (|u'.name - u''.name|)$

4.3 Linking changes through the levels



RESULTS

Automatize detection & semantic description of TSN changes: feature by feature + cluster of changes + chain of changes + 2 case studies

✓ **Europe** - Nomenclature of Territorial Units for Statistics (NUTS)

4 versions, 4 levels, 7619 geographic units

122,000 resources describing European Units

39,306 resources describing matching and changes

✓ **Australia** - Australian Statistical Geography Standard (ASGS)

2 versions, 4 levels, 5356 geographic units



<http://purl.org/steamer/nuts>

WHAT'S NEXT?

Geo-visualisation of Territorial Changes Over Time on the basis of GitHub, Inc. initiatives

