

WIKIDATA LAB XXXV

Modeling Elections

Wikidata Ontology



Who are we?



Éder Porto

Applied mathematician, programmer, designer and wikimedian



Amador Alvarez

Wikimedian, focused in Infoboxes & Wikidata.
former IT & Organization Director, now retired

What happened before infobox election?

Since 2016, the Catalan Wikipedia opted to use **Wikidata to feed the Infoboxes**.

- We had > 3.000 ultra-specialized manual infoboxes
- We used 5-6 different solutions for mapping
- Without any class of harmonization: layout, sizes, colors, etc.

We're 1.300 editors for 713.000 articles.

So, we decided to take advantage of this change to make a full re-engineering, reduce maintenance and be more efficient.

CAwiki Infobox status

cawiki articles:		713.118	a 22-10-2022			0			
CORE INFOBOXES				(15-08-2021)	LONG TAIL				
Group	#	total	% of cawiki	Previous %	Group	#	total	% of cawiki	
1	4	367.230	51,50%	51,70%	esport	3	9.030		
{{infotaula persona}}	1	179.792			geo & chem	4	8.449		
{{Infotaula geografia política}}	1	123.194			astro	4	5.694		
{{infotaula d'organització}}	1	32.872			social	6	5.615		
{{Indret}}	1	31.372			tech	7	5.372		
2	4	149.540	20,97%		20,73%	bio	9	4.828	
{{Infotaula d'ésser viu}}	1	70.619				politic & symbol	5	4.431	
{{infotaula edifici}}	1	46.117				divers	10	4.429	
{{Infotaula pel·lícula}}	1	17.610				militar	1	3.389	
{{Infotaula d'esdeveniment}}	1	15.194				music	4	3.306	
3	6	25.147	3,53%	2,30%	lang	5	3.100		
{{Infotaula obra musical}}	1	5.537			instr	8	2.477		
{{Infotaula de llibre}}	1	4.499			media	2	2.442		
{{Infotaula personatge}}	1	4.448			transport	3	2.302		
{{Infotaula d'objecte astronòmic}}	1	3.955			physics	5	786		
{{infotaula d'obra artística}}	1	3.844			econom	2	491		
{{Infotaula de publicacions periòdiques}}	1	2.864			Total Long tail	78	66.141	9,27%	
Total large used	14	541.917	75,99%	74,73%					
Total Long tail	78	66.141	9,27%	9,78%					

Data dictionary & Ontology

- A **data dictionary** is a centralized repository of information that describes the data or properties of entities in a specific information system, such as meaning, relationships with other data, origin, use, and format.
- **Ontology** is a type of representation in the form of a conceptual scheme that formulates a formal specification of the concepts contained in a domain or knowledge area, giving meaning to each of them and the relationships among them.
- In other words, if the dictionary describes the characteristics of the contents of one or more entities or tables, the ontology shows the forms of interrelationship of entities in the whole of the area of knowledge represented.

Data dictionary & Ontology

Example

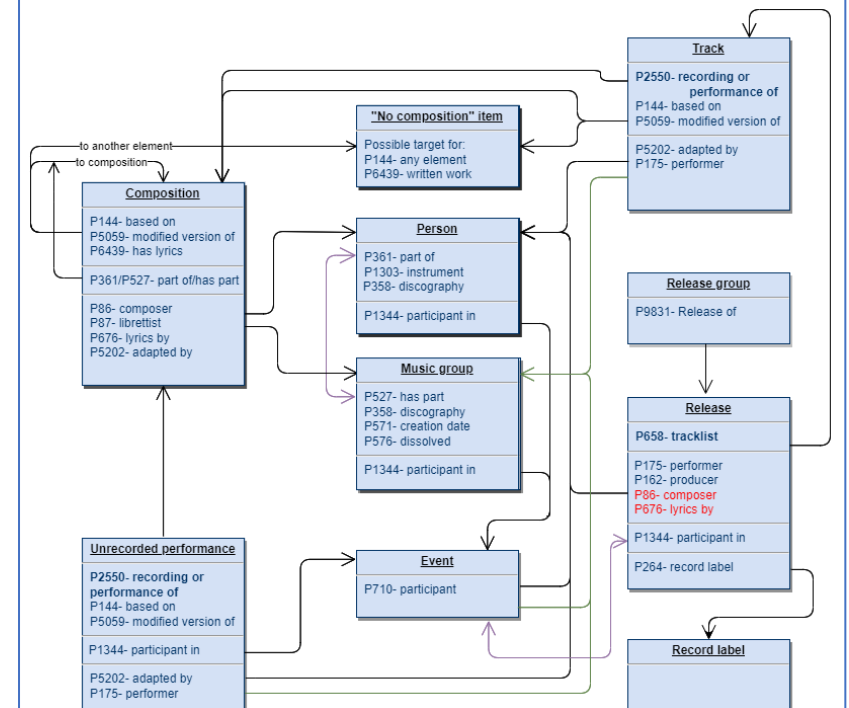
Track properties [modifica | modifica el codi]

These properties may apply to items with instància de (P31) of *pista musical* (Q7302866), *videoclip* (Q193977) or their subclasses.

Títol	ID	Tipus de dada	Descripció	Exemples	Inversa
àudio	P51	Fitxer multimèdia de Commons	fitxer d'àudio: fitxer de so	Estudi op. 25 < <i>àudio</i> > Chopin, Frédéric - Opus 25 - Twelve Grand Etudes - a minor (2).ogg	-
gènere artístic	P136	Element	gènere i by genre: gènere d'una obra, o gènere a què es dedica l'artista. No s'hauria d'utilitzar per relacionar obres amb els seus temes, per la qual cosa s'ha d'utilitzar el paràmetre "tema principal" (P921)	Night Train < <i>gènere artístic</i> > blues	-
sèrie	P179	Element	sèrie, sèrie, seqüència d'esdeveniments, sèrie de llibres i sèrie pictòrica: el subjecte és part d'una sèrie	Brown Sugar < <i>sèrie</i> > The Rolling Stones European Tour 1970	-
llengua del terme, de l'obra o del nom	P407	Element	llengua: llengua original (per a persones utilitzeu llengua materna P364 o llengua usada P1412)	Mazurek Dąbrowskiego < <i>llengua del terme, de l'obra o del nom</i> > polonès	-
productor	P162	Element	productor: productor de la pel·lícula o obra musical	Wuthering Heights < <i>productor</i> > Andrew Powell	-
data de publicació	P577	Punt en el temps	data de publicació: data en què es publicà l'obra	Quintet La Truita < <i>data de publicació</i> > 7 December 1820	-
tonalitat	P826	Element	clau i tonalitat: tonalitat principal d'una composició musical	Sonata per a piano núm. 31 < <i>tonalitat</i> > la bemoll major	-
classificat a	P2291	Element	llista d'èxits musicals: chart where the element reached a position	Ei Eerika pääse taivaaseen < <i>classificat a</i> > Suomen virallinen lista - Latauslista	-
propietari dels drets	P3931	Element	propietari dels drets i legal holder of rights for a creative work: person or organisation who holds the copyright of a work according to the Berne Convention	Love Me Do < <i>propietari dels drets</i> > Sony Music Publishing	-

WikiProject Music: Structure among entities

- Only properties related with link among entities are shown
- "Event" is not a music entity, but it may defines musical events.
- "No composition item" is not an entity, but a kind of valid items to point from P144/P6439 properties
- Property names in red need to be verified



Our case: Elections

What do we mean by "elections"?

- Election day?
- The act of voting?
- The election result: from my municipality, region, state?
- The entire electoral process?
-
- If we look at a WP “good article”, we find a set of many of these views.
- What should we have in WD?

What does an election item have ?

Modeling Elections

Amical Wikimedia

w.wiki/5rGD 25-10-22

9

Election
results

Vectors :

Electoral process

Current election

Pre-election
phase

Election
results

Post-election
phase

Vectors :

Electoral process

- preparation of applications,
- primaries within political parties,
- surveys,
- debates between candidates, etc.

Current election

Pre-election phase

Election results

Post-election phase

- use of the results
- transformation of votes into seats
- constitution of chambers
- voting for a president/mayor among elected members,
- etc.

Vectors :

Same type and territory series

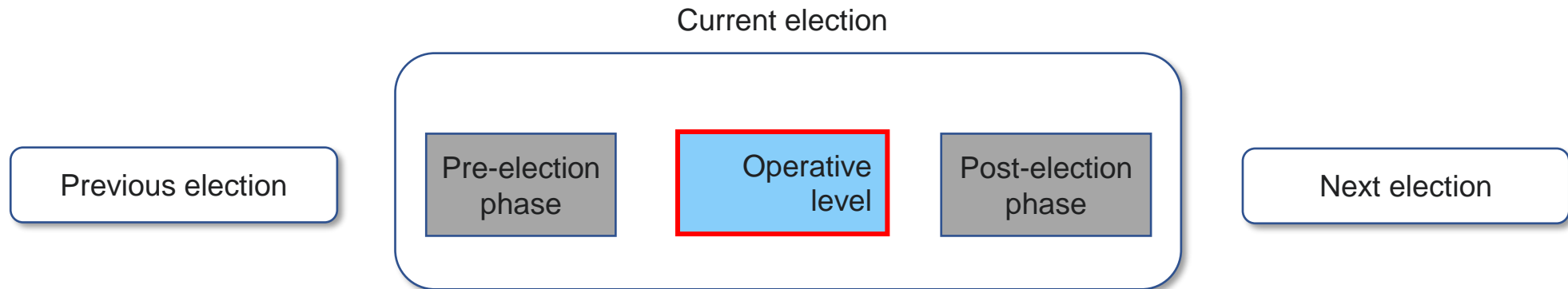
It is a chronic event:

- It is part of a series of elections that are repeated in the same territory
- It is part of a typology of elections that happen everywhere (presidential, general, municipal, etc.)



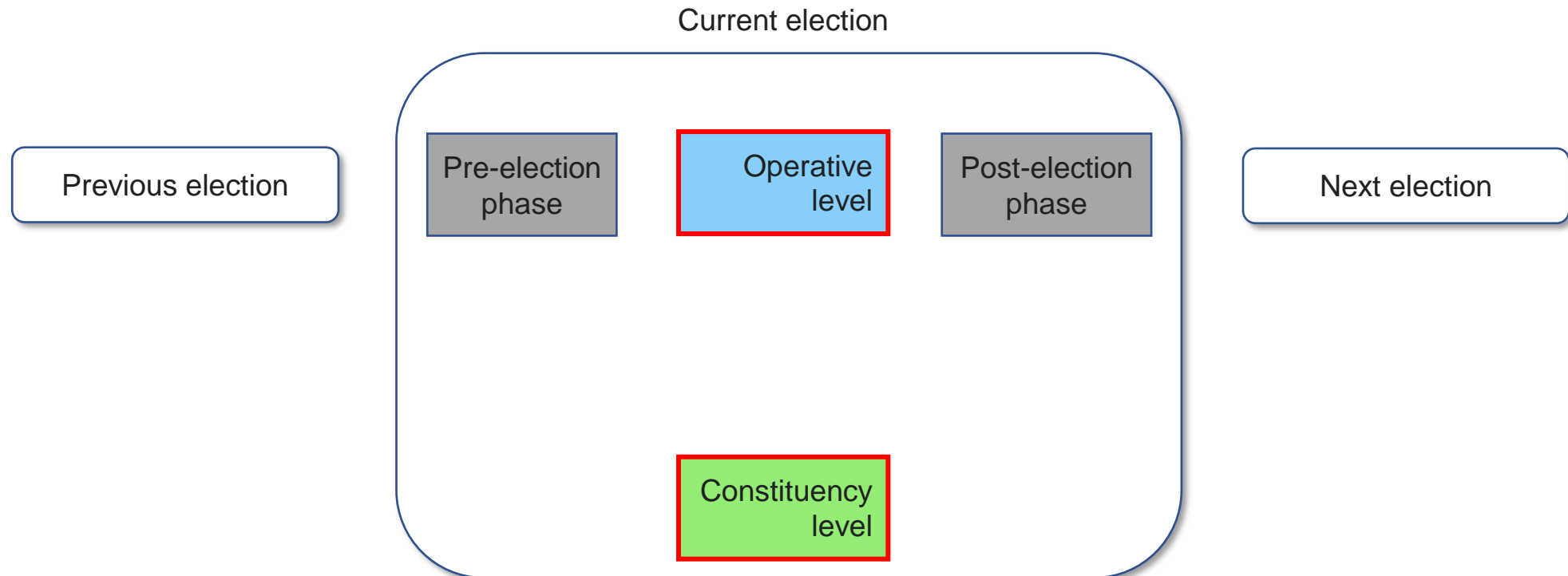
Vectors :

Presentation level → Operative



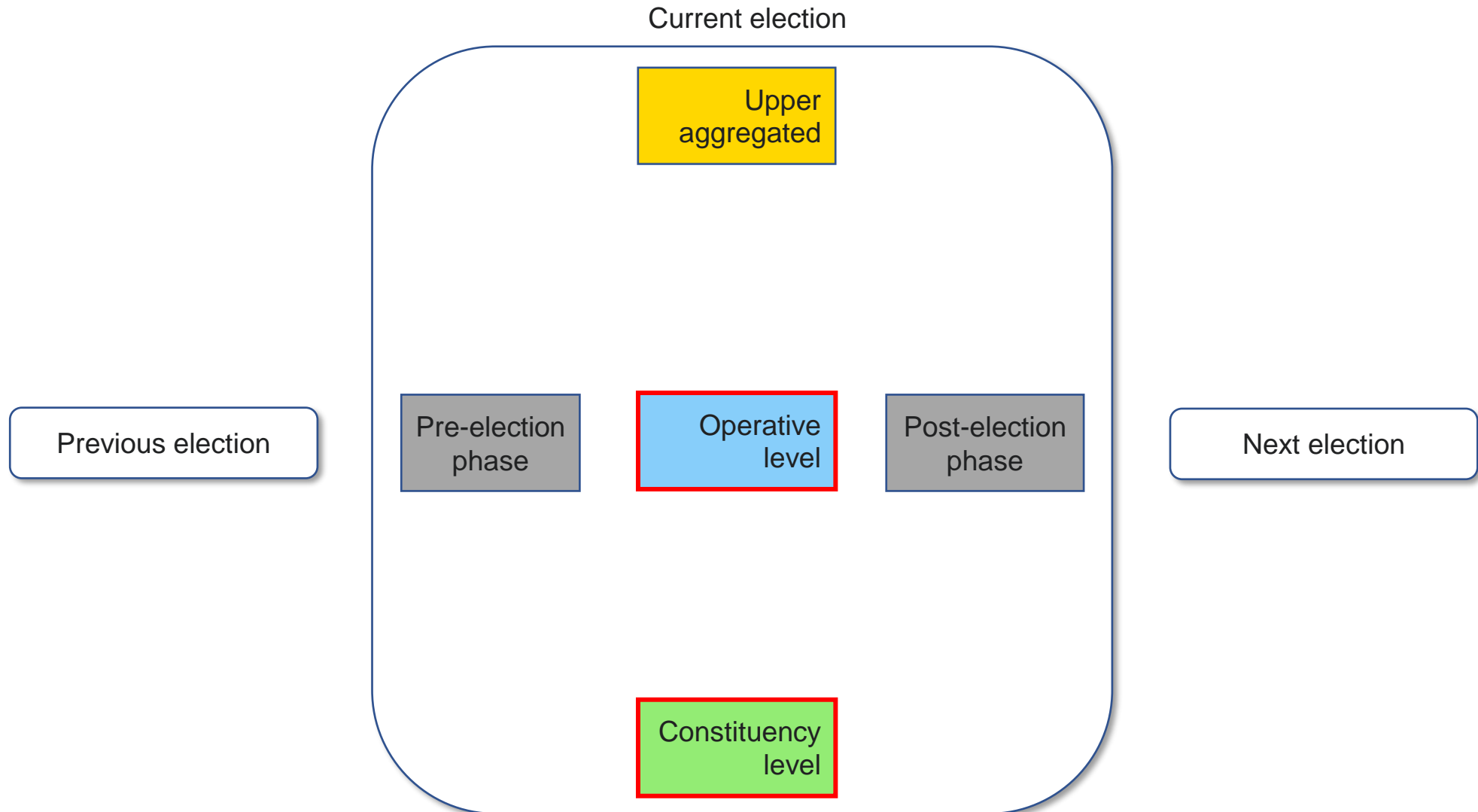
Vectors :

Presentation level \rightarrow Constituency



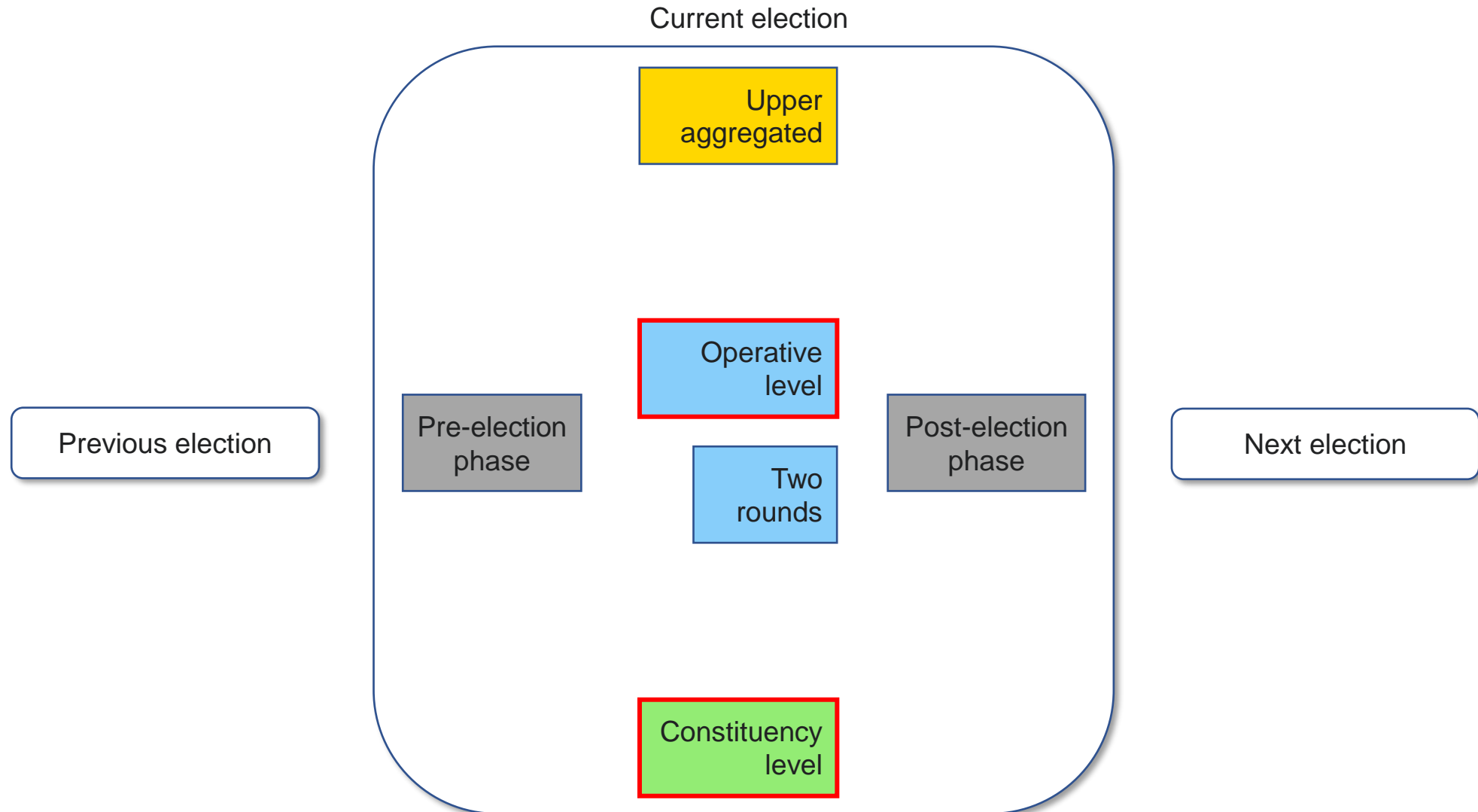
Vectors :

Upper aggregated



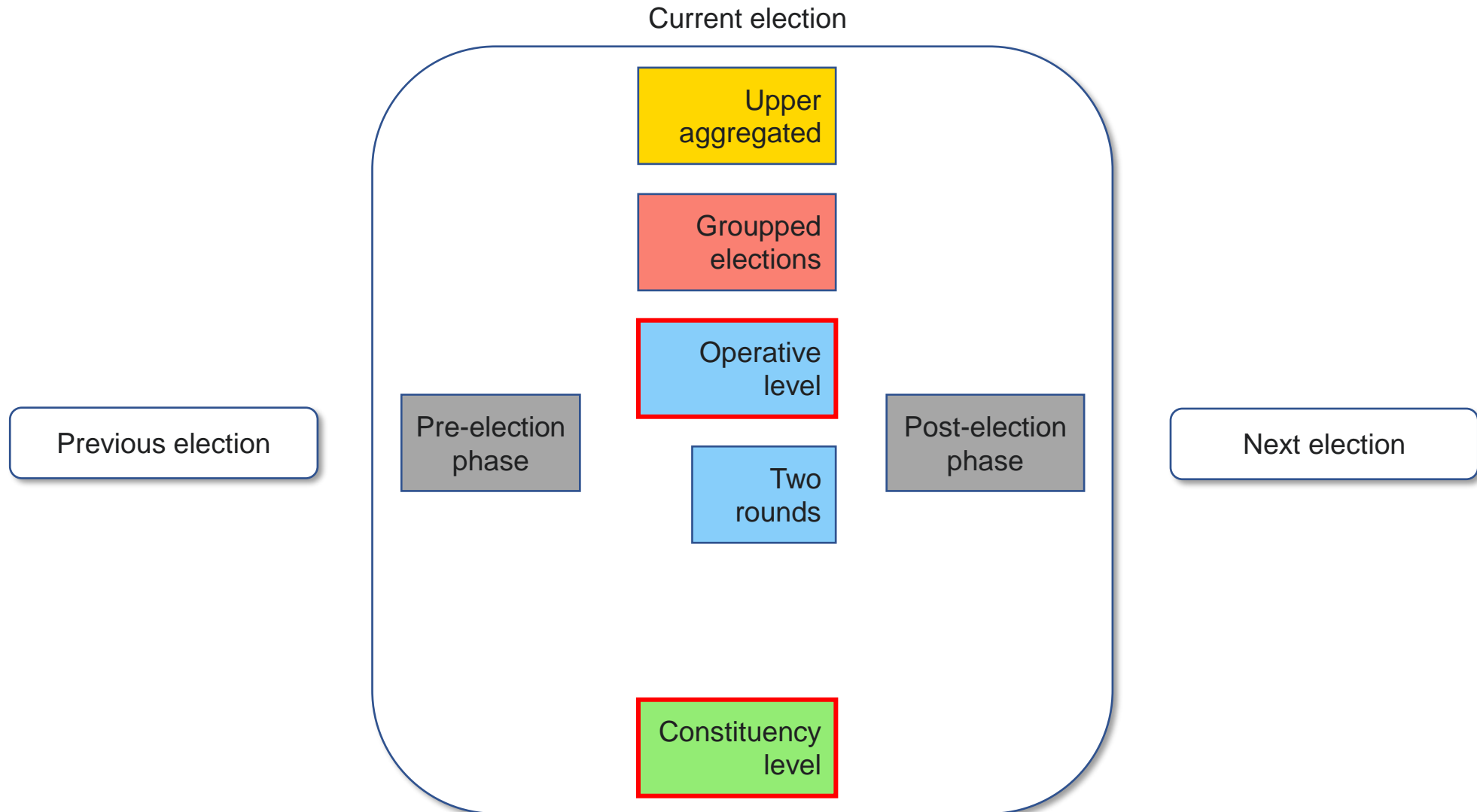
Vectors :

Two rounds

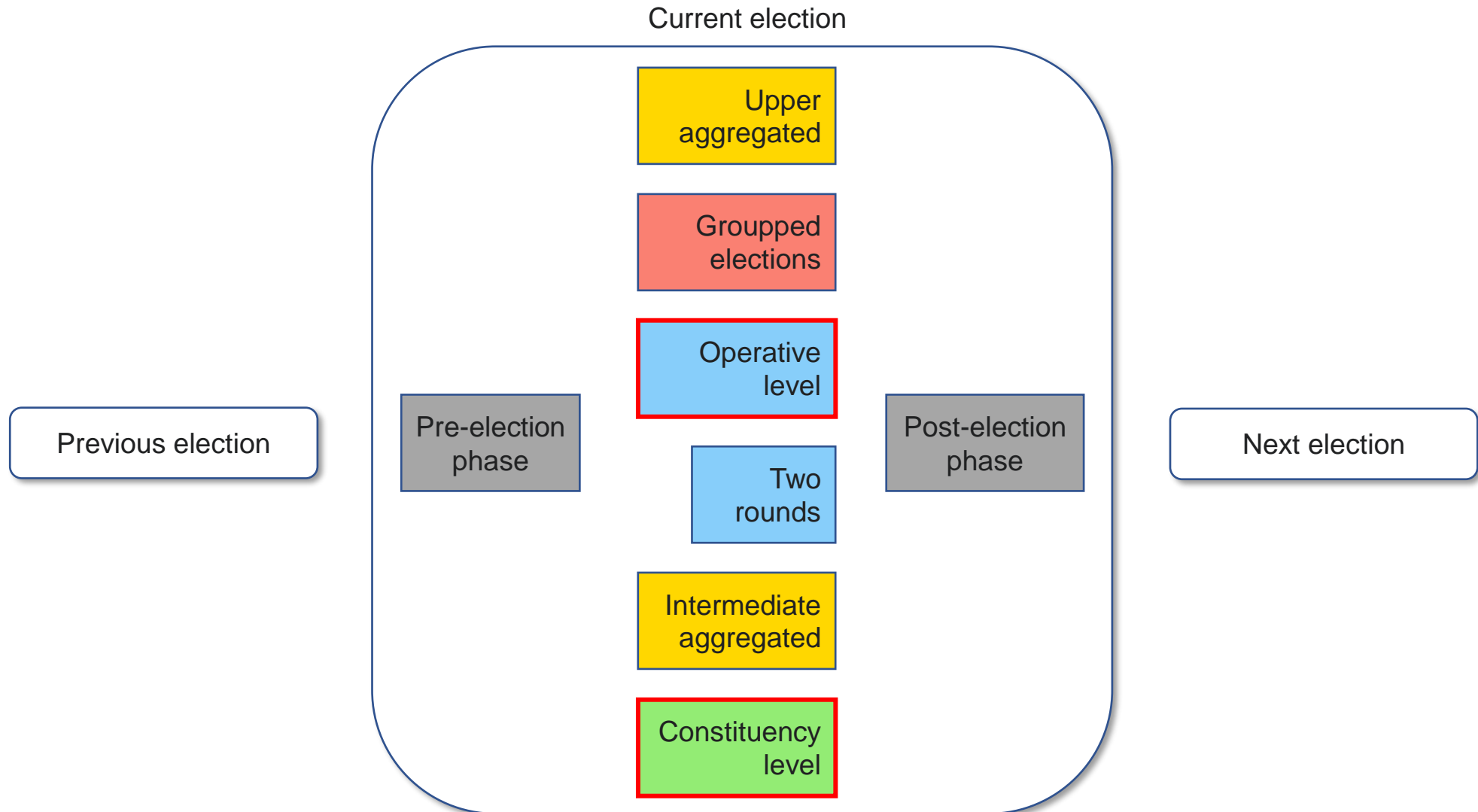


Vectors :

Grouped elections



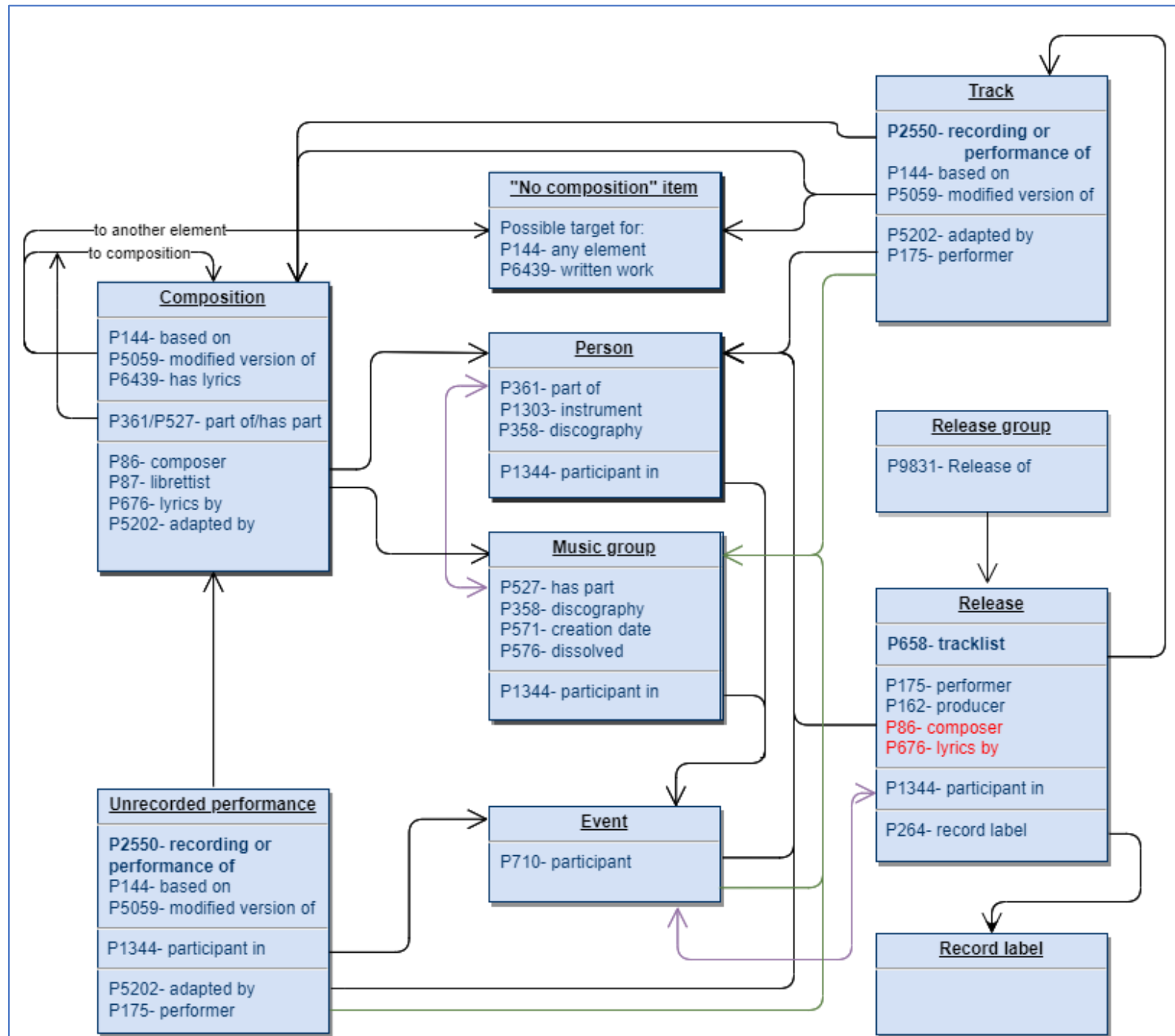
Vectors: Intermediate aggregated



Our case: Ontology

Ontology: Music

- Different entities
- With its own data dictionary
- Linked thru specific properties



Ontology: Elections

- The unique entity is **Results**
- 90% similar data dictionary
- Linked thru generic properties

[Full ontology document](#)

Each kind of items contain the whole specific information of its level.

They are autonomous without depending on the existence of other level items of the same election.

- 1) **Green & Blue** levels should exist. Sometimes are the same (i.e. local elections)
- 2) **Yellow** levels are optional and can be multiple
- 3) **Salmon** level may exist in special double event.



Questions ?

Examples

Modeling Elections

Amical Wikimedia

[w.wiki/5rGD](https://www.wiki/5rGD) 25-10-22

23

[Template:Infobox_election/example_intro](#)

Track properties [\[modifica \]](#) [\[modifica el codi \]](#)



These properties may apply to items with instància de (P31) of *pista musical* (Q7302866), *videoclip* (Q193977) or their subclasses.

Títol	ID	Tipus de dada	Descripció	Exemples	Inversa
àudio	P51	Fitxer multimèdia de Commons	fitxer d'àudio: fitxer de so	Estudi op. 25 <àudio> Chopin, Frédéric - Opus 25 - Twelve Grand Etudes - a minor (2).ogg	-
gènere artístic	P136	Element	gènere i by genre: gènere d'una obra, o gènere a què es dedica l'artista. No s'hauria d'utilitzar per relacionar obres amb els seus temes, per la qual cosa s'ha d'utilitzar el paràmetre "tema principal" (P921)	Night Train <gènere artístic> blues	-
sèrie	P179	Element	sèrie, sèrie, seqüència d'esdeveniments, sèrie de llibres i sèrie pictòrica: el subjecte és part d'una sèrie	Brown Sugar <sèrie> The Rolling Stones European Tour 1970	-
llengua del terme, de l'obra o del nom	P407	Element	llengua: llengua original (per a persones utilitzeu llengua materna P364 o llengua usada P1412)	Mazurek Dąbrowskiego <llengua del terme, de l'obra o del nom> polonès	-
productor	P162	Element	productor: productor de la pel·lícula o obra musical	Wuthering Heights <productor> Andrew Powell	-
data de publicació	P577	Punt en el temps	data de publicació: data en què es publicà l'obra	Quintet La Truita <data de publicació> 7 December 1820	-
tonalitat	P826	Element	clau i tonalitat: tonalitat principal d'una composició musical	Sonata per a piano núm. 31 <tonalitat> la bemoll major	-
classificat a	P2291	Element	llista d'èxits musicals: chart where the element reached a position	Ei Eerika pääse taivaaseen <classificat a> Suomen virallinen lista - Latauslista	-
propietari dels drets	P3931	Element	propietari dels drets i legal holder of rights for a creative work: person or organisation who holds the copyright of a work according to the Berne Convention	Love Me Do <propietari dels drets> Sony Music Publishing	-



GRACIES

<https://ca.wikipedia.org/wiki/Usuari:Amadalvarez>
amadalvarez@yahoo.es

<w.wiki/5rGD>