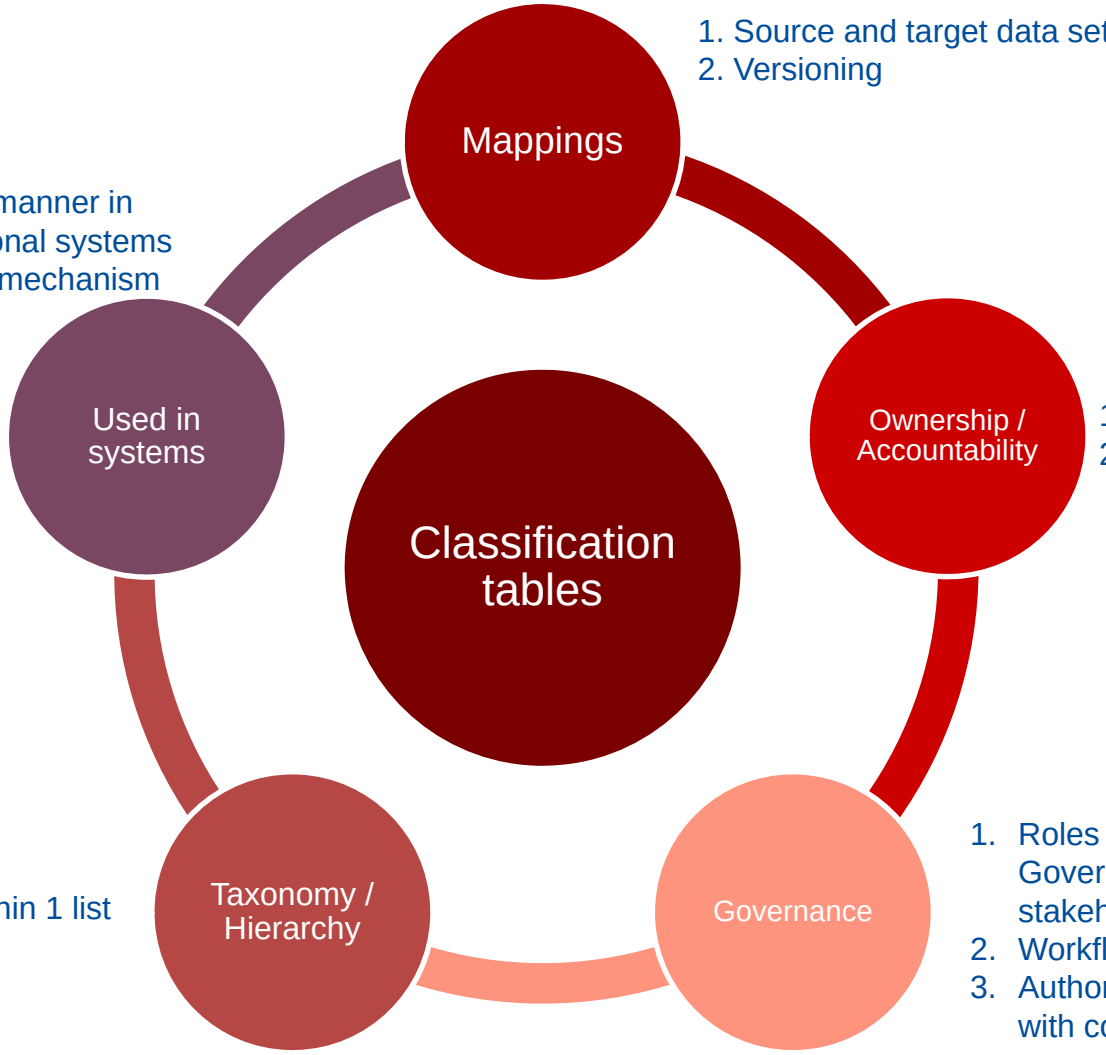
Semantics in Data Governance of Technology Classifications

EWI Focus - Gebruik van classificaties voor O&O-monitoring

The Challenges of Data Governance: Classification Tables



1. Source and target data sets
2. Versioning

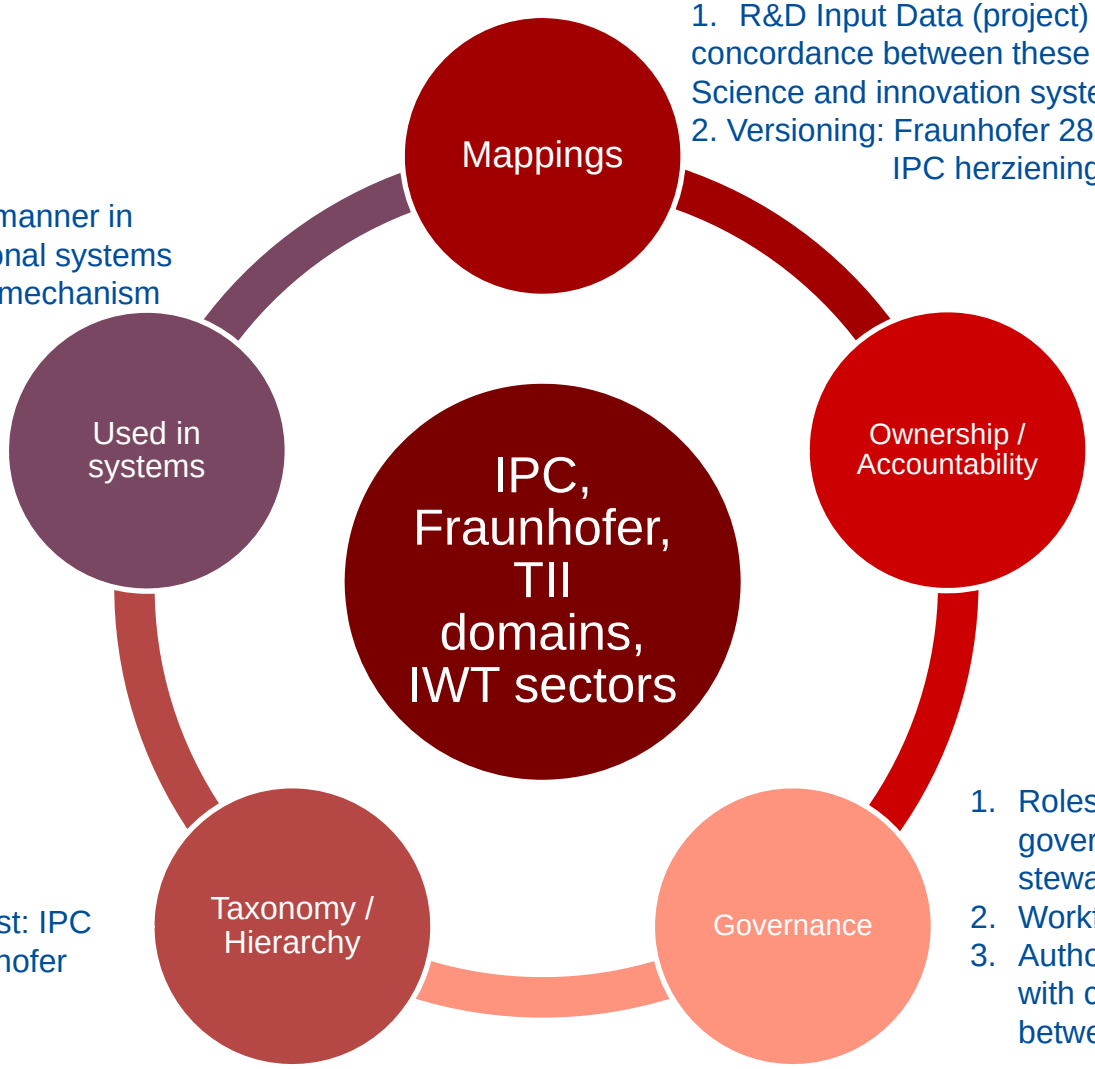
1. Can be defined internally
2. Can be specified externally

1. Roles and rights for data Governance (e.g. stewards – stakeholders)
2. Workflows
3. Authorities can provide us with conversion tables

1. Hierarchy within 1 list
2. Derived lists

1. Used in read-only manner in analytical and operational systems
2. Publish / subscribe mechanism
3. Import and export

The Challenges of Data Governance: Technology Classification Tables



1. R&D Input Data (project) vs. Output data (patent): concordance between these classifications => analysis of Science and innovation system
2. Versioning: Fraunhofer 28 (1992) ➔ Fraunhofer 35 (2006)
   IPC herziening

1. Used in read-only manner in analytical and operational systems
2. Publish / subscribe mechanism

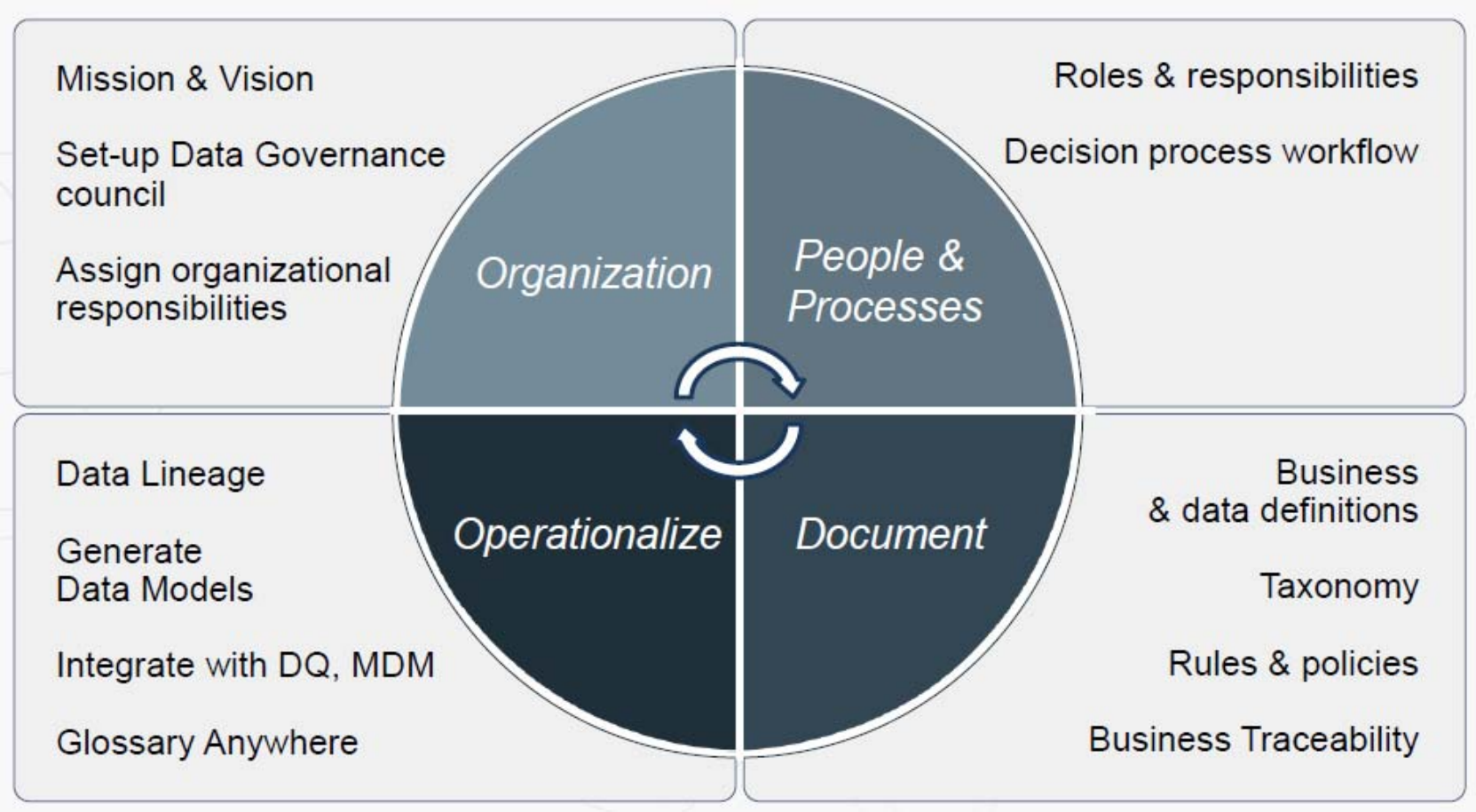
1. Can be defined internally (IWT – ECOOM)
2. Can be specified externally (International Patent Office)

1. Roles and rights in governance: stewards/stakeholders
2. Workflows
3. Authorities have provided us with conversion tables between IPC and Fraunhofer

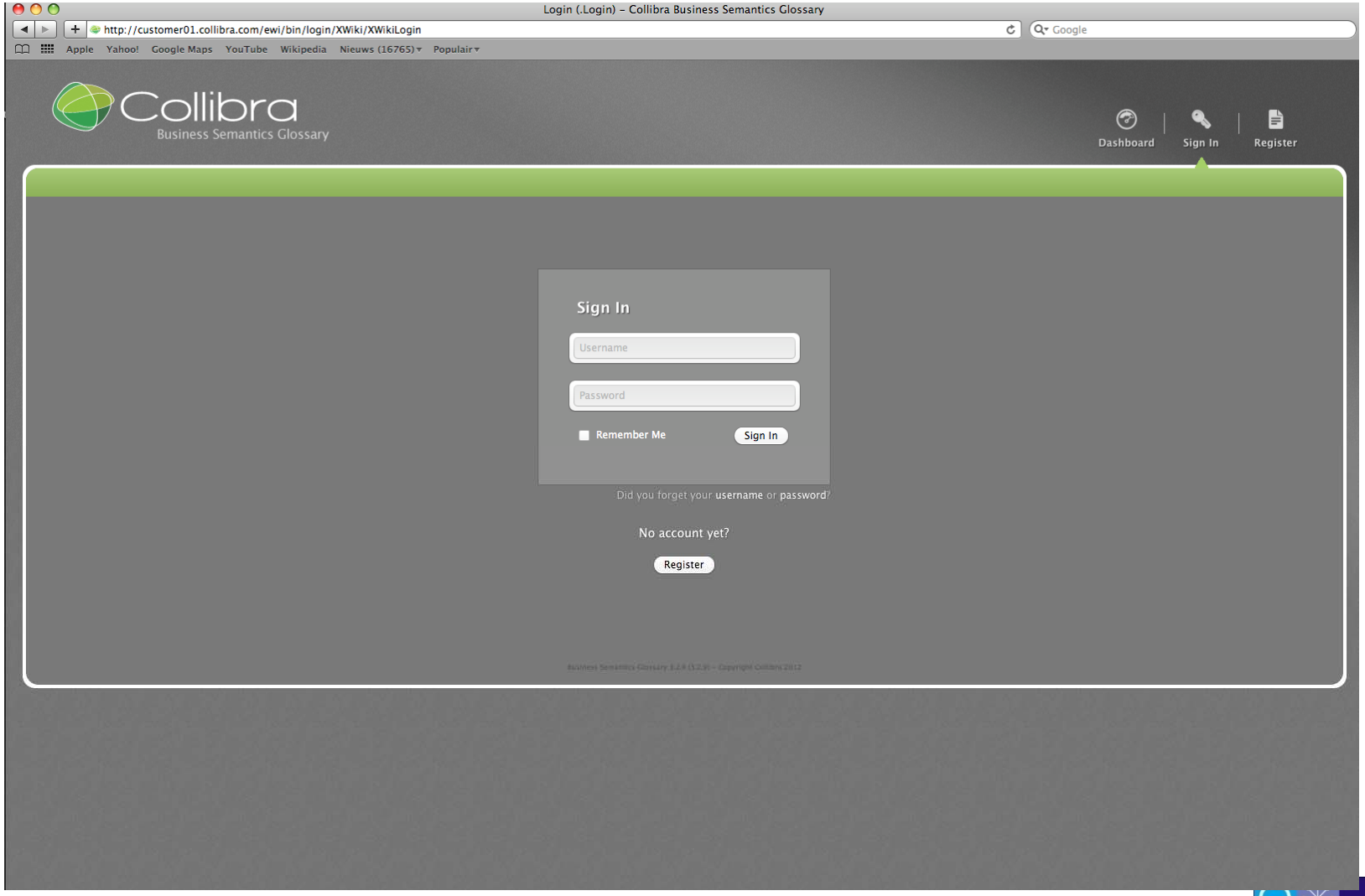
1. Hierarchy within 1 list: IPC
2. Derived lists: Fraunhofer

Semantics in Data Governance

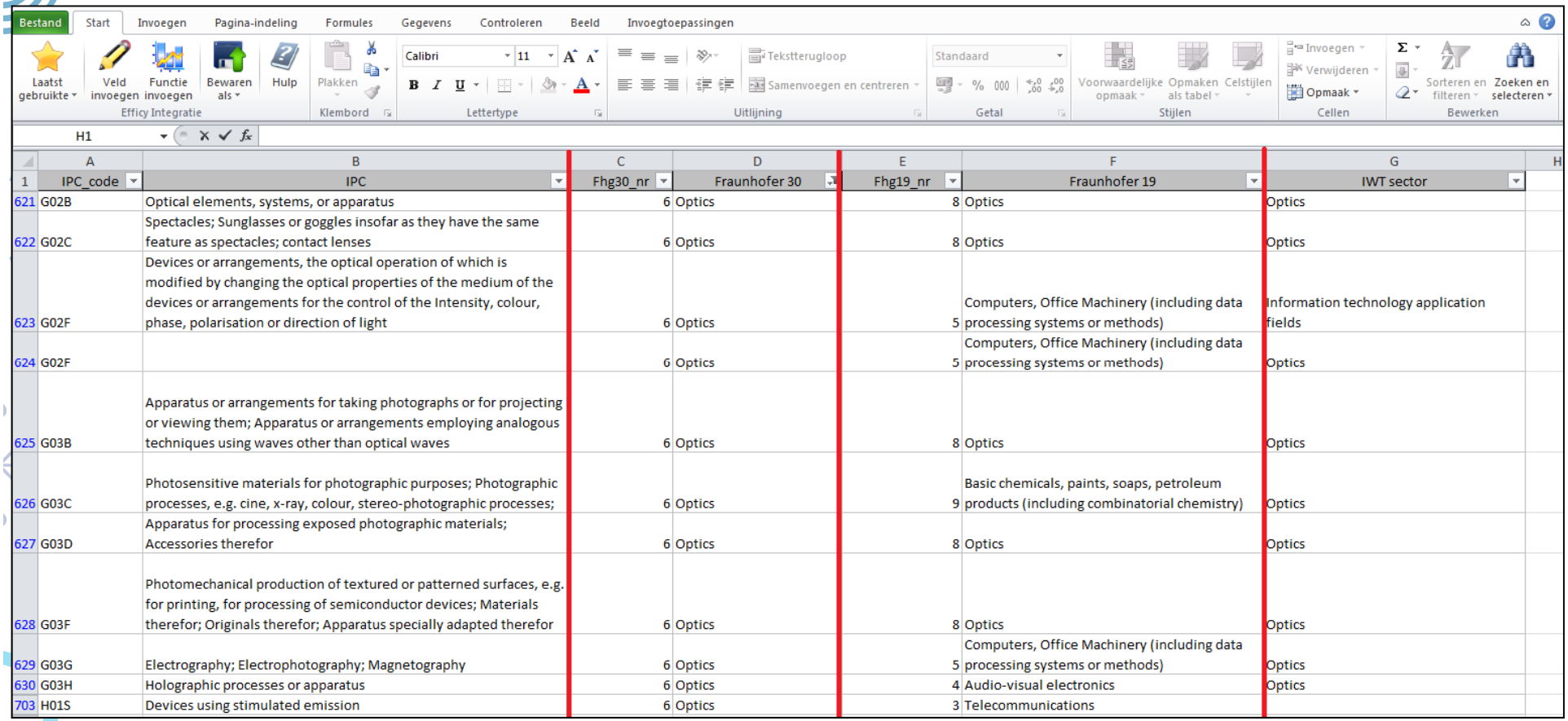
Four components of Semantics in Data Governance



Business Semantics Glossary (Collibra)

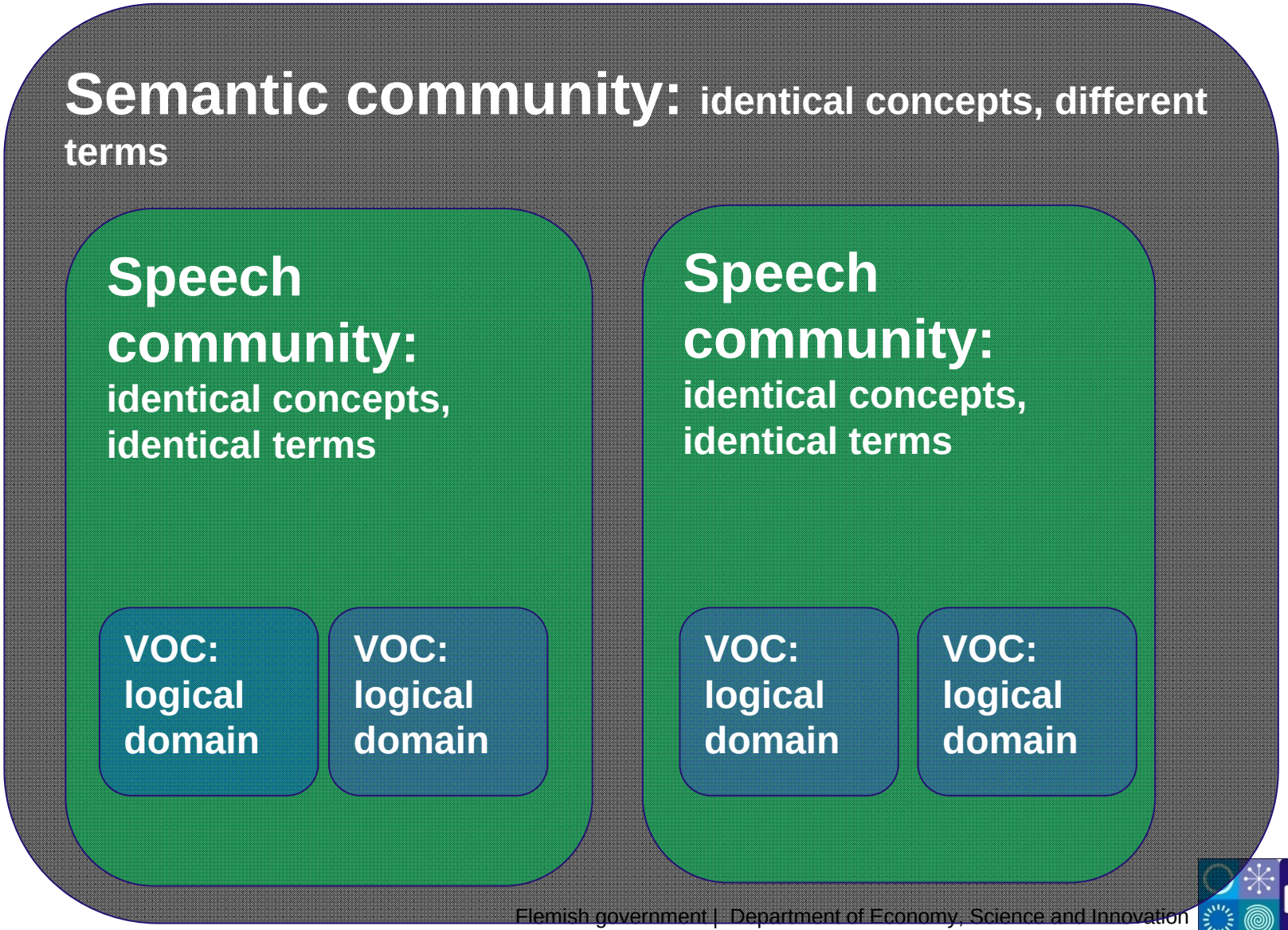


Business Case: Optics

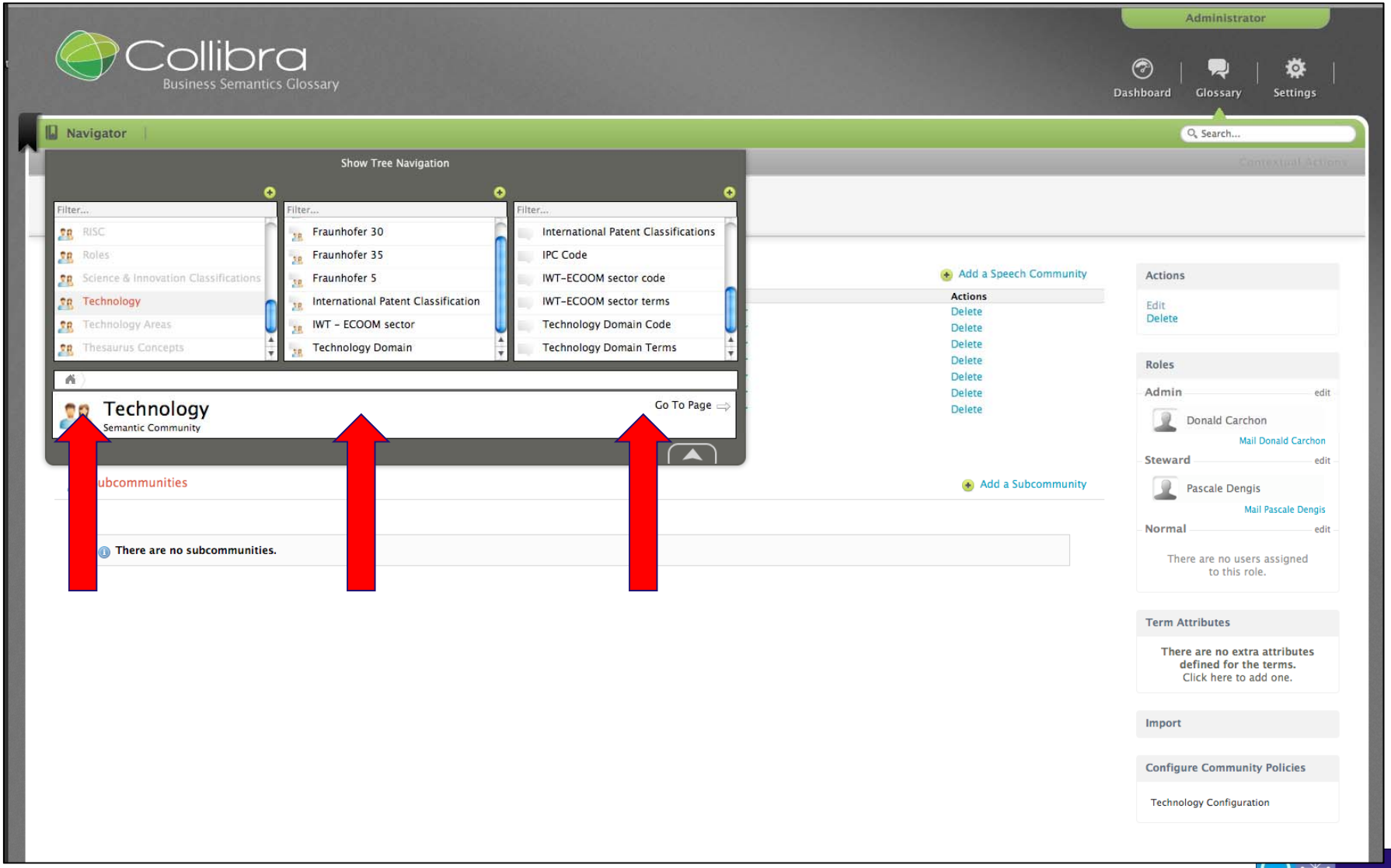


	A	B	C	D	E	F	G
1	IPC_code	IPC	Fhg30_nr	Fraunhofer 30	Fhg19_nr	Fraunhofer 19	IWT sector
621	G02B	Optical elements, systems, or apparatus	6	Optics	8	Optics	Optics
622	G02C	Spectacles; Sunglasses or goggles insofar as they have the same feature as spectacles; contact lenses	6	Optics	8	Optics	Optics
623	G02F	Devices or arrangements, the optical operation of which is modified by changing the optical properties of the medium of the devices or arrangements for the control of the Intensity, colour, phase, polarisation or direction of light	6	Optics	5	Computers, Office Machinery (including data processing systems or methods)	Information technology application fields
624	G02F		6	Optics	5	Computers, Office Machinery (including data processing systems or methods)	Optics
625	G03B	Apparatus or arrangements for taking photographs or for projecting or viewing them; Apparatus or arrangements employing analogous techniques using waves other than optical waves	6	Optics	8	Optics	Optics
626	G03C	Photosensitive materials for photographic purposes; Photographic processes, e.g. cine, x-ray, colour, stereo-photographic processes;	6	Optics	9	Basic chemicals, paints, soaps, petroleum products (including combinatorial chemistry)	Optics
627	G03D	Apparatus for processing exposed photographic materials; Accessories therefor	6	Optics	8	Optics	Optics
628	G03F	Photomechanical production of textured or patterned surfaces, e.g. for printing, for processing of semiconductor devices; Materials therefor; Originals therefor; Apparatus specially adapted therefor	6	Optics	8	Optics	Optics
629	G03G	Electrography; Electrophotography; Magnetography	6	Optics	5	Computers, Office Machinery (including data processing systems or methods)	Optics
630	G03H	Holographic processes or apparatus	6	Optics	4	Audio-visual electronics	Optics
703	H01S	Devices using stimulated emission	6	Optics	3	Telecommunications	

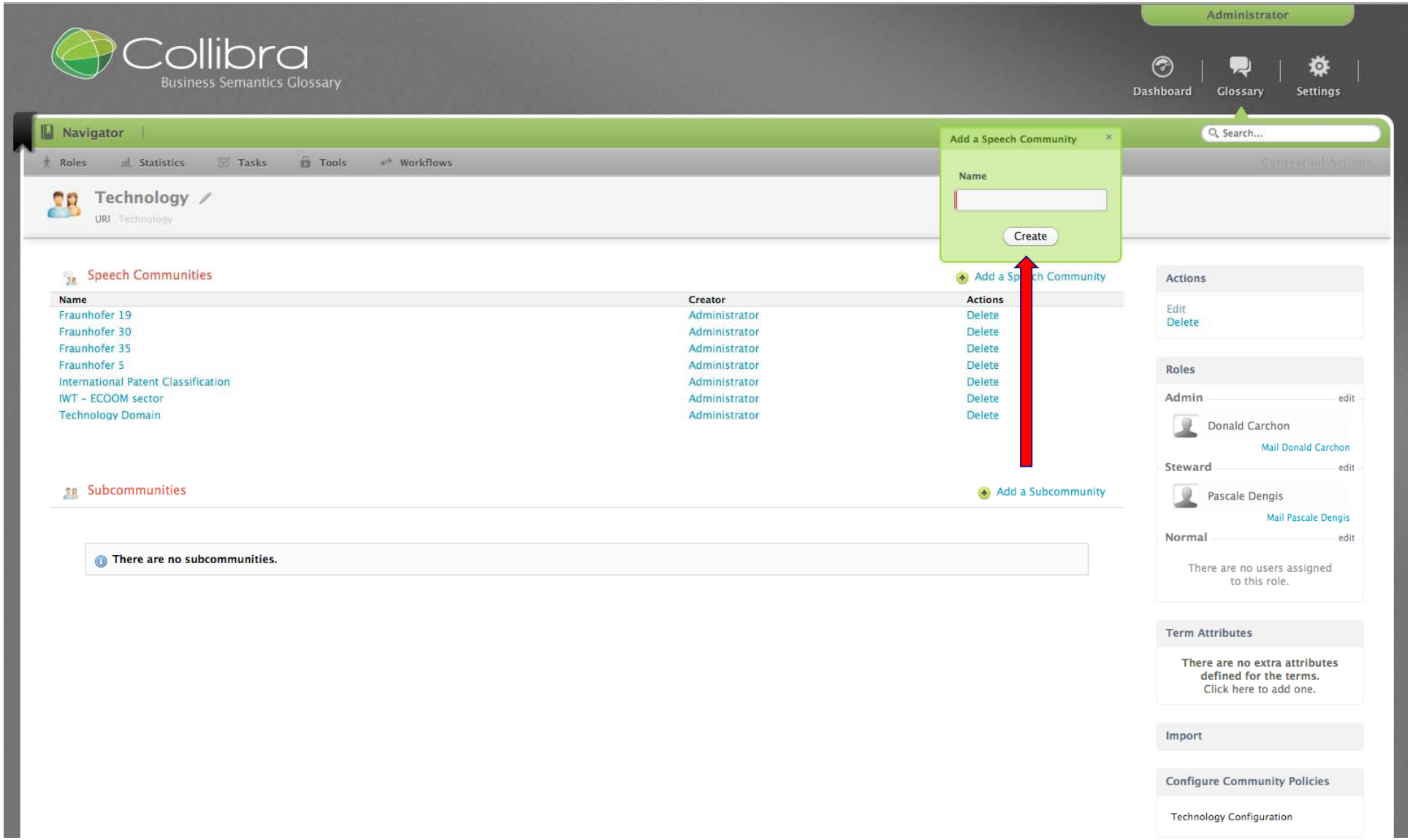
Governance – SBVR (OMG Standaard)



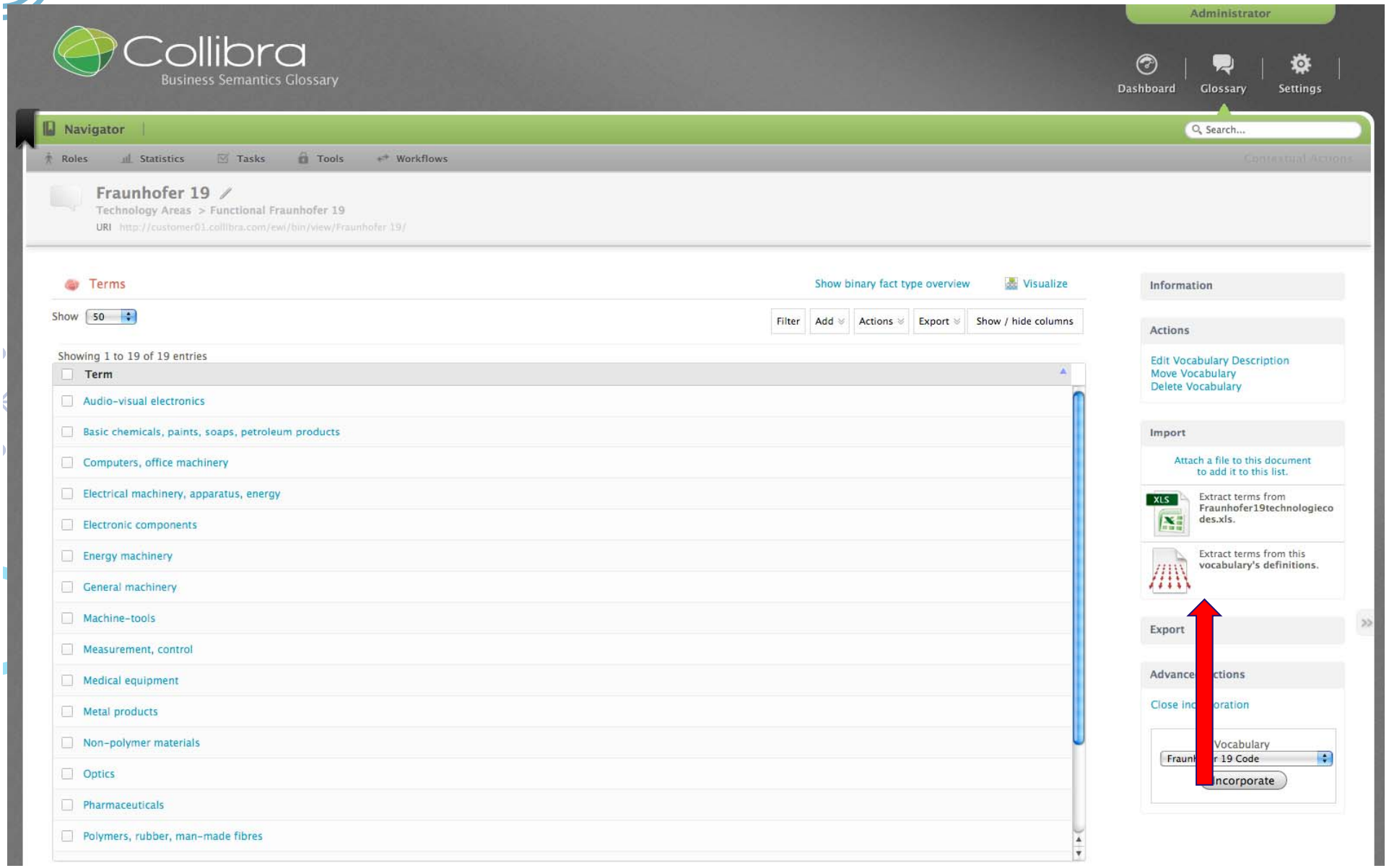
How to Organise Technology Community? Semantic community



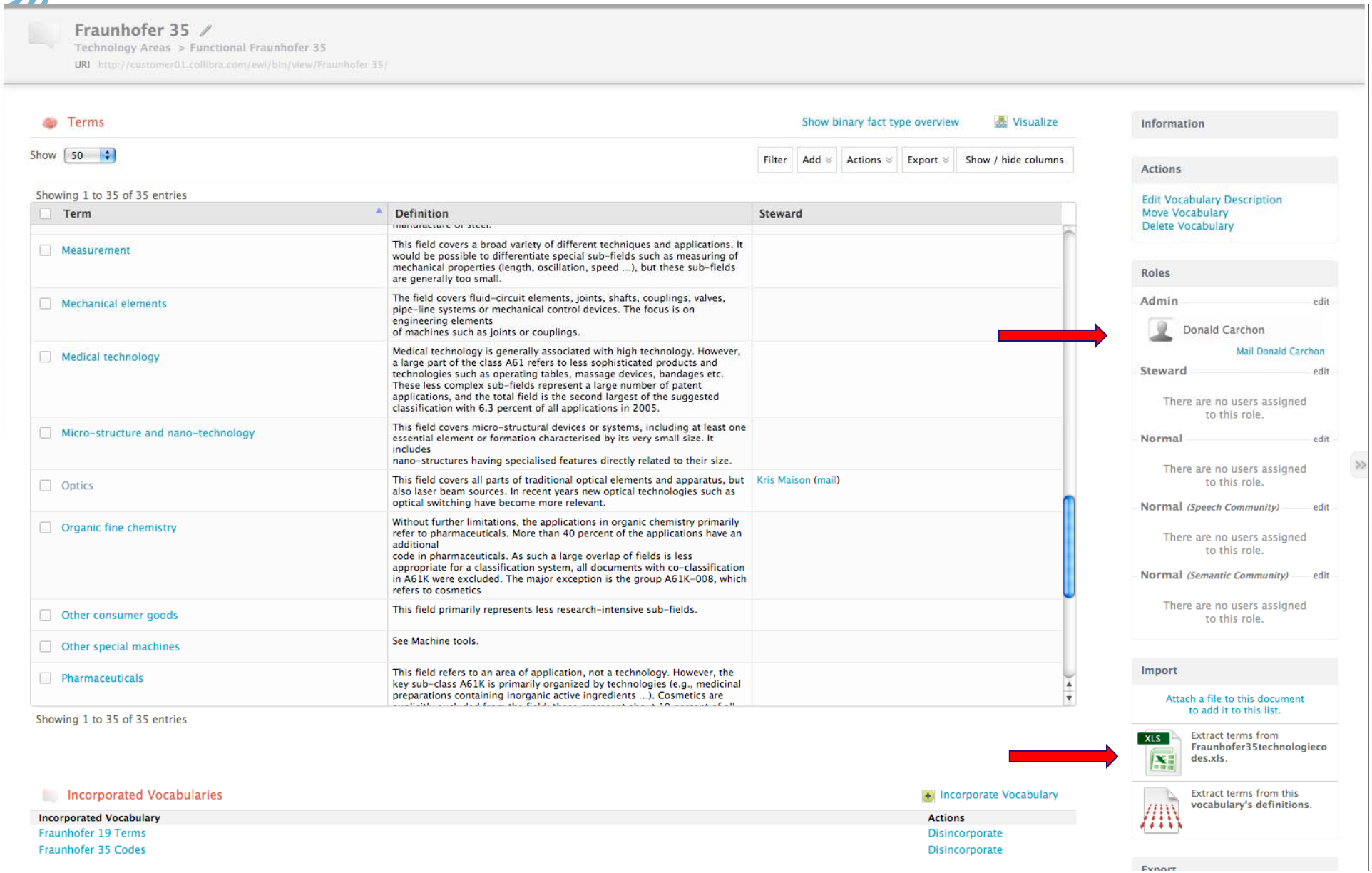
How to Organise Technology Community? Speech Community



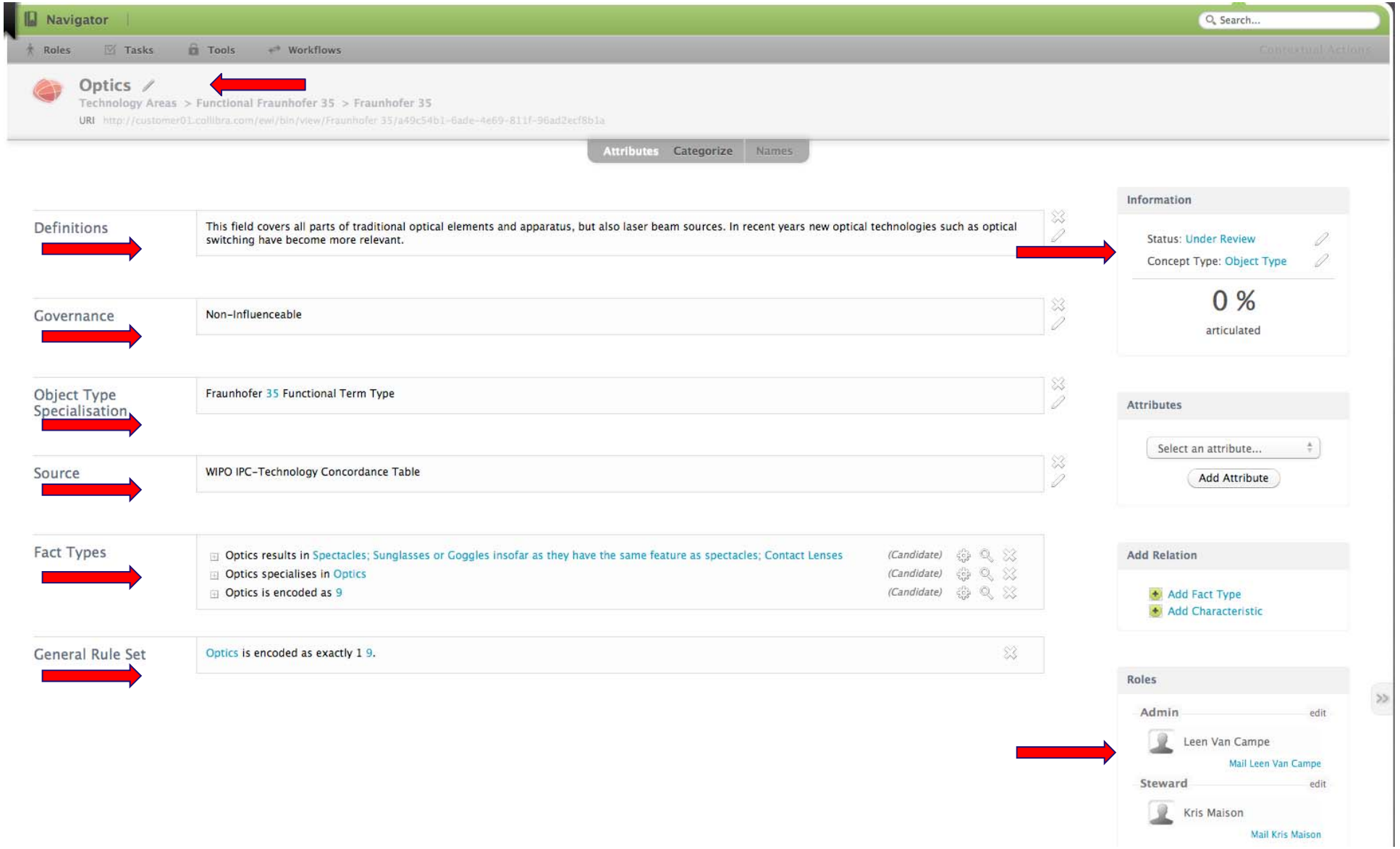
How to Organise Technology Community? Vocabulary



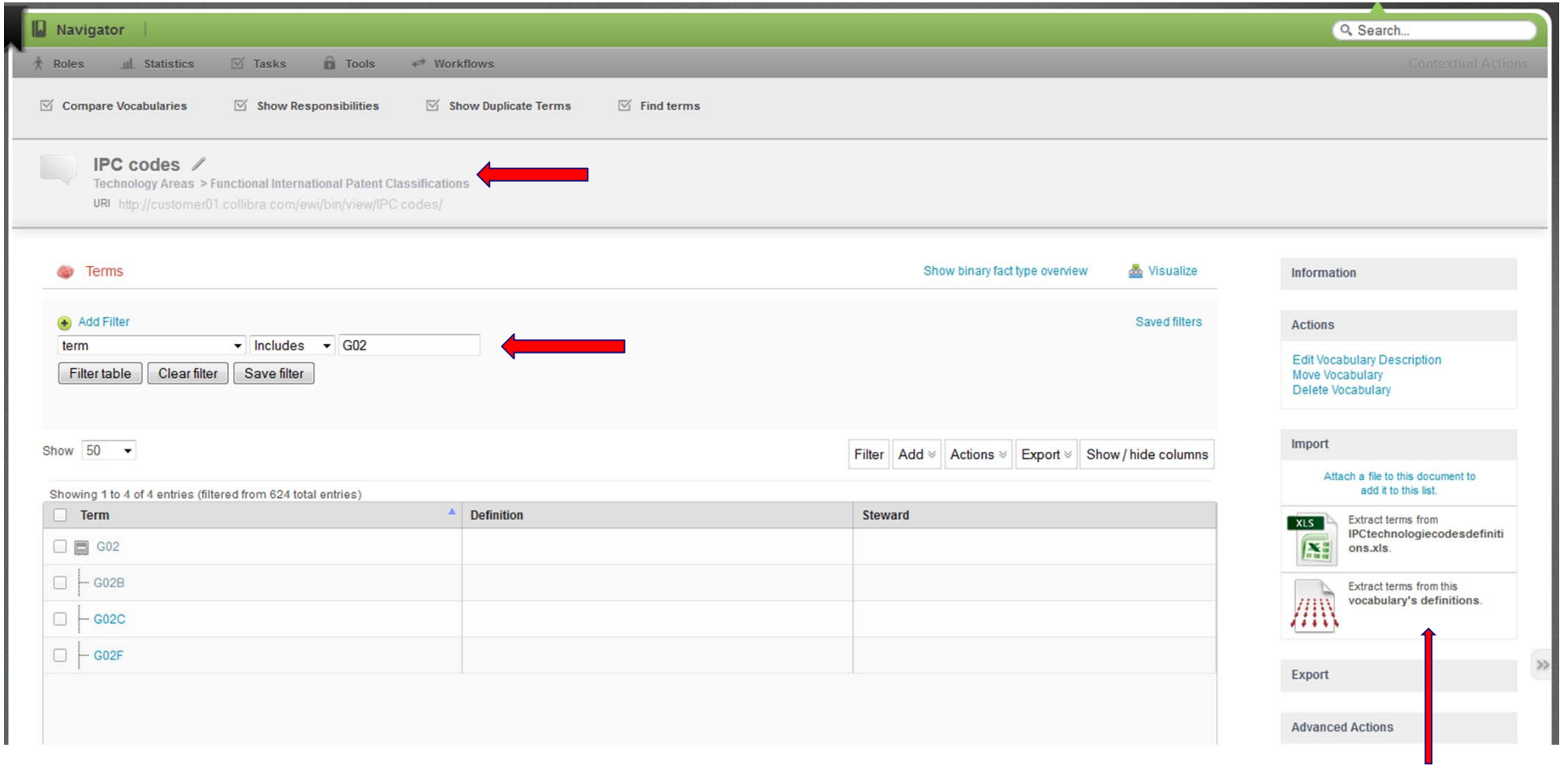
Document: Terms and Definitions



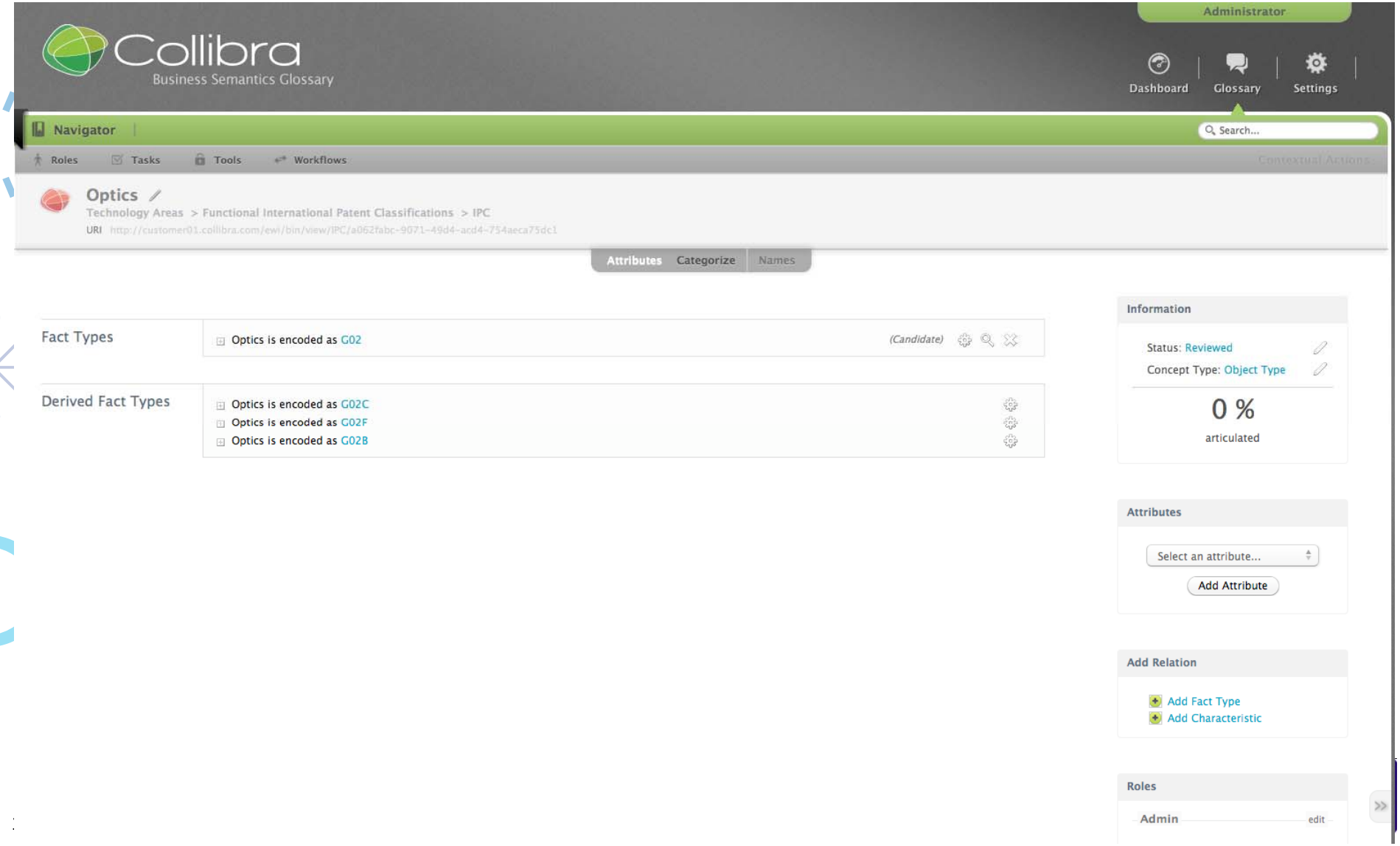
Document: Terms and Definitions



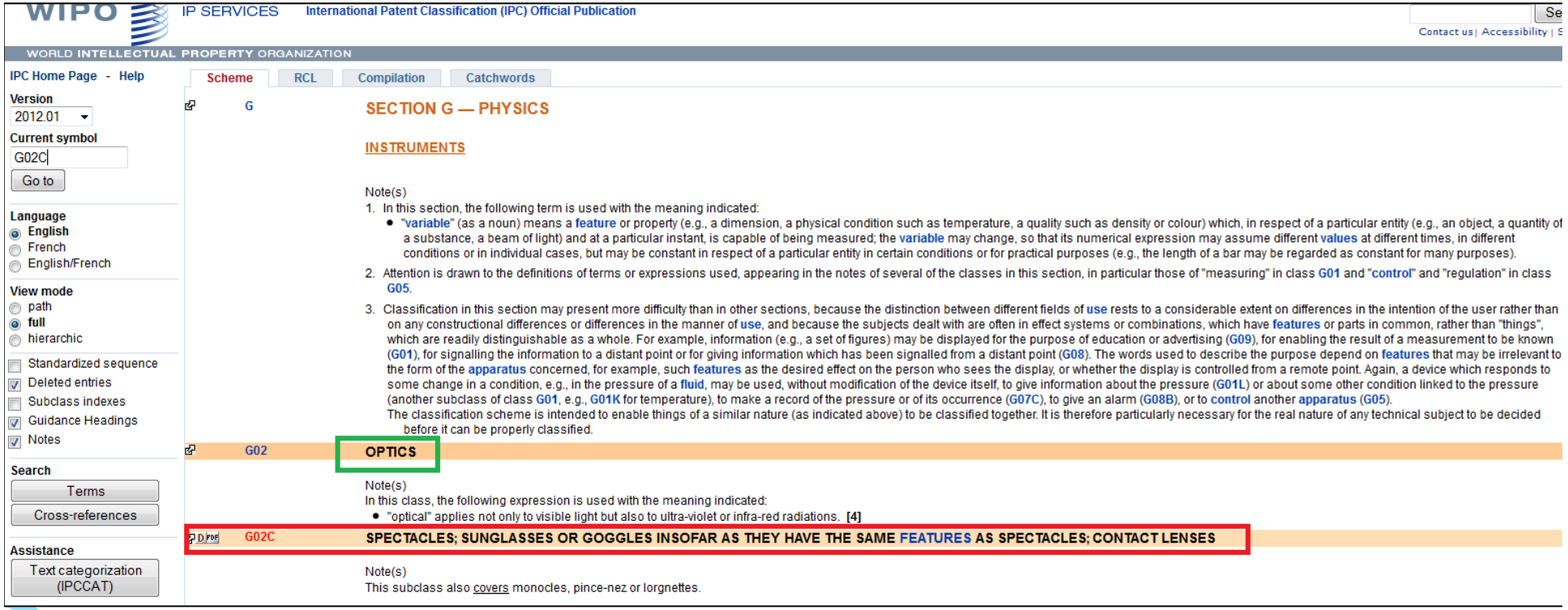
Document: Codes within Hierarchy



Document: Relation between Term and Code



Document: Categorisation of Terms



WIPO IP SERVICES International Patent Classification (IPC) Official Publication

WORLD INTELLECTUAL PROPERTY ORGANIZATION

IPC Home Page - Help

Version 2012.01

Current symbol G02C

Go to

Language: English / French / English/French

View mode: path / full / hierarchic

Standardized sequence / Deleted entries / Subclass indexes / Guidance Headings / Notes

Search: Terms / Cross-references

Assistance: Text categorization (IPCCAT)

Scheme | RCL | Compilation | Catchwords

G — SECTION G — PHYSICS

INSTRUMENTS

Note(s)

1. In this section, the following term is used with the meaning indicated:
   - "variable" (as a noun) means a feature or property (e.g., a dimension, a physical condition such as temperature, a quality such as density or colour) which, in respect of a particular entity (e.g., an object, a quantity of a substance, a beam of light) and at a particular instant, is capable of being measured; the variable may change, so that its numerical expression may assume different values at different times, in different conditions or in individual cases, but may be constant in respect of a particular entity in certain conditions or for practical purposes (e.g., the length of a bar may be regarded as constant for many purposes).
2. Attention is drawn to the definitions of terms or expressions used, appearing in the notes of several of the classes in this section, in particular those of "measuring" in class G01 and "control" and "regulation" in class G05.
3. Classification in this section may present more difficulty than in other sections, because the distinction between different fields of use rests to a considerable extent on differences in the intention of the user rather than on any constructional differences or differences in the manner of use, and because the subjects dealt with are often in effect systems or combinations, which have features or parts in common, rather than "things", which are readily distinguishable as a whole. For example, information (e.g., a set of figures) may be displayed for the purpose of education or advertising (G09), for enabling the result of a measurement to be known (G01), for signalling the information to a distant point or for giving information which has been signalled from a distant point (G08). The words used to describe the purpose depend on features that may be irrelevant to the form of the apparatus concerned, for example, such features as the desired effect on the person who sees the display, or whether the display is controlled from a remote point. Again, a device which responds to some change in a condition, e.g., in the pressure of a fluid, may be used, without modification of the device itself, to give information about the pressure (G01L) or about some other condition linked to the pressure (another subclass of class G01, e.g., G01K for temperature), to make a record of the pressure or of its occurrence (G07C), to give an alarm (G08B), or to control another apparatus (G05).
   The classification scheme is intended to enable things of a similar nature (as indicated above) to be classified together. It is therefore particularly necessary for the real nature of any technical subject to be decided before it can be properly classified.

G02 **OPTICS**

Note(s)
In this class, the following expression is used with the meaning indicated:
- "optical" applies not only to visible light but also to ultra-violet or infra-red radiations. [4]

G02C **SPECTACLES; SUNGLASSES OR GOGGLES INSOFAR AS THEY HAVE THE SAME FEATURES AS SPECTACLES; CONTACT LENSES**

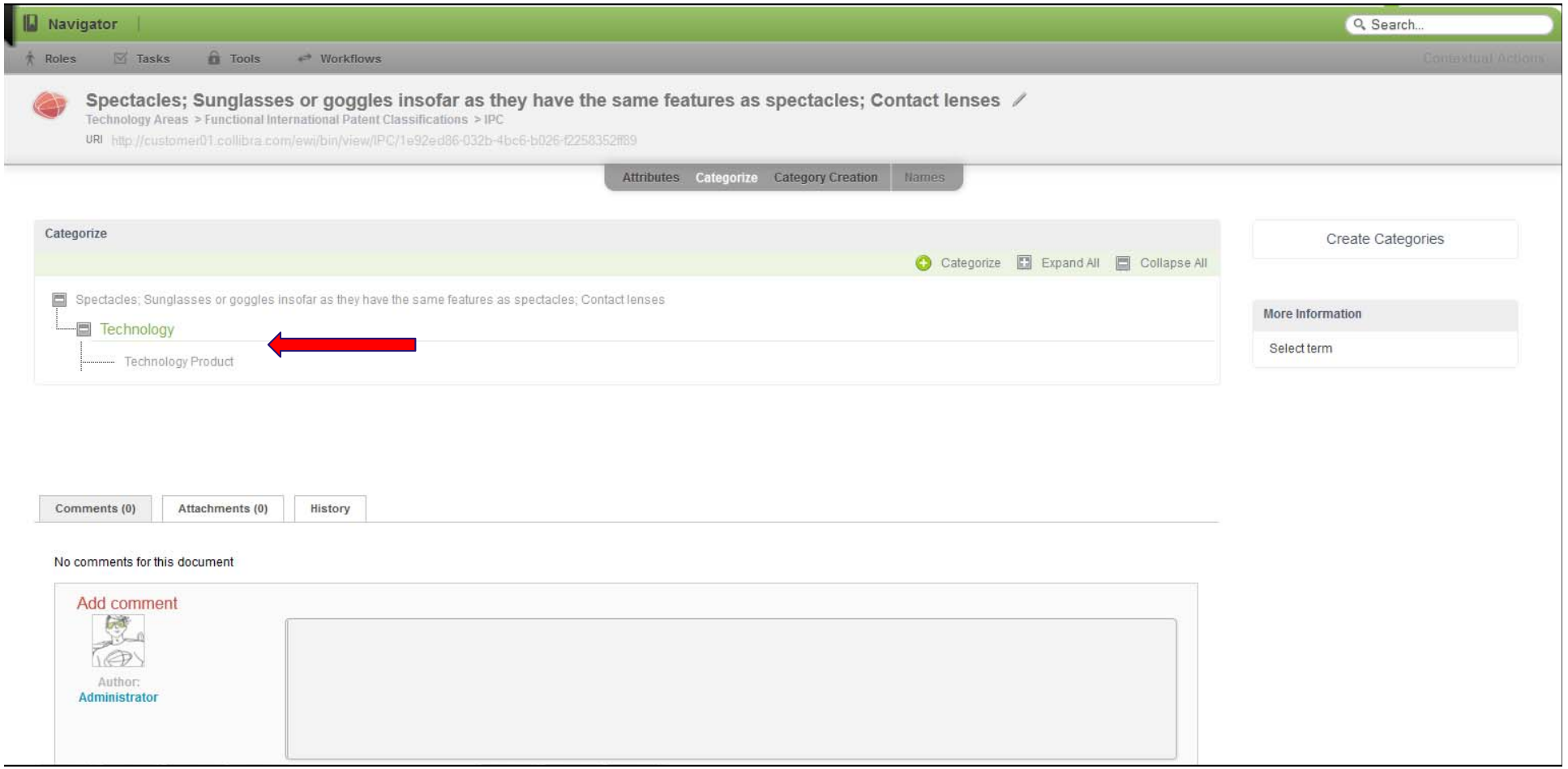
Note(s)
This subclass also covers monocles, pince-nez or lorgnettes.

< Analysis of the Technology classifications:

- Different concepts are being used in one and the same list
- Need for categorisation of terms – separation of concepts!

  e.g. “Basic Materials”: basic materials sector – products? – technology?

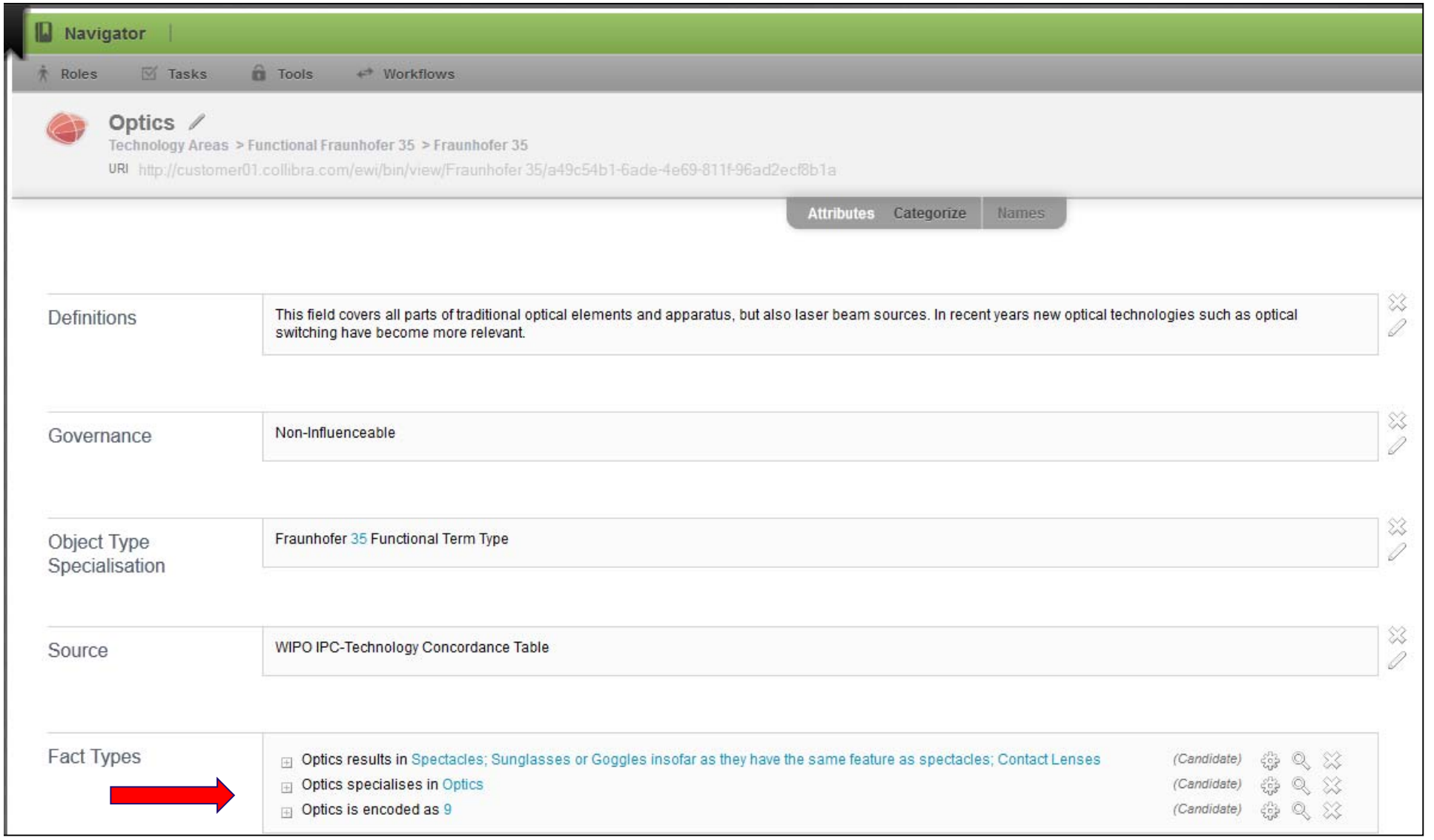
Document: Categorisation of Terms



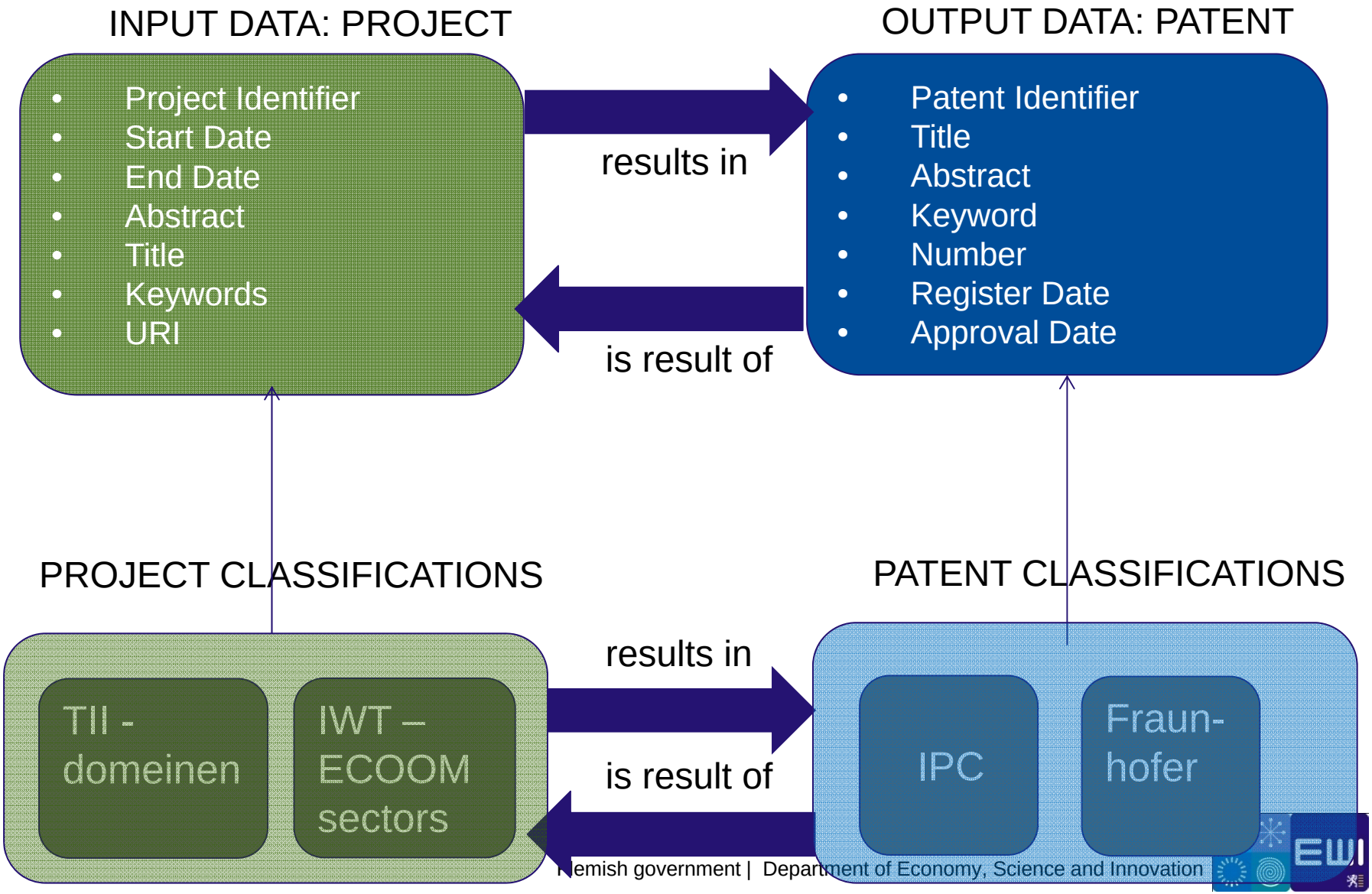
Document: Explicit rules of Mapping

1. Explicit Rules of Mapping Based on Explicit
   - Definitions of the Terms
   - Categorisation of the Terms
2. Different types of mappings
   - Concepts of the same category: 1:1 of 1:n ➔ broader then, narrower then, specialisation of, maps on, specialised in
     - Optics (*Technology Product*) in Fraunhofer specialises in Optics (*Technology Product*) in IPC
   - Concepts of a different category: define the meaning of the relation
     - Optical instruments e.g. telescopes (*Technology Product*) are the result of Optics (*Technology Domain*)

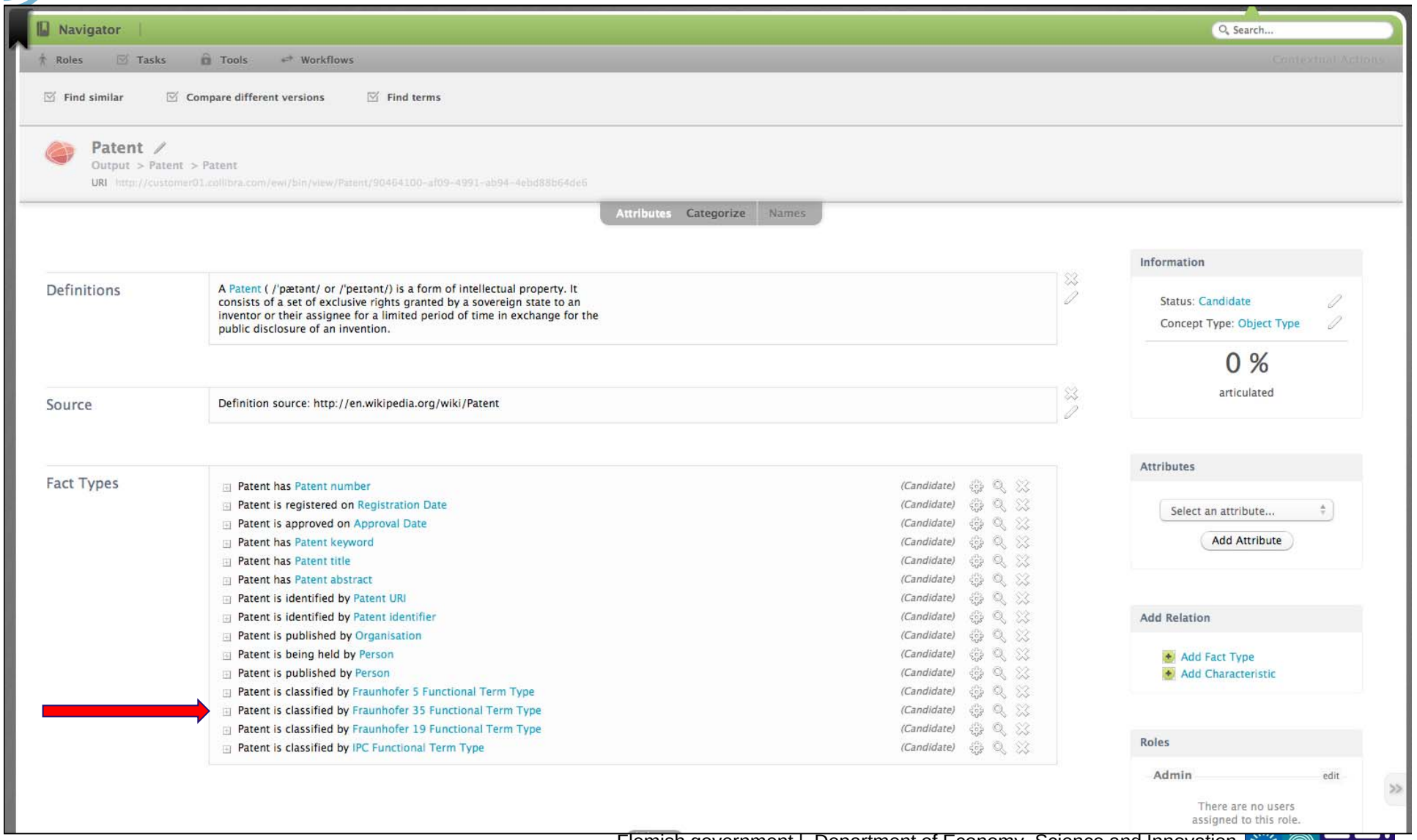
Document: Explicit rules of Mapping



Document: The metamodel of the classification tables



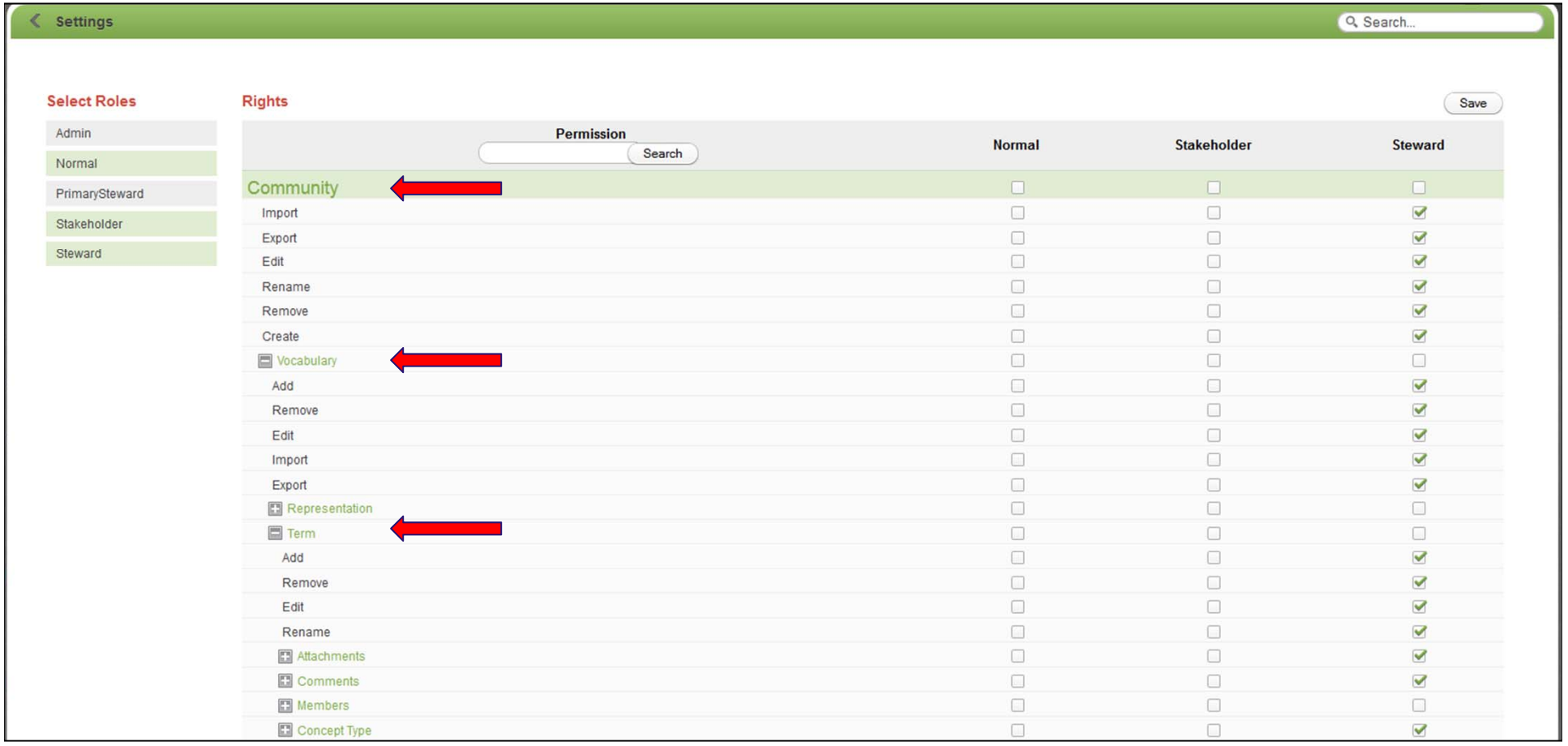
Document: The metamodel of Technology Classifications



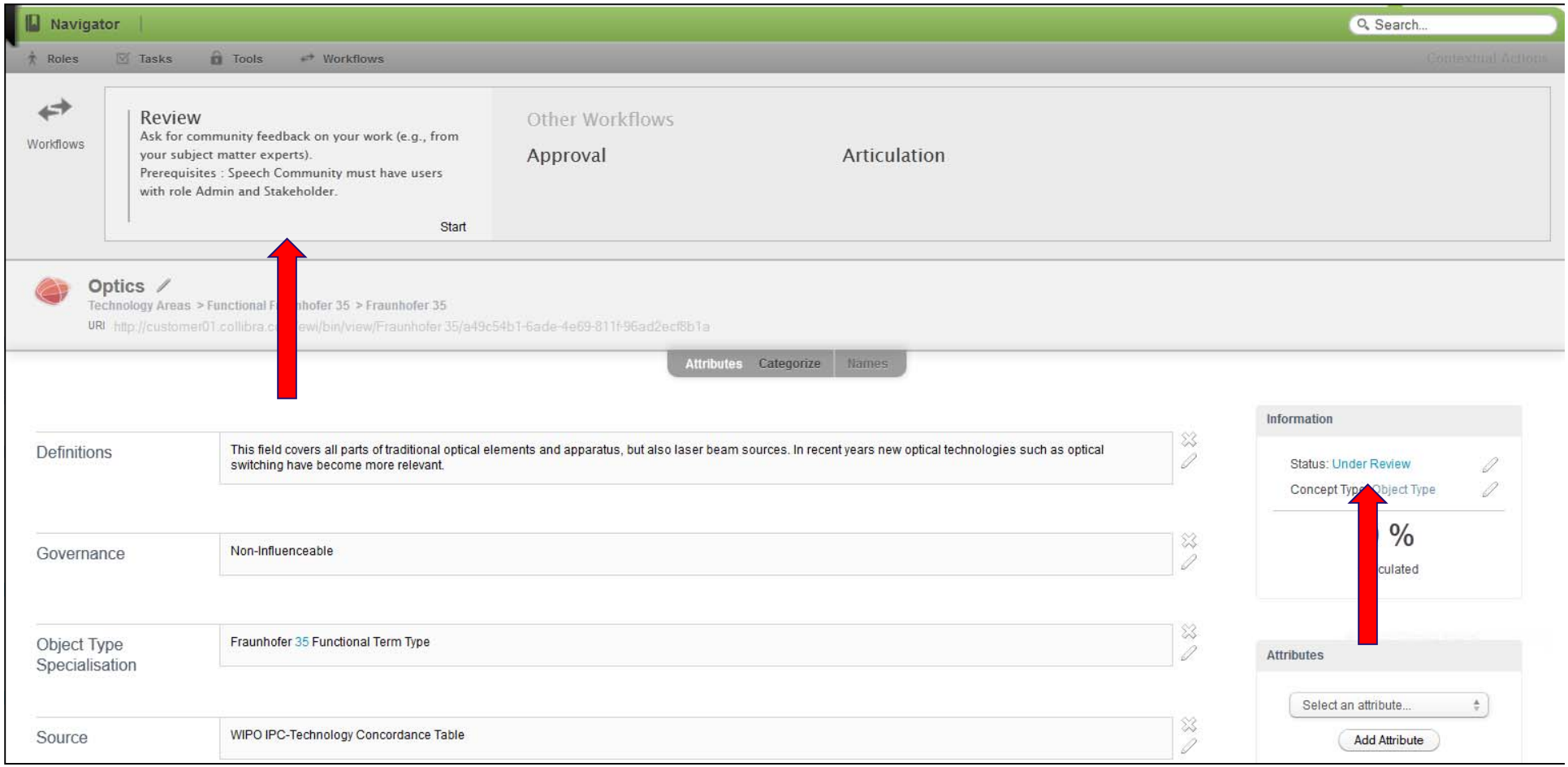
The Five Steps of “Document”

1. Define all Terms – Respect the Code Hierarchy
2. Make the Relation between the Codes and Terms Explicit
3. Separate Different Concepts: Categorize the Terms
    1. If non-influenceable ➔ deduce categorisation
    2. If influenceable ➔ choose the right term!
4. Make the Rules of Mapping Explicit
    1. Based on Definitions (Term Vocabulary)
    2. Based on Categories
5. Take into Account the Metamodel of Classifications

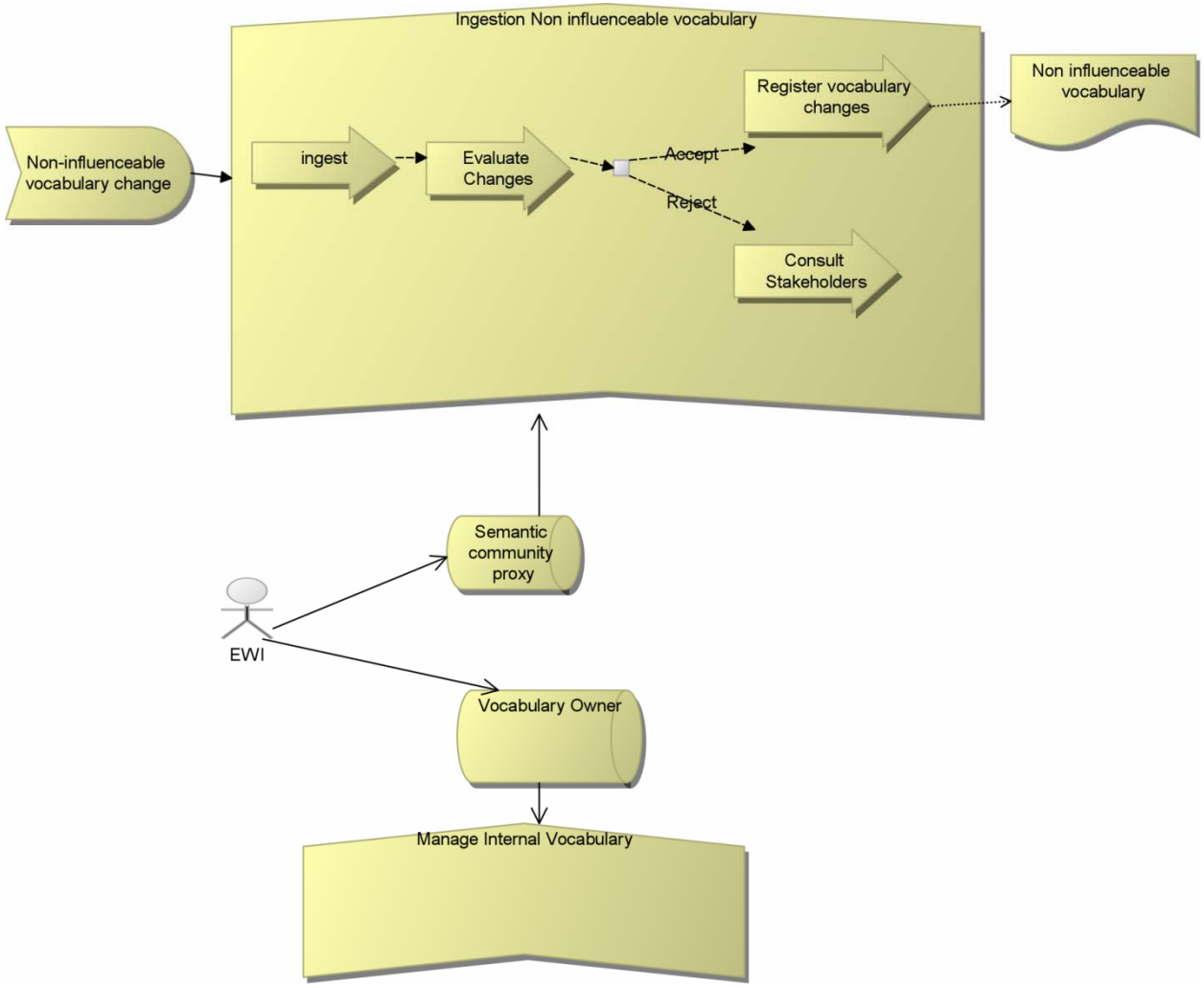
Roles and Rights: Collaborative platform



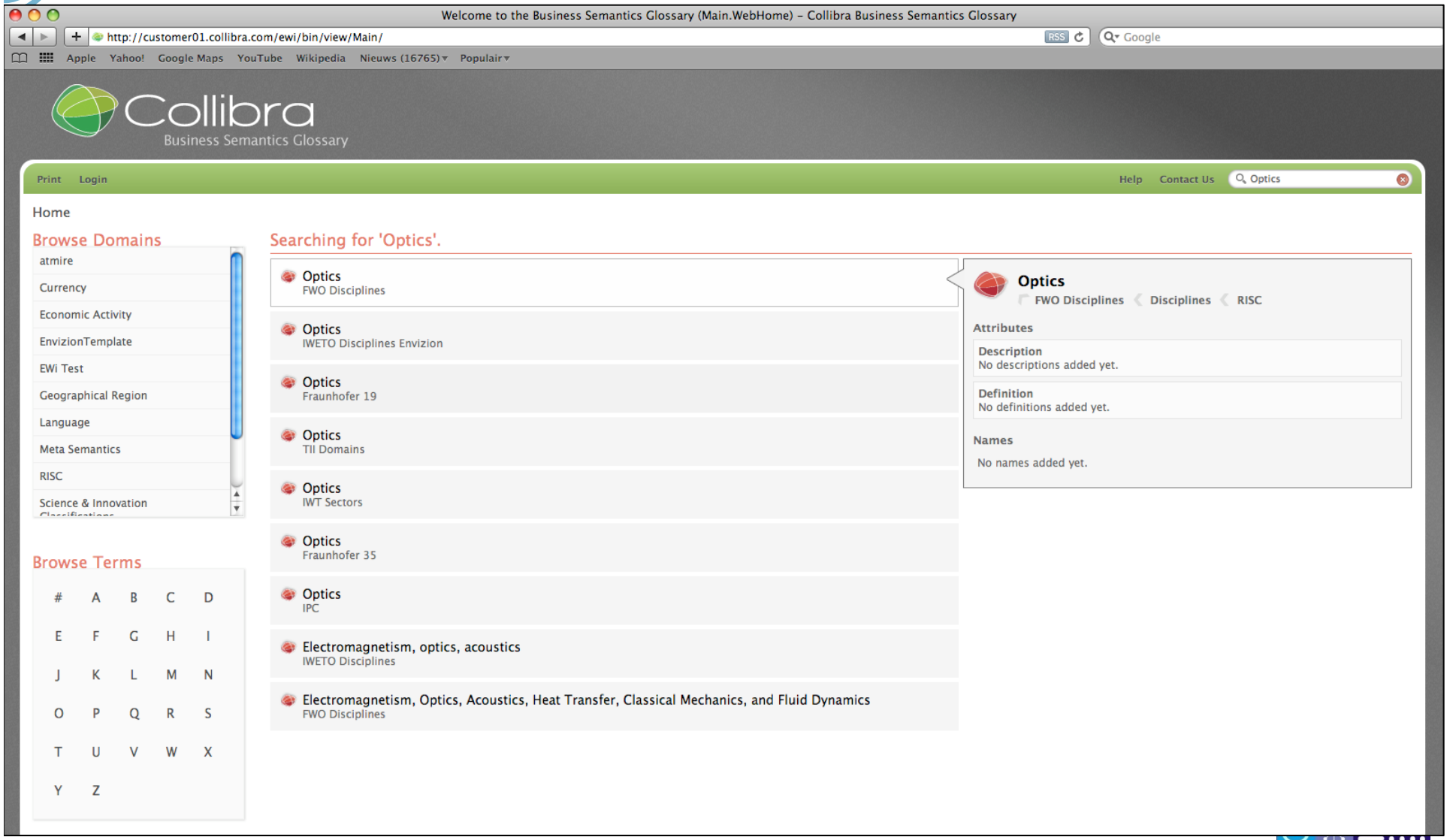
Workflows: Collaborative Platform



Workflow Example: Collaborative Platform

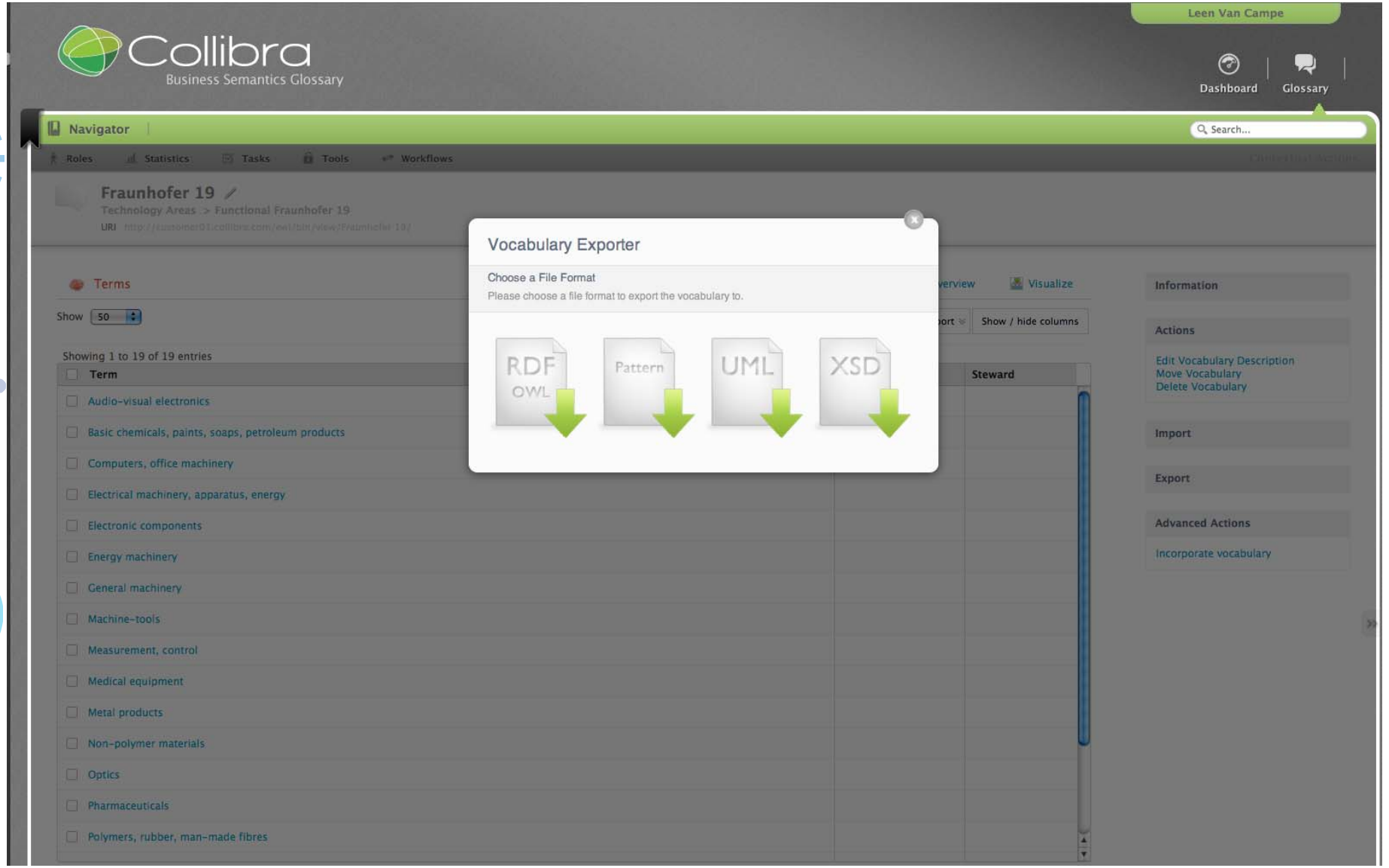


Who can consult the Business Semantics Glossary? Online Platform



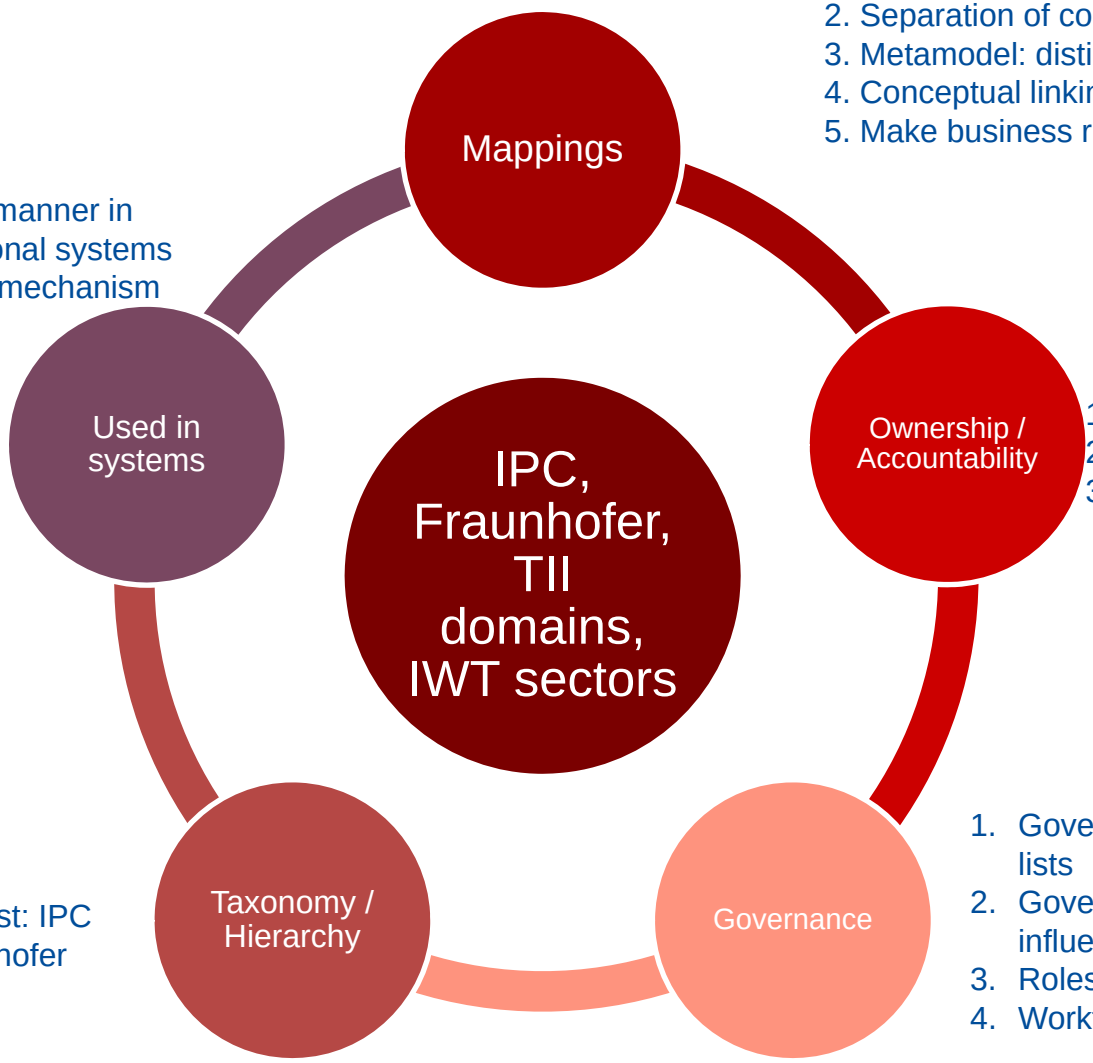
4. Operationalise

Integration with systems / export possibilities



Challenges met?

1. Principles of Mappings
    1. Business discussion about the meaning of Concepts
    2. Separation of concepts: categorisation
    3. Metamodel: distinction between IN and OUT
    4. Conceptual linking (not between terms)
    5. Make business rules explicit



1. Used in read-only manner in analytical and operational systems
2. Publish / subscribe mechanism

1. Community structure SBVR
2. Roles/rights made explicit
3. Workflows

1. Governance Influenceable lists
2. Governance of non-influenceable lists
3. Roles and Rights
4. Workflow support

1. Hierarchy within 1 list: IPC
2. Derived lists: Fraunhofer

Conclusions

EWI's Interest to Collaborate on Classifications:

1. **Organisation**
   - SBVR Community structure for Technology
     - International Patent Classifications
     - Fraunhofer
     - TII Domains
     - IWT sectors
   - SBVR Community structure for Science Domains
     - FWO Disciplines
     - IWETO Disciplines
     - IWETO Science Domains
   - SBVR Community structure for Funding Sources
     - IWETO Funding Sources
   - *Governance council: stakeholders, domain experts, stewards*
2. **Document (Q1 en Q2 in 2013)**
   - Definitions of Fraunhofer, definitions of IPC
   - Funding Sources : versioning + categorisation
   - Science Domains: *Frascati categorisation*
   - Information model / metamodel of classification tables (Patent, Project, Researcher, Funding)
   - *Definitions of FWO Disciplines*?
   - *Definitions of IWT sectors?*
3. **People and Processes**
   - *Roles and rights*
   - *Workflows*
4. **Operationalise**
   - Glossary anywhere
   - Import and export: Q2 in 2013
   - API integration with systems

Thanks

Leen Van Campe
Departement EWI
leen.vancampe@ewi.vlaanderen.be