

## Nanotechnologies – Vocabulary

Platform for Ontology Engineering and Evolution for Nanotechnology

# **AUTOMATIC CONSTRUCTION OF NANOTECHNOLOGY ONTOLOGY STANDARDS**

**ICSU–CODATA WORKSHOP ON THE DESCRIPTION  
OF NANOMATERIALS**, Ministère de l'Enseignement  
Supérieur et de la Recherche, Paris, France, 2012-02-23/24

**NORDIC QUANTUM COMPUTING GROUP AS © 2012**

## Outline

- ❑ Introduction and motivation
- ❑ Development of example nano carbon nomenclature and integration of data in an ontology for nanomaterials
  - Nomenclature for NanoCarbon
  - Integration of physical/chemical properties and data in an OWL 2 ontology for NanoCarbon
  - Ontology evaluation and further development
- ❑ Ontology evolution process
  - Ontology learning
  - Ontology merging
  - Ontology updating
- ❑ Project structure
- ❑ Conclusions

## Ontology Standards for Nanotechnology

- ❑ Ontologies are the central elements in any open e-infrastructure for information exchange across disciplines, organizational and geographical borders
- ❑ In the nanotechnology research field there is still no ontology standard in place, in spite of a strong need to coordinate and relate the diverse research conducted in this interdisciplinary field
- ❑ Ontologies are used to realise improved data integration and interoperability and can potentially bring big savings for both the industry and the public sectors
- ❑ Realising this potential requires efficient ontology construction, merging and evolution technologies

## Ontology Standards for Nanotechnology, cont.

- ❑ The integration of people and applications into more efficient collaborative processes is central to the work on ontology engineering and evolution
- ❑ To make this happen, the applications and the information sources must all understand the annotations with respect to a common language, the ontology
- ❑ This ontology is an explicit and commonly accepted terminology standard that allows us to describe the content of data and services in a machine-processable and unambiguous way

## Current stakeholders (new partners are welcome)

- ❑ Nanoscience, Nanomaterials and Nanotechnology
  - Norwegian Academy of Science and Letters, Standards Norway (ISO), Norwegian National Technical Committee (IEC), Norwegian technical universities, research institutes and industry. Cooperation (finance and technology) with Ministry of Trade and Industry and Ministry of Defense.
  - Broad outreach to international network in research, industry and finance.
- ❑ Computer Science and Information Technology
  - Oxford University, Manchester University, Karlsruhe Institute of Technology, Accenture (global)
- ❑ Computers and Law
  - Clifford Chance LLP, Norwegian Research Center for Computers and Law (NRCCL, University of Oslo), Simonsen Law, Arntzen de Besche

# Project background

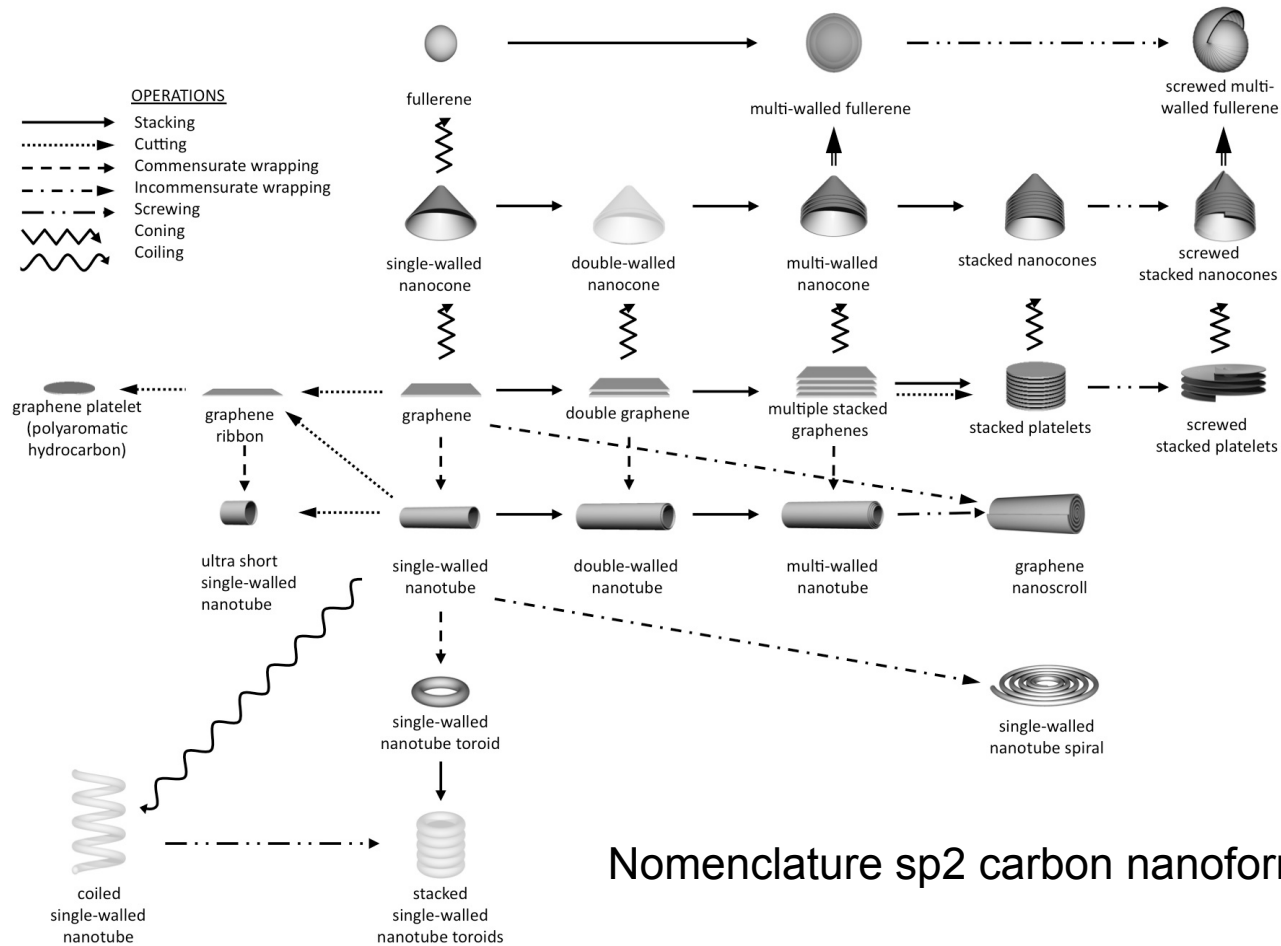
- ❑ Used the following documents
  - ISO/IEC 80004 Nanotechnology Vocabulary Series (specified in OWL2)
  - Nomenclature sp<sup>2</sup> carbon nanoforms (I. Suarez-Martinez, N. Grobert, Chr. P. Ewels)
  - Informal standards, like the PAS, and scientific articles from the nanotechnology domain
- ❑ Ontology lifecycle model used
  - Both manual modeling tasks and automatic tool support
  - The ontology learning workbench (OWB) maps existing standards onto OWL2 (Ontology Web Language) structures and suggests new ontology structures from mining scientific domain documentation
  - Results from the OWB can be viewed, evaluated and modified with any ontology editor (e.g., Protégé)
  - Whereas the OWB speeds up the analysis of domain documents, the decisions on what to include in the final ontology still rests with the ontology modelers (nano domain experts)

## Difficulties with the use of ontologies

- ❑ Ontologies are being developed independently for different subdisciplines of nanotechnology
- ❑ New ontological elements (terminologies and definitions) continuously emerge from researchers and standardization groups
- ❑ Some old elements become irrelevant or find more popular corresponding synonyms
- ❑ The dynamic nature of this process requires continuous update of existing ontologies
- ❑ Collaboration between distinct subdisciplines requires safe techniques for merging of independently developed ontologies

Managing ontologies manually is not a simple task

# Nomenclature for NanoCarbon



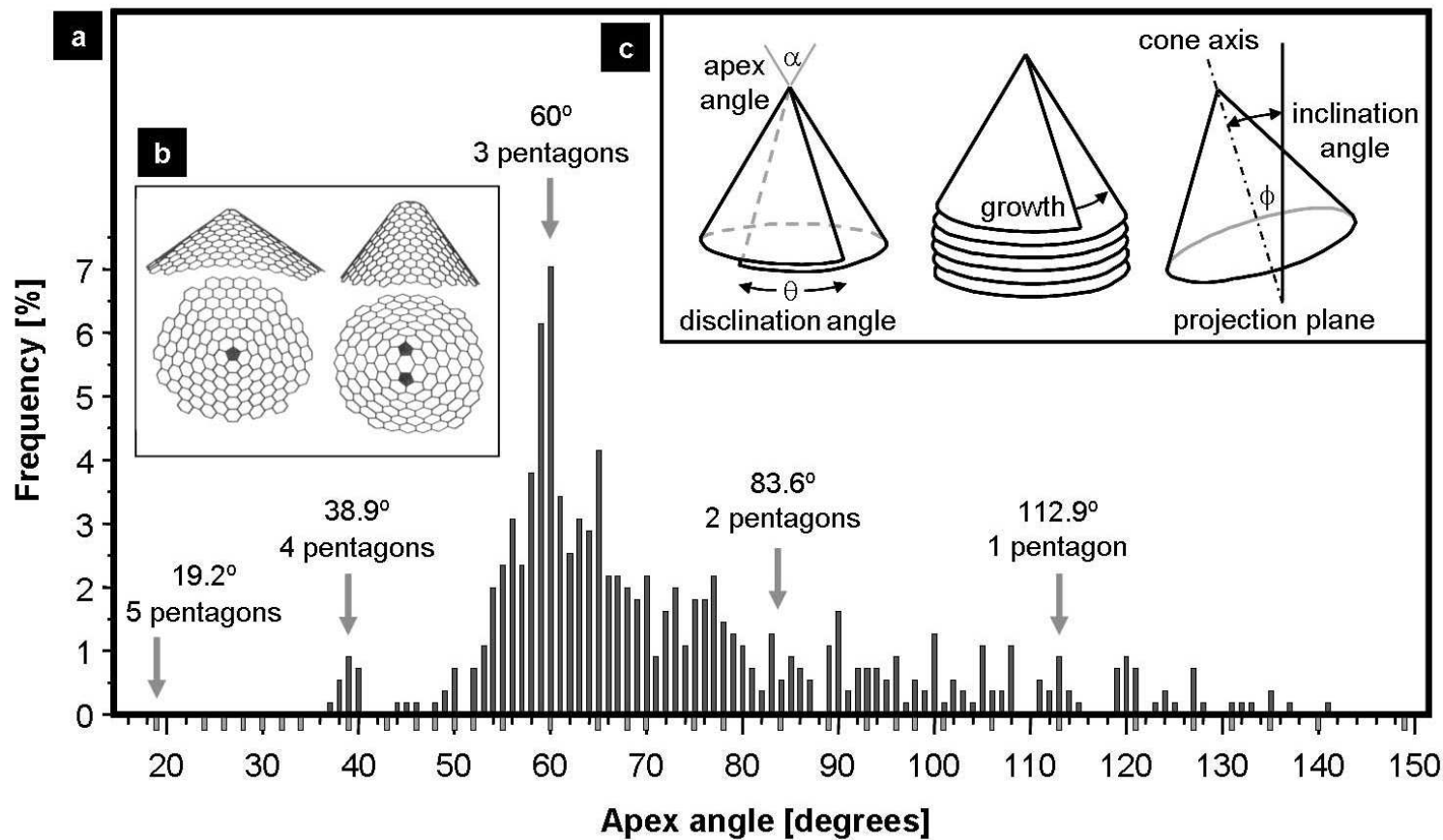
Nomenclature sp<sup>2</sup> carbon nanoforms



## Nomenclature for NanoCarbon, cont.

- ❑ The nomenclature is developed in accordance with the guidelines in; *Nomenclature for  $sp^2$  carbon nanoforms*
- ❑ Exception: The term *helical nanocones* is used instead of *screwed stacked nanocones*, as we think the latter is incorrect; screwed cones cannot be exploded into single graphene cones – they consist of a single layer wrapped around itself. (See next picture)

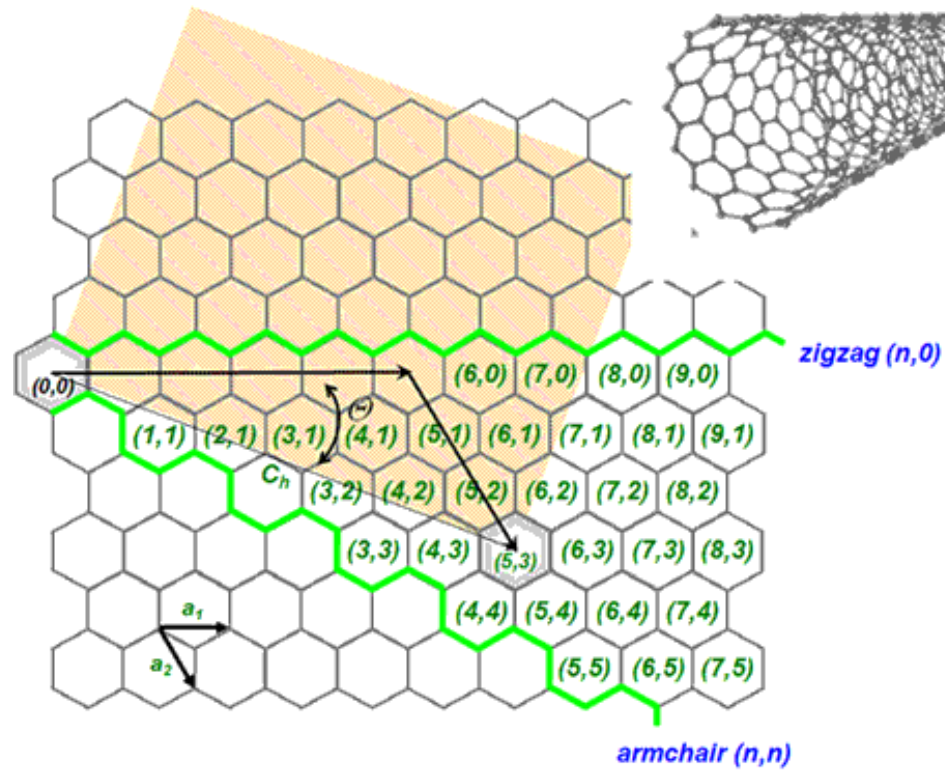
# Nomenclature for NanoCarbon, cont.



## Single Walled Carbon Nano Tubes

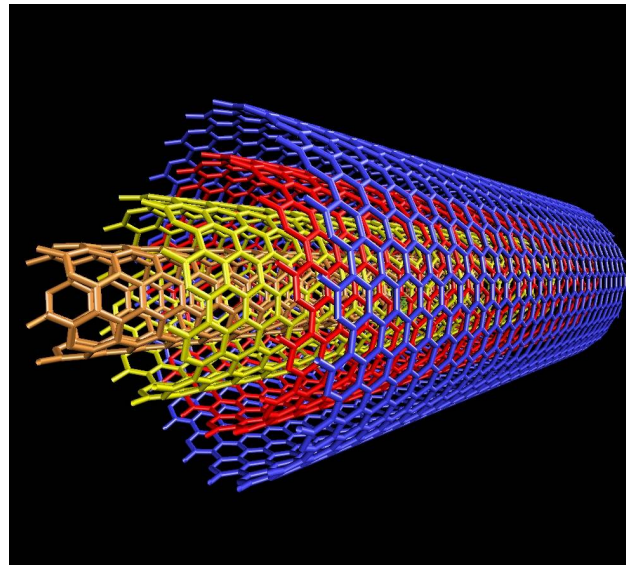
- ❑ SWNT[R13/L100/M6/N5] = Single Walled Carbon Nano Tube with Radius  $R = 13$  nm, Length  $L = 100$  nm and chiral vector (see next picture)  $(M,N) = (6,5)$
- ❑ This specifies completely an open-ended tube without functionalized surface or defects. Suggestions to alternative syntax are welcome; e.g. SWNT\_R{13}L{100}...etc

# Chiral Vector



## Multi-Walled Carbon Nanotubes

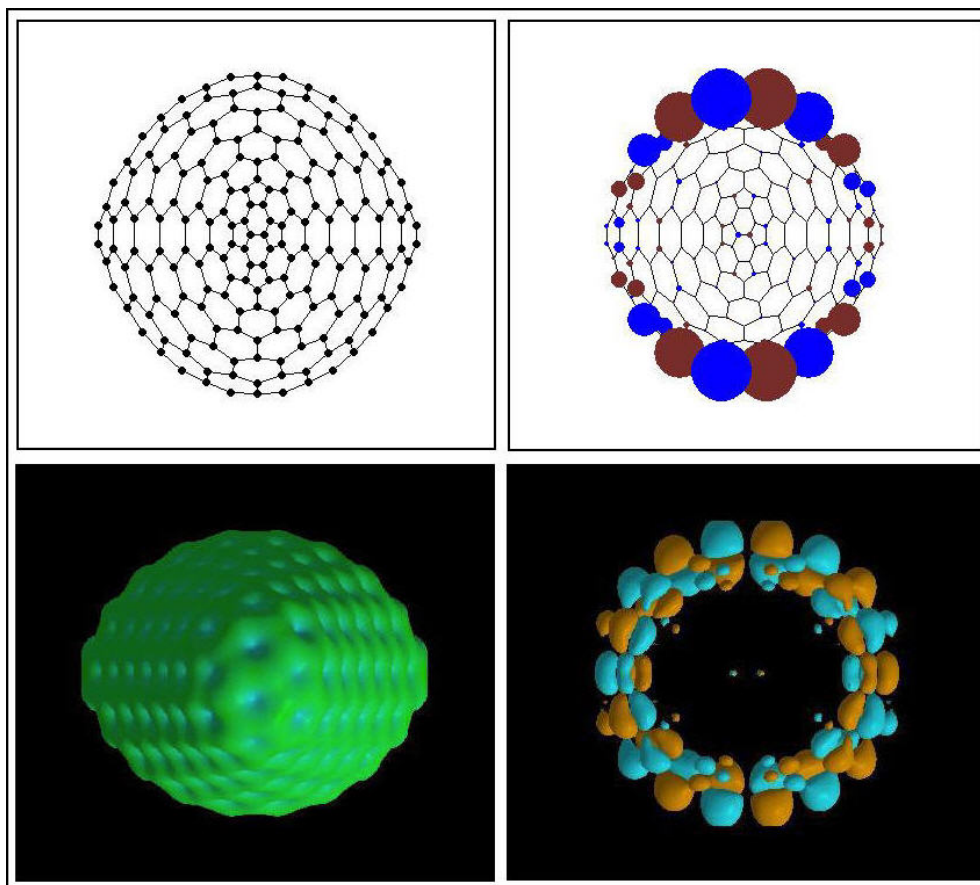
- MWCNT[R34/L460/M7/N1/W5] = Carbon Nano Tube with # Walls  $W = 5$  where the outer wall has Radius  $R = 34$  nm, Length  $L = 460$  nm and chiral vector  $(M,N)=(7,1)$ .



## Graphene Cones

- ❑ Single Walled Carbon Nano Cones are called *Graphene Cones* in the ontology, as they can be transformed from graphene by removing 1 to 5 sectors of disclination angle  $60^\circ$  and rejoining the resulting dangling bonds.
- ❑ GC[D240/R250] = Graphene Cone transformed from circular graphene of Radius  $R=250$  nm and total disclination angle  $D = 240^\circ = 4 \times 60^\circ$ .

# Integration of physical/chemical properties and data in an ontology for NanoCarbon



# Model Ontology for Nano Carbon

ScreenHunter\_05 Feb. 20 23:45.png 02/20 23:45 PC2743 henningh ScreenHunter

File Edit Ontologies Reasoner Tools Refactor Tabs View Window Help

Ontology1329742973296 (http://www.semanticweb.org/ontologies/2012/1/ontology1329742973296.owl)

Entities Classes Object Properties Data Properties Individuals OWL Viz

Class Hierarchy: NanoCarbon

- Thing
  - NanoCarbon
    - Carbon\_Nano\_Conos
      - Graphene\_Conos\_(GC[...])'
      - 'Helical\_Conos\_(HC[...])'
      - 'Stacked\_Conos\_(SGC[...])'
    - Carbon\_Nano\_Tubes
      - ...//SWCNT[...]
      - ...@SWCNT[...]
      - MWCNT[...]
      - SWCNT[...]
    - Graphene
    - Graphene\_Platelet
    - Graphene\_Ribbon
    - Stacked\_Graphene
    - Stacked\_Platelets
  - Properties
    - Chemical
      - Inert
      - Reactive
    - Electronic
      - Insulator
      - Metallic
      - Semiconductor
    - Mechanical
      - Stiff
      - Tensile

OWL Viz: NanoCarbon

Asserted model Inferred model

```

graph TD
    Thing -- is-a --> Properties
    Thing -- is-a --> NanoCarbon
    Properties -- is-a --> Mechanical
    Properties -- is-a --> Electronic
    Properties -- is-a --> Chemical
    Mechanical -- is-a --> Stiff
    Mechanical -- is-a --> Tensile
    Electronic -- is-a --> Insulator
    Electronic -- is-a --> Metallic
    Electronic -- is-a --> Semiconductor
    Chemical -- is-a --> Inert
    Chemical -- is-a --> Reactive
    NanoCarbon -- is-a --> Stacked_Graphene
    NanoCarbon -- is-a --> Stacked_Platelets
    NanoCarbon -- is-a --> Carbon_Nano_Tubes
    NanoCarbon -- is-a --> Graphene
    NanoCarbon -- is-a --> Graphene_Platelet
    NanoCarbon -- is-a --> Carbon_Nano_Conos
    NanoCarbon -- is-a --> Graphene_Ribbon
    Carbon_Nano_Tubes -- is-a --> SWCNT1[...//SWCNT[...]]
    Carbon_Nano_Tubes -- is-a --> SWCNT2[...@SWCNT[...]]
    Carbon_Nano_Tubes -- is-a --> SWCNT3[SWCNT[...]]
    Carbon_Nano_Tubes -- is-a --> MWCNT[MWCNT[...]]
    Carbon_Nano_Conos -- is-a --> GC['Graphene_Conos_(GC[...])']
    Carbon_Nano_Conos -- is-a --> HC['Helical_Conos_(HC[...])']
    Carbon_Nano_Conos -- is-a --> SGC['Stacked_Conos_(SGC[...])']
    
```

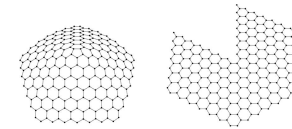


# Representation of Nano Carbon entities. Below: SWNCT with methyl-functionalized surface

The screenshot shows the Protégé ontology editor interface. The top menu bar includes File, Edit, Ontologies, Reasoner, Tools, Refactor, Tabs, View, Window, and Help. The address bar shows the URL: <http://www.semanticweb.org/ontologies/2012/1/1/Ontology1329742973296.owl>. The interface is divided into several panes:

- Class hierarchy (left):** Shows a tree structure starting with 'Thing' and 'NanoCarbon'. Under 'NanoCarbon', there are sub-classes like 'Carbon\_Nano\_Cones', 'Carbon\_Nano\_Tubes', 'Graphene', 'Graphene\_Platelet', 'Graphene\_Ribbon', 'Stacked\_Graphene', and 'Stacked\_Platelets'. There are also 'Properties' listed: 'Chemical' (Inert, Reactive), 'Electronic' (Insulator, Metallic, Semiconductor), and 'Mechanical' (Stiff, Tensile).
- Members list (center):** Shows the inferred members list for the selected class. The selected class is 'CH3/SWCNT[R38/L160/M7/N1]'. The members list is currently empty.
- Annotations (top right):** Shows annotations for the selected class. A comment is visible: "SWNCT[38/L160/M7/N1] (see SWNT[...]) with methyl-functionalized surface."
- Property assertions (bottom right):** Shows property assertions for the selected class. The data property assertions are:
  - hasLength "160 nm"
  - hasYoungsModulus "0.54 Tpa"
  - hasRadius "38 nm"
  - hasSolvent "Chloroform, Acetone"
  - hasChiralVector "(7,1)"
  - hasBandGap "0.23 keV"

# Graphene Cone



File Edit Ontologies Reasoner Tools Refactor Tabs View Window Help ScreenHunter\_06 Feb. 21 06:58.png 02/21 06:58 PC2743 henningh ScreenHunter

Ontology1329742973296 (http://www.semanticweb.org/ontologies/2012/1/1/ontology1329742973296.owl)

Entities Classes Object Properties Data Properties Individuals OWL Viz

Class hierarchy: Polar

- Thing
  - NanoCarbon
    - Carbon\_Nano\_Cones
      - Graphene\_Cones\_(GC[...])'
      - Helical\_Cones\_(HC[...])'
      - Stacked\_Cones\_(SGC[...])'
    - Carbon\_Nano\_Tubes
      - Graphene
      - Graphene\_Platelet
      - Graphene\_Ribbon
      - Stacked\_Graphene
      - Stacked\_Platelets
  - Properties
    - Chemical
      - Polar
      - Inert
      - Reactive
    - Electronic
      - Insulator
      - Metallic
      - Semiconductor
    - Mechanical
      - Stiff
      - Tensile

Members list (Inferred)

Members list (Inferred): GC[D240/R250]

Members list: GC[D240/R250]

Annotations Usage

Annotations: GC[D240/R250]

Annotations +

comment

"Graphene Cone with total disclination angle D of 240=4x60 deg and radius R=250 nm"

Description: GC[D240/R250]

Types +

- 'Graphene\_Cones\_(GC[...])'
- Polar
- Semiconductor
- Stiff

Same individuals +

Different individuals +

Property assertions: GC[D240/R250]

Object property assertions +

Data property assertions +

- hasHOMO-LUMO-Gap "0.73 keV"
- hasDipoleMoment "316 Debye"
- hasYoungsModulus "0.86 Tpa"

Negative object property assertions +

Negative data property assertions +

## Heterogeneities in ontologies for nanotechnology

- ❑ Heterogeneity can be:
  - assembly of higher order structures not having same size with disparate geometric arrangements,
  - that represents dissimilar level of perfection.
  - both static and dynamic.
- ❑ Transferability to other systems / challenges:
  - Hierarchical structure is probably possible, but the shape of this is likely to be material dependent.
  - Should also be compatible with other relevant nomenclature
    - Eg. crystallographic in the case of crystals can be formed
    - Chemical (eg, self-organizing forms)
    - Biological (eg, PDB database, gene sequences, etc.) when such building blocks included in nanostructures.

## Heterogeneities in ontologies for nanotechnology

- ❑ Heterogeneities and structural defects can hardly be addressed through the ontology alone.
- ❑ Complete specification in these cases can however be achieved by a link under Datatype Properties in the OWL2 interface to a molecular file, e.g. in Protein Data Bank (PDB) format.
- ❑ The molecular file defines the complex completely, and can be displayed graphically. If possible, a standard naming convention is still desirable for these cases.

## Ontology evaluation and further development (input from ICSU CODATA workshop attendees)

- ❑ Is the nomenclature convenient? Complete specification through nomenclature alone seems infeasible due to e.g. the plethora of topological variations at the open ends of tubes and cones.
- ❑ Building model ontologies is a good method to test nomenclatures. For example, integration of numerical data requires a more specific nomenclature than qualitative properties.
- ❑ Further development requires expert input on dependencies of physical/chemical properties on structural details.

# Project: Ontology Engineering and Evolution for Nanotechnology

- A **collaborative environment** to support the use and development of nanotechnology ontologies
  - **To learn** new ontological elements
    - highlighting **emerging terms** among researcher and standardization groups
    - identifying **relations** between them
  - **To combine** ontologies
    - helping to **merge existing ontologies** from different subdisciplines of nanotechnology
    - locating **overlapping** and **conflicting** regions
  - **To update** ontologies
    - finding ontology elements that become **out-of-date** or are **irrelevant** for researchers and standardization groups
    - identifying elements that should be **renamed**
    - identifying new elements that should be **included**

## Ontology Learning

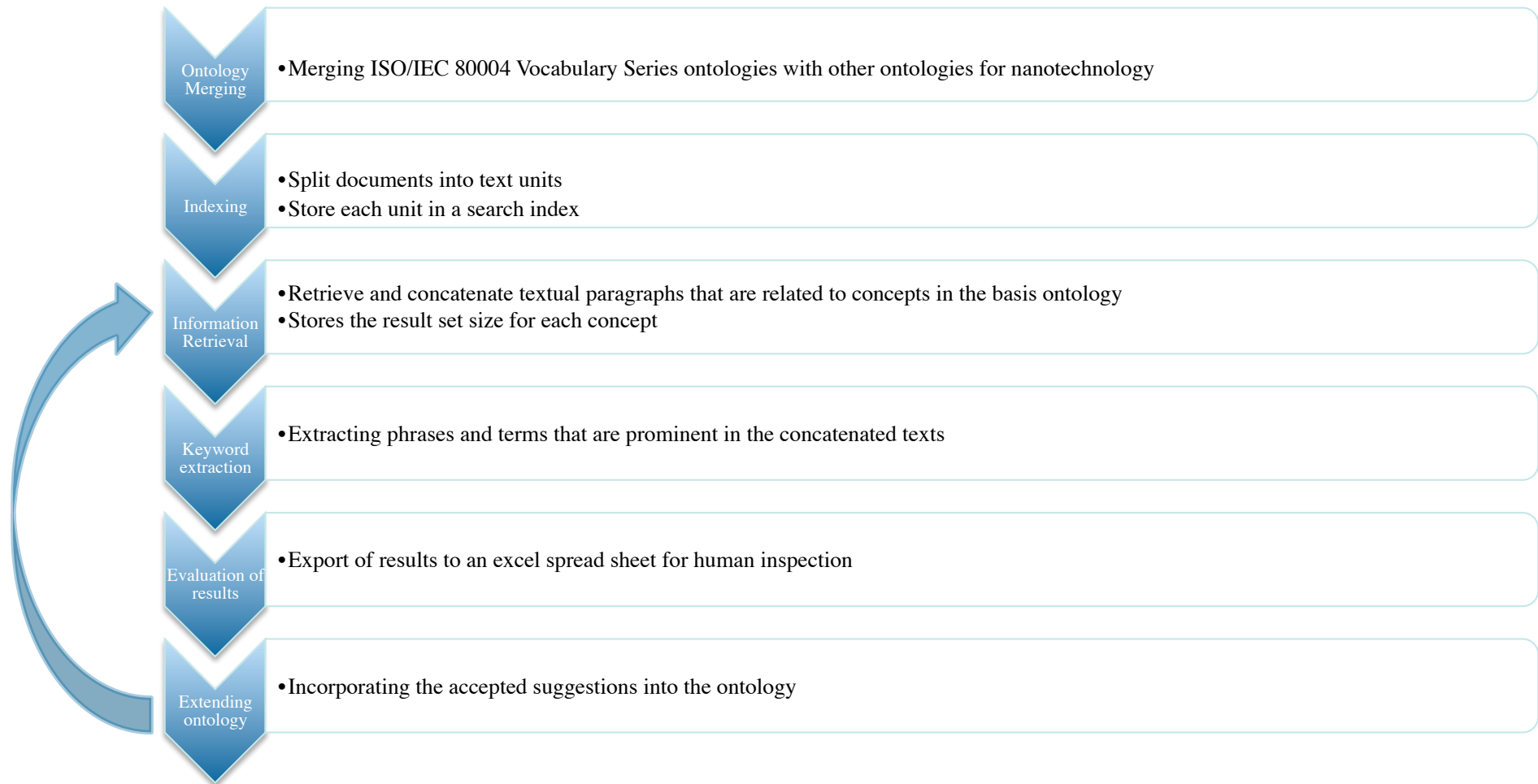
- ❑ Goal: (semi-)automatically identify and extract relevant concepts and relations from a given corpus to form an ontology
- ❑ Documents are parsed; indexed and subject to statistical analyses for concept extraction;
- ❑ This produces statistical reports including
  - suggestion scores (how often has a candidate term been suggested), and
  - the suggested neighborhood in an existing ontology

## Ontology Merging

- ❑ Some ontologies are developed in totally different scientific and industrial contexts
- ❑ However, the inter-disciplinary nature of the field requires collaboration between academics and industrial partners
- ❑ Safe merging of ontologies is important and tools are needed to
  - detect possible inconsistencies
  - assist in the repair of these inconsistencies



# Merging of Ontologies

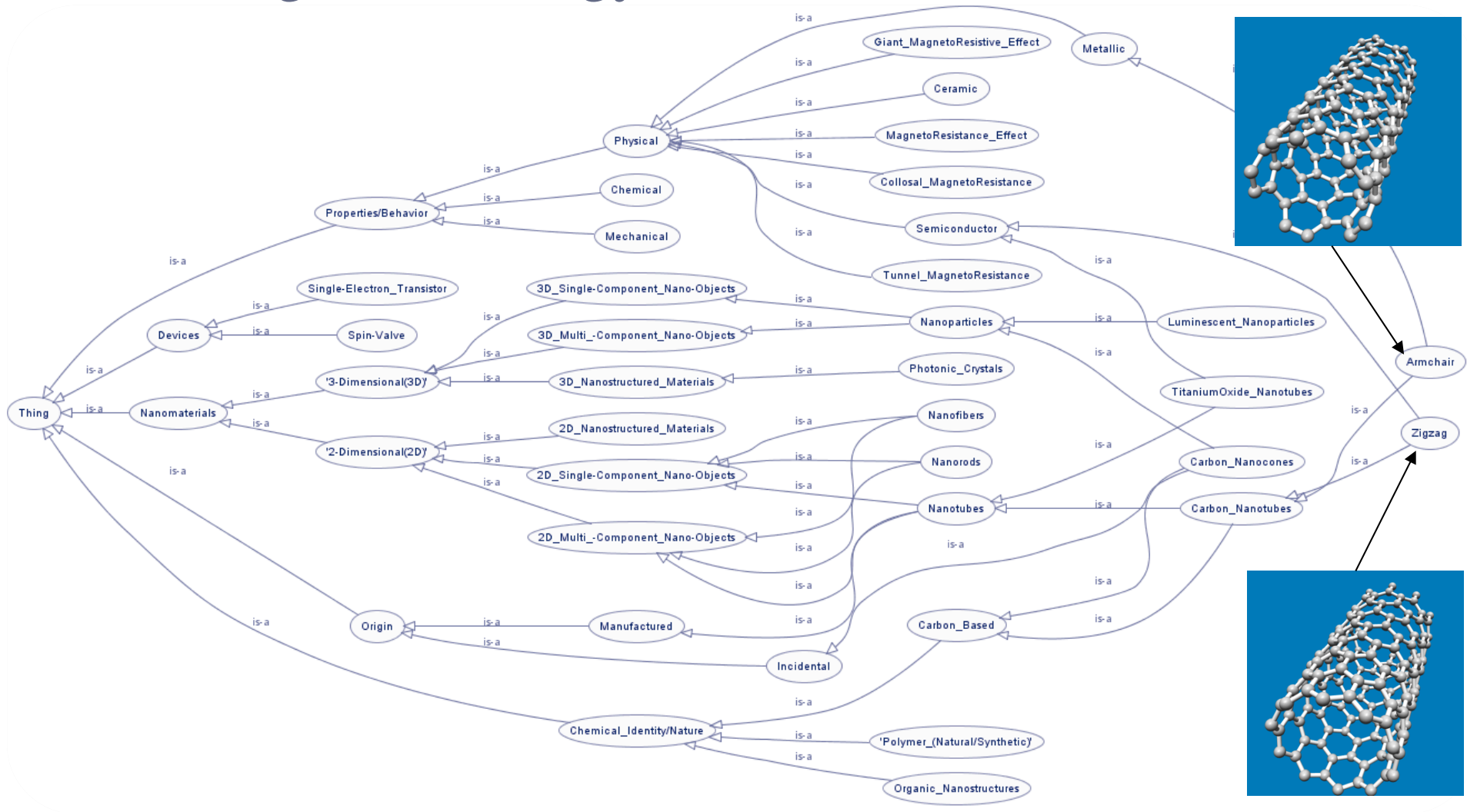




## Ontology Updating

- ❑ Ontologies evolve naturally
- ❑ Most industries experience that ontologies are very dynamic and grow over time
- ❑ Manual update is complex; costly and time consuming
- ❑ Document analysis tools can suggest the removal; renaming; or inclusion of new terms in an ontology

# Extending the ontology (1)



# Extending the ontology (2)

Ontology1329852355625.owl (http://www.semanticweb.org/ontologies/2012/1/Ontology1329852355625.owl) - [/Users/APM/Library/Mail Downloads/merged.owl]

Active Ontology | Entities | Classes | Object Properties | Data Properties | Individuals | OWLViz | DL Query

Class hierarchy: NanoCarbon

- Thing
  - NanoCarbon
    - Carbon\_Nano\_Conos
      - 'Graphene\_Conos\_(GC[...])'
      - 'Helical\_Conos\_(HC[...])'
      - 'Stacked\_Conos\_(SGC[...])'
    - Carbon\_Nano\_Tubes
      - ...//SWCNT[...]
      - ...@SWCNT[...]
      - MWCNT[...]
      - SWCNT[...]
    - Graphene
    - Graphene\_Platelet
    - Graphene\_Ribbon
    - Stacked\_Graphene
    - Stacked\_Platelets
  - Properties
  - chemical\_identity/nature
  - nanomaterial
    - engineered\_nanomaterial
    - incidental\_nanomaterial
    - manufactured\_nanomaterial
  - nano-object
  - nanostructured\_material
  - properties/behavior
  - vocabulary
    - 80004-10\_nanophotonic\_components
    - 80004-1\_core\_terms
    - 80004-3\_carbon\_nano-objects
    - 80004-4\_nanostructured\_materials
    - 80004-5\_nano-bio\_interface
    - 80004-6\_nanoscale\_measurement\_and\_instrumentation
    - 80004-7\_healthcare
    - 80004-8\_nanomanufacturing\_processes

OWLViz: NanoCarbon

Asserted model | Inferred model

```

classDiagram
    class Thing
    class NanoCarbon
    class Carbon_Nano_Tubes
    class Stacked_Graphene
    class Graphene_Ribbon
    class Stacked_Platelets
    class Graphene_Platelet
    class Graphene
    class Carbon_Nano_Conos
    class MWCNT["MWCNT[...]"]
    class SWCNT["...//SWCNT[...]"]
    class ATSWCNT["...@SWCNT[...]"]
    class SWCNT2["SWCNT[...]"]
    class Helical["'Helical_Conos_(HC[...])'"]
    class Graphene_GC["'Graphene_Conos_(GC[...])'"]
    class Stacked_SGC["'Stacked_Conos_(SGC[...])'"]

    Thing --|> NanoCarbon
    NanoCarbon --|> Carbon_Nano_Tubes
    NanoCarbon --|> Stacked_Graphene
    NanoCarbon --|> Graphene_Ribbon
    NanoCarbon --|> Stacked_Platelets
    NanoCarbon --|> Graphene_Platelet
    NanoCarbon --|> Graphene
    NanoCarbon --|> Carbon_Nano_Conos
    Carbon_Nano_Tubes --|> MWCNT
    Carbon_Nano_Tubes --|> SWCNT
    Carbon_Nano_Tubes --|> ATSWCNT
    Carbon_Nano_Tubes --|> SWCNT2
    Carbon_Nano_Conos --|> Helical
    Carbon_Nano_Conos --|> Graphene_GC
    Carbon_Nano_Conos --|> Stacked_SGC
    
```

No Reasoner set  Show Inferences

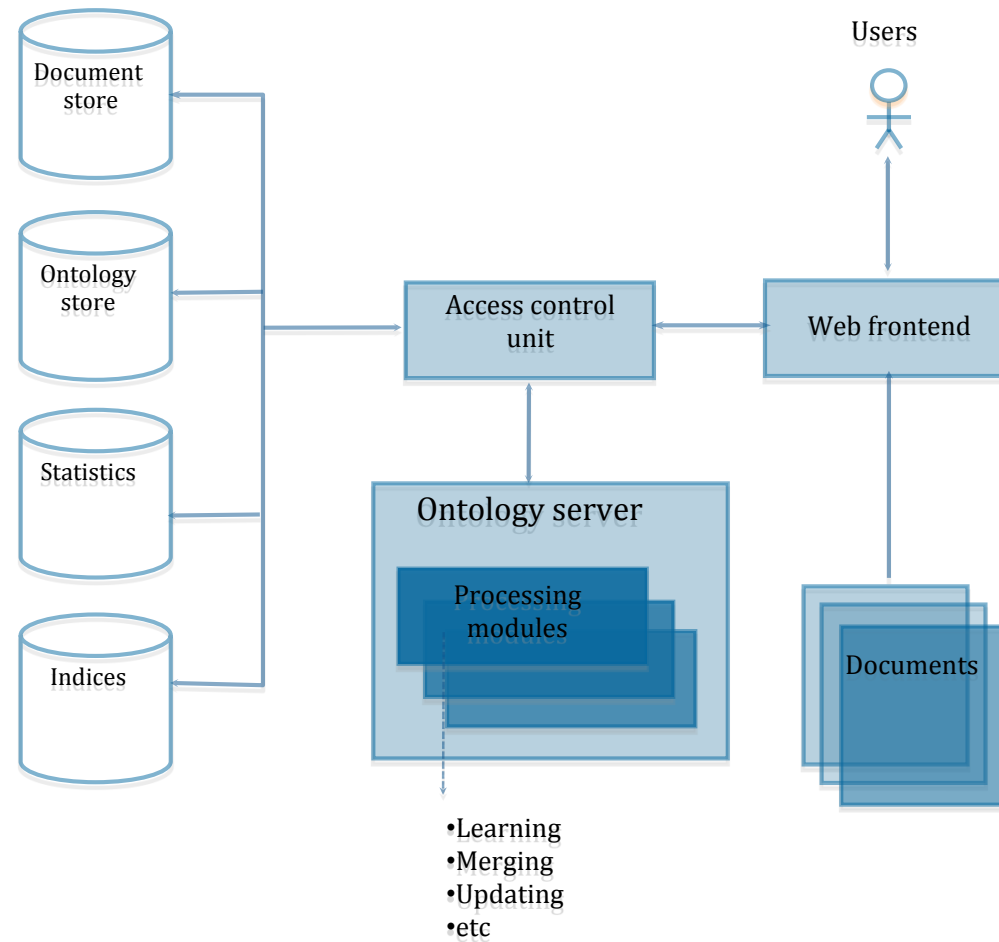
## Integrating these techniques

- ❑ Learning, merging and updating ontologies are relevant for all cases of nanotechnology
- ❑ Industrial collaborators need a unified tool-supported platform integrating all of these aspects of ontology design
- ❑ Research collaborators can contribute with cutting-edge techniques and benefit from tests in real-case scenarios
- ❑ An iterative development process that gradually integrates the techniques can mitigate risks and offer real benefits to all concerned

# Project: Ontology Engineering and Evolution for Nanotechnology

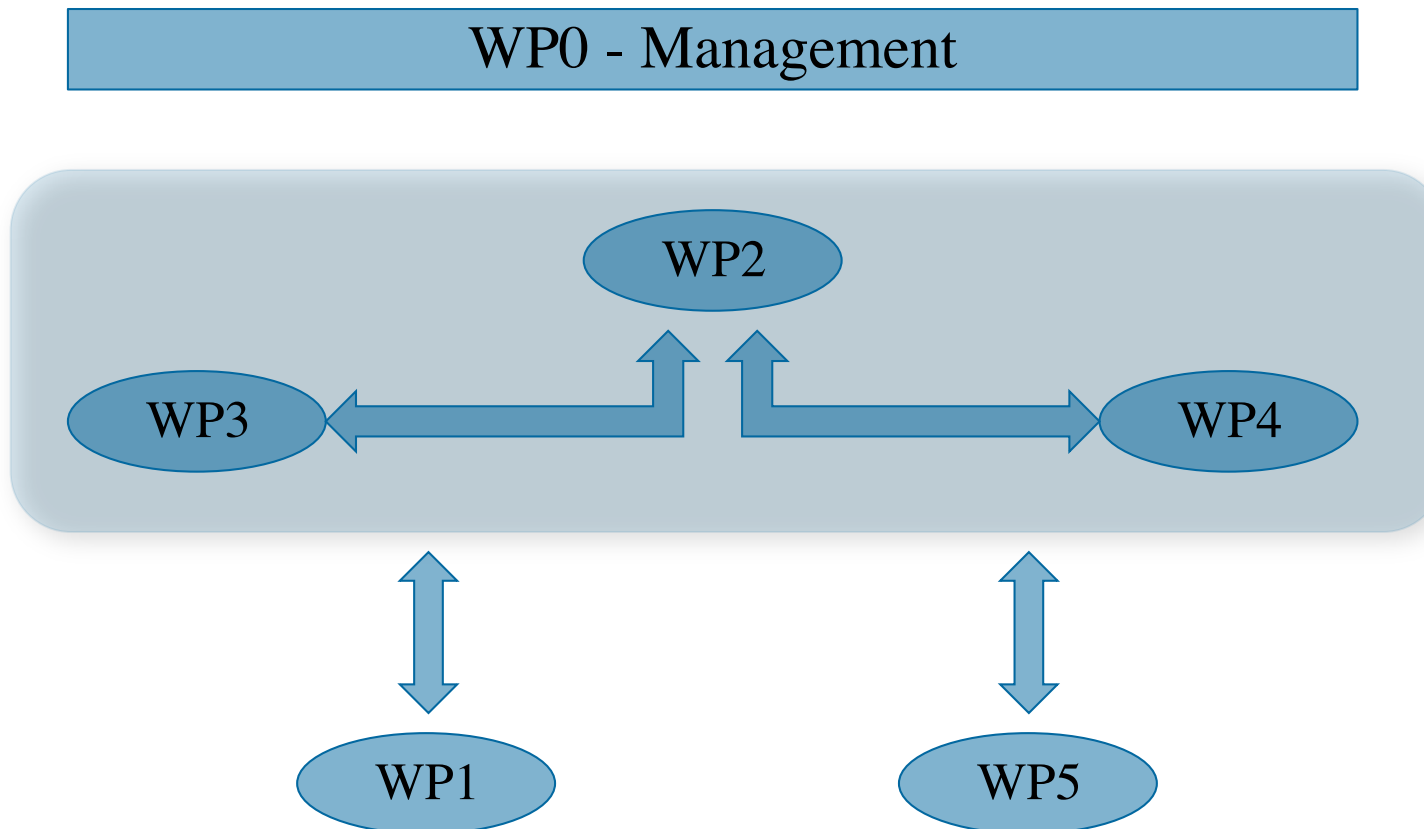
- Goal:
  - To provide a collaborative environment for the development and use of ontologies and the integration of advanced techniques for ontology design
  - To demonstrate ontology learning in the nanotechnology domain
- Approach:
  - Development of a core basis and subsequent incorporation of more sophisticated components
  - Use of existing ontologies in nanotechnology as a case study to support the development
- Ontology basis:
  - ISO/IEC TS 80004 Nanotechnology Vocabulary Series
  - IUPAC nomenclatures (InChI)
  - Crystallographic/chemical/biological nomenclatures
  - Gene Ontology (GO) and other ontologies, e.g. NPO
- Document corpus:
  - Documents on nanotechnology in relevant fields
  - Working documents from ISO/TC 229, IEC/TC 113, CEN/TC 352, OECD WPMN

# Platform architecture





# Project structure



## Project structure

- ❑ WP0 – Project management
  - NQCG
- ❑ WP1 – Ontology Evolution Platform
  - NQCG
- ❑ WP2 – Integrated Ontology Evolution Framework
  - Oxford/Manchester: Prof. Ian Horrocks and Prof. Uli Sattler
- ❑ WP3 – Ontology Learning and Extraction
  - KIT: Prof. Rudi Studer
- ❑ WP4 – Collaborative Ontology Engineering
  - Oxford/Manchester: Prof. Ian Horrocks and Prof. Uli Sattler
- ❑ WP5 – Ontology Engineering for the Nanotechnology Domain
  - NQCG: Dr. Henning Heiberg-Andersen

## Project structure

- ❑ WP1 – Ontology Evolution Platform
  - Specification, design and implementation of the platform
- ❑ WP2 – Integrated Ontology Evolution Framework
  - Representation language and structures
- ❑ WP3 – Ontology Learning and Extraction
  - Data acquisition, linguistic processing, ontology learning, trend-detection
- ❑ WP4 – Collaborative Ontology Engineering
  - Composition and decomposition framework for collaborative engineering
- ❑ WP5 – Ontology Engineering for the Nanotechnology Domain
  - Use case study and evaluation aspects

# Integrated Ontology Evolution Framework

- ❑ Integrated ontology representation language
  - extension of OWL 2 to handle information about
    - provenance
    - trust of ontology components (e.g., axioms and modules)
    - composition of an ontology
  - so as to allow
    - representation of uncertainties
    - a more powerful OWL 2 import and export mechanism

# Ontology Learning and Extraction

- ❑ Tools and techniques for detection of occurrences of ontology components (classes, individual, etc) in natural language text
  - Data acquisition: document crawlers and data infrastructure for textual knowledge
  - Linguistic processing
  - Ontology learning and mapping
  - Ontology-based trend detection
    - identify trends and emerging topics in newly published documents

# Collaborative Ontology Engineering

- ❑ Development of methods that allow independent manipulation of parts of ontologies with subsequent re-integration into the overall ontology
- ❑ Ontology decomposition, version control and reconciliation
  - analysis of the logical modular structure of the ontologies
  - support for version management
  - support for conflict resolution

## Conclusions

- ❑ Ontology learning should be used as an effective tool in
  - Development of standard nanotechnology terminologies and nomenclature systems
  - Maintenance of standard nanotechnology terminologies and nomenclature systems
  - Harmonization of different standard nanotechnology terminologies and nomenclature
  
- ❑ Collaborative ontology engineering platform can effectively assist ICSU
  - Construct and maintain nomenclature systems for nanoscience and nanotechnology
  - Collaborative environment to support the use and development of nano ontologies
  - Internet platform for collaboration and cooperation in nano research and development
  
- ❑ A broad representation of experts should take part in this work
  - ICSU, CODATA and Scientific Unions like IUPAC
  - Materials Research Societies
  - ISO/TC 229 and IEC/TC 113
  - OECD Working Party on Manufactured Nanomaterials and WPN
  - IP5 Foundation Projects (EPO, JPO, KIPO, SIPO, USPTO)
  - World Intellectual Property Organization (WIPO)
  - National academies of science and national standards development organizations
  - Organizations involved in developing classifications for nanotechnology
  - Leading scientific journals and publishers in relevant domains