DODO: A Mechanism Helping to Dynamically Construct Domain Ontologies for Services Integration

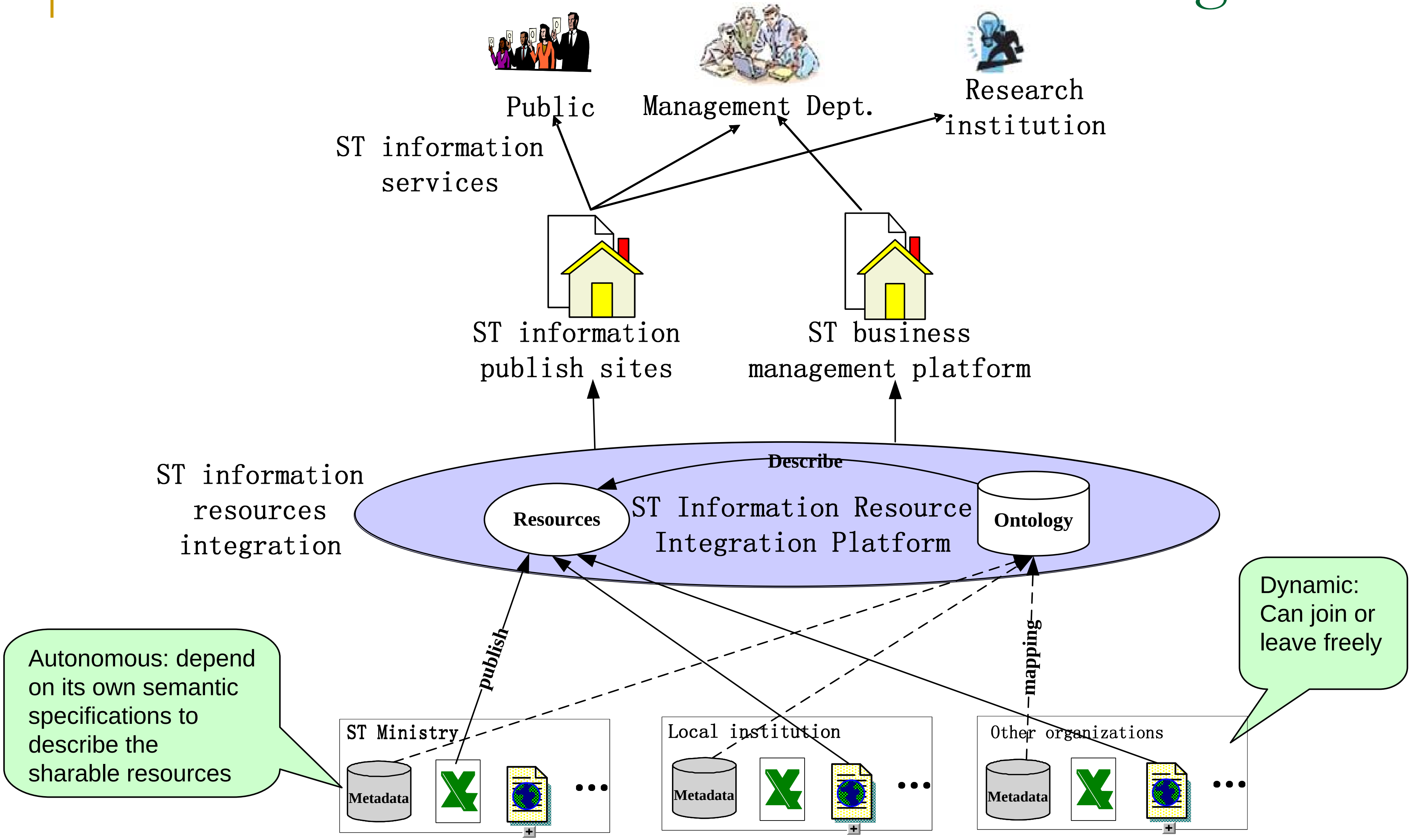
Chen Liu，Wanghu Chen，Yanbo Han
liuchen@software.ict.ac.cn
Research Centre for Grid and Service Computing,
Institute of Computing Technology, Chinese Academy of Sciences

Outline

- Background and Challenges
- Related Works
- Rationale of the DODO Mechanism
- Application of the DODO Mechanism
- Conclusion

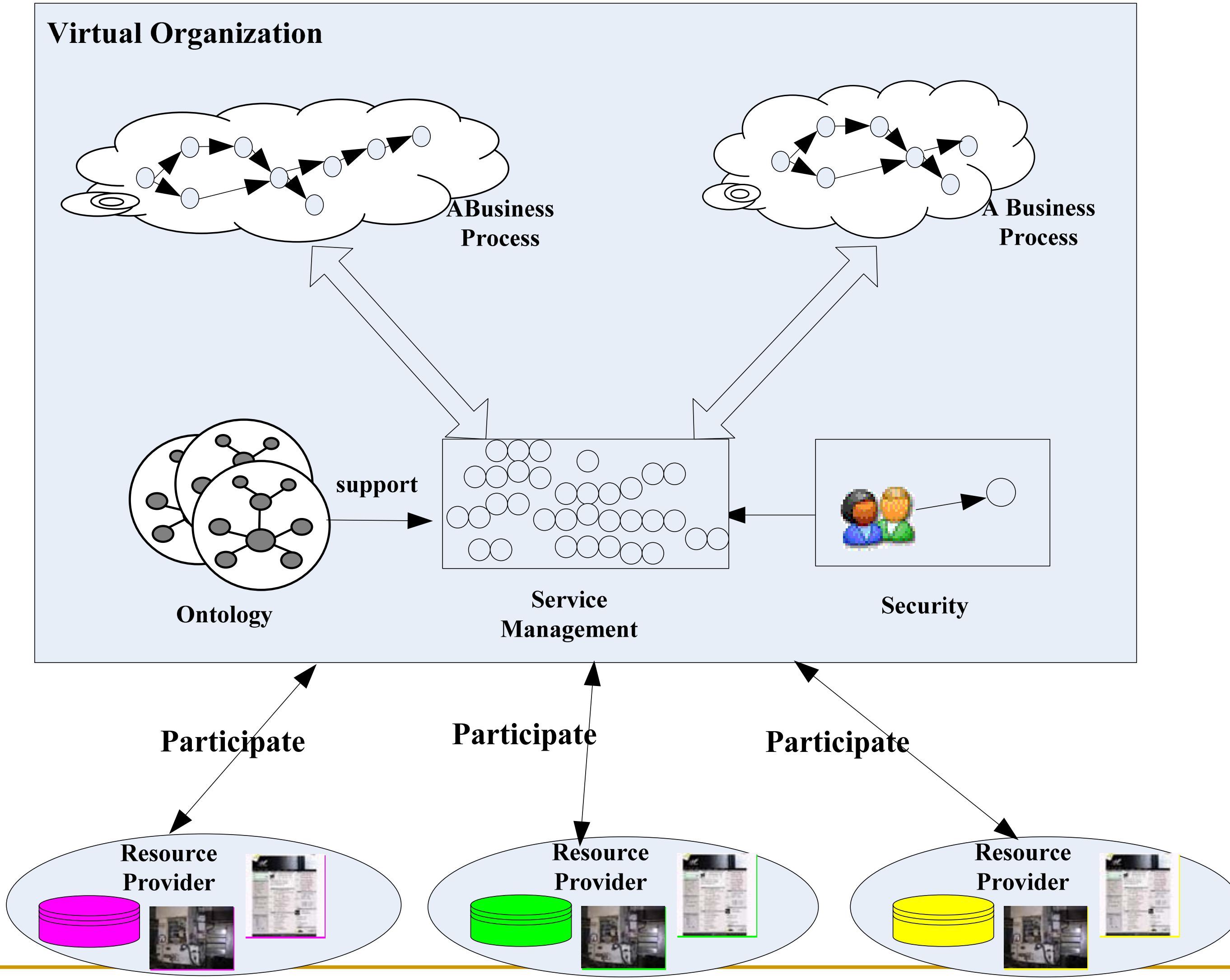
Background and Challenges

A scenario : ST* information resources integration



*ST : science and technology

Our solutions: virtual organization



The characteristics of a VO

- Goal driven: A VO is sponsored according to a business goal. The sponsor is called the owner.
- Dynamic: Each Resource provider can join or leave a VO freely.
- Autonomous: Each member of a VO is autonomous.
- Process-oriented: The services in a VO can be composed into business processes to satisfy users' requirements.

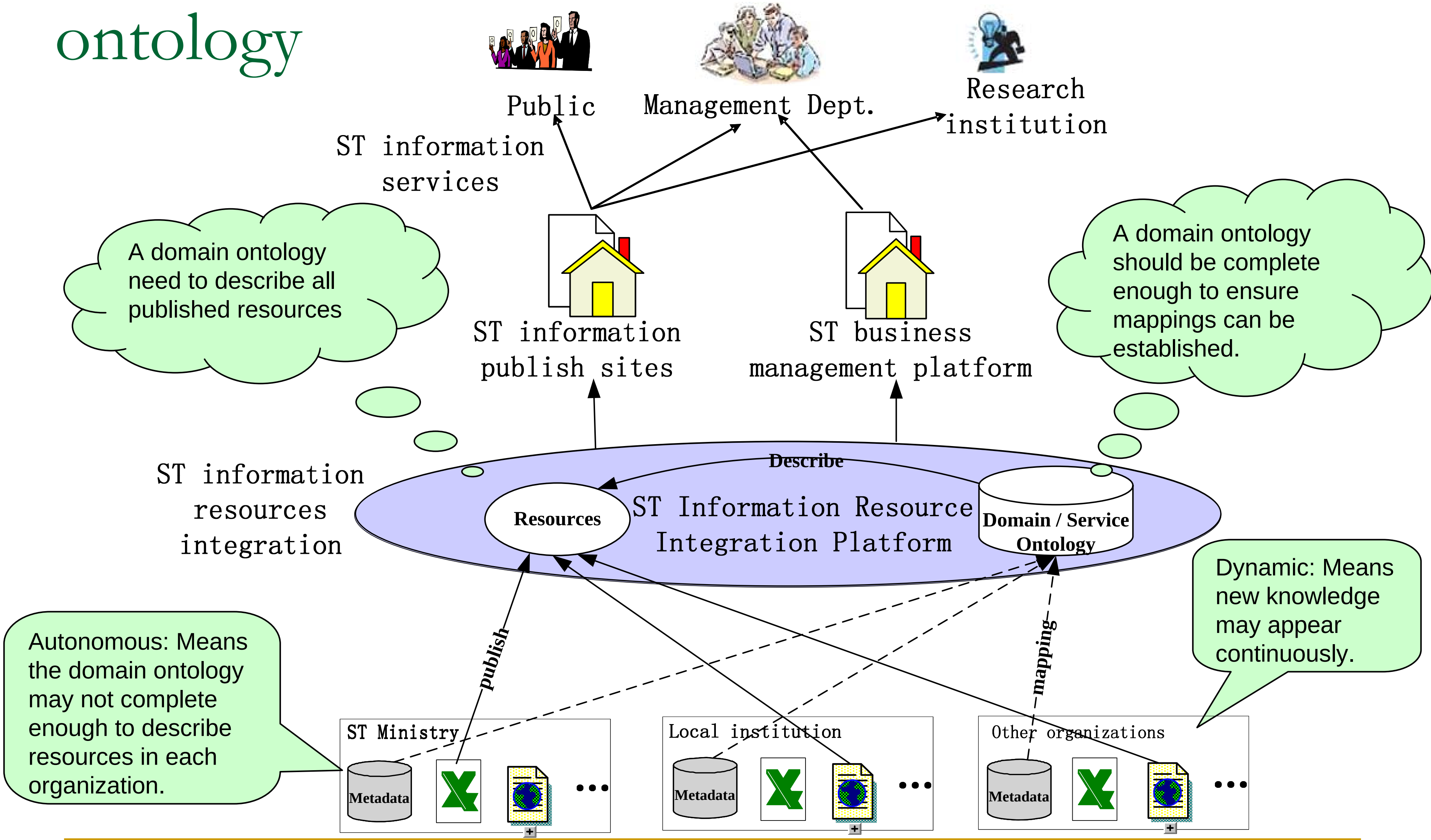
The functions of the ontology

- The ontology is mainly used to describe the services and their related resources. Only after adding semantic annotations to services, a VO can:
  - Help users to find their needed services exactly
  - Enable semi-automatic service composition

The semantic description of web services

- Describe the semantic web services depending on two kinds of ontologies:
  - A generic web service ontology: define what is a web service, such as owl-s
  - A domain ontology: describe the resources which related to the services

Back to the scenario : requirements for a domain ontology



ST : science and technology

Challenges for constructing a domain ontology in a VO

- The dynamics and autonomy of a VO plague the construction of a domain ontology
  - Cannot pre-construct a global ontology.
  - Cannot construct a stable ontology.
- Conclusion
  - A domain ontology in a VO should keep evolving.

Related Works

Traditional approaches

- Build-from-scratch approaches
  - Domain experts analyze the application domain and construct ontologies from scratch.
  - Enterprise Ontologies and skeletal methodology [7], TOVE ontology and evaluation methodology [8] as well as CHEMICAL and Methontology methodology [9].
- Build-by-integrating-others approaches
  - reusing other available ontologies to build a new ontology [11-12].
  - merging different ontologies about the same subject into a single one that “unifies” all of them [13-14].

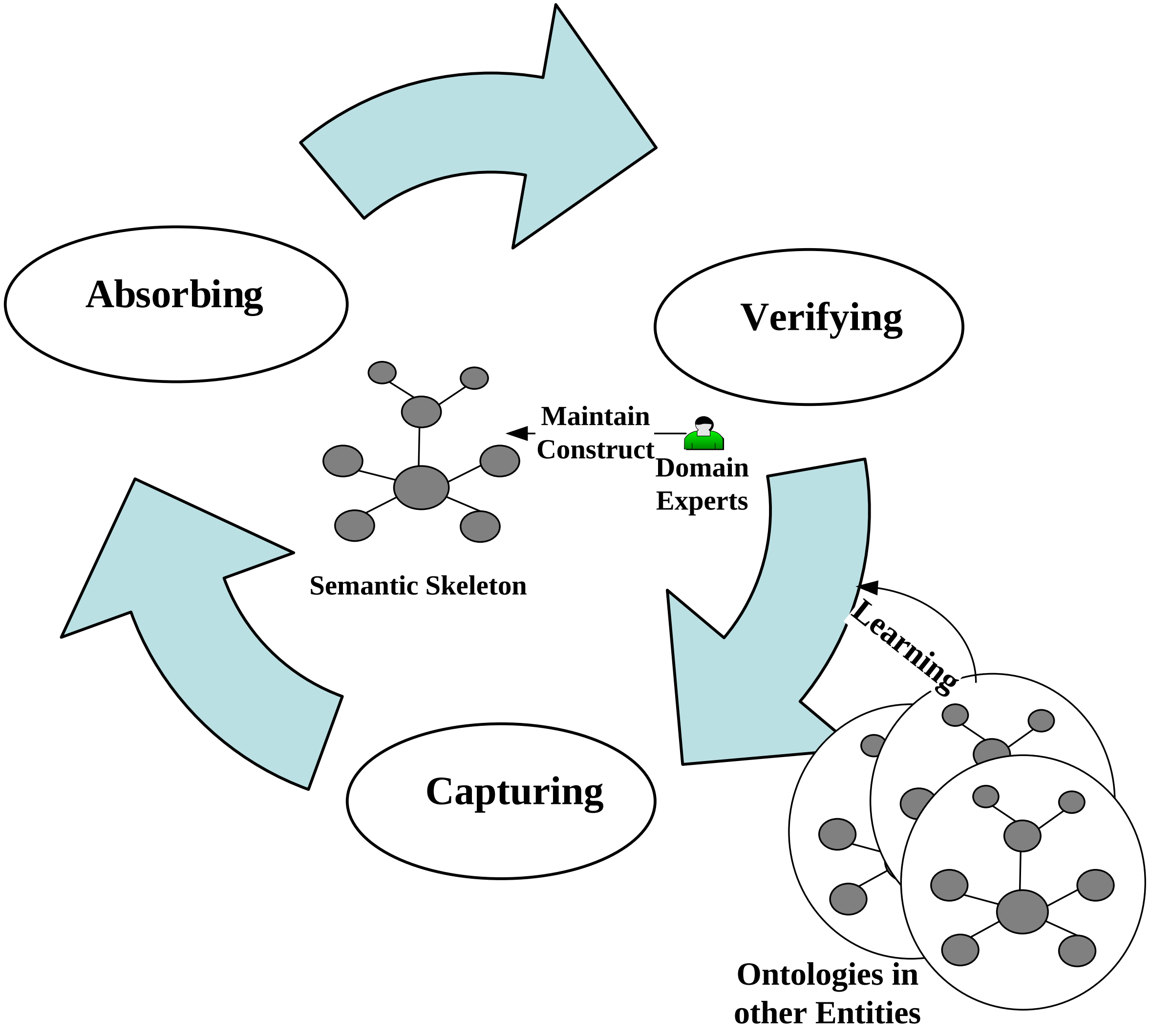
Analysis

- Build-from-scratch approaches are incapable
  - Construct ontologies beforehand and result in inflexible and hard-to-change ontologies.
- Build-by-integrating-others approaches are incapable
  - High cost and apt to lose original semantics heavily. [6]

Rationale of the DODO Mechanism

The process of constructing a domain ontology

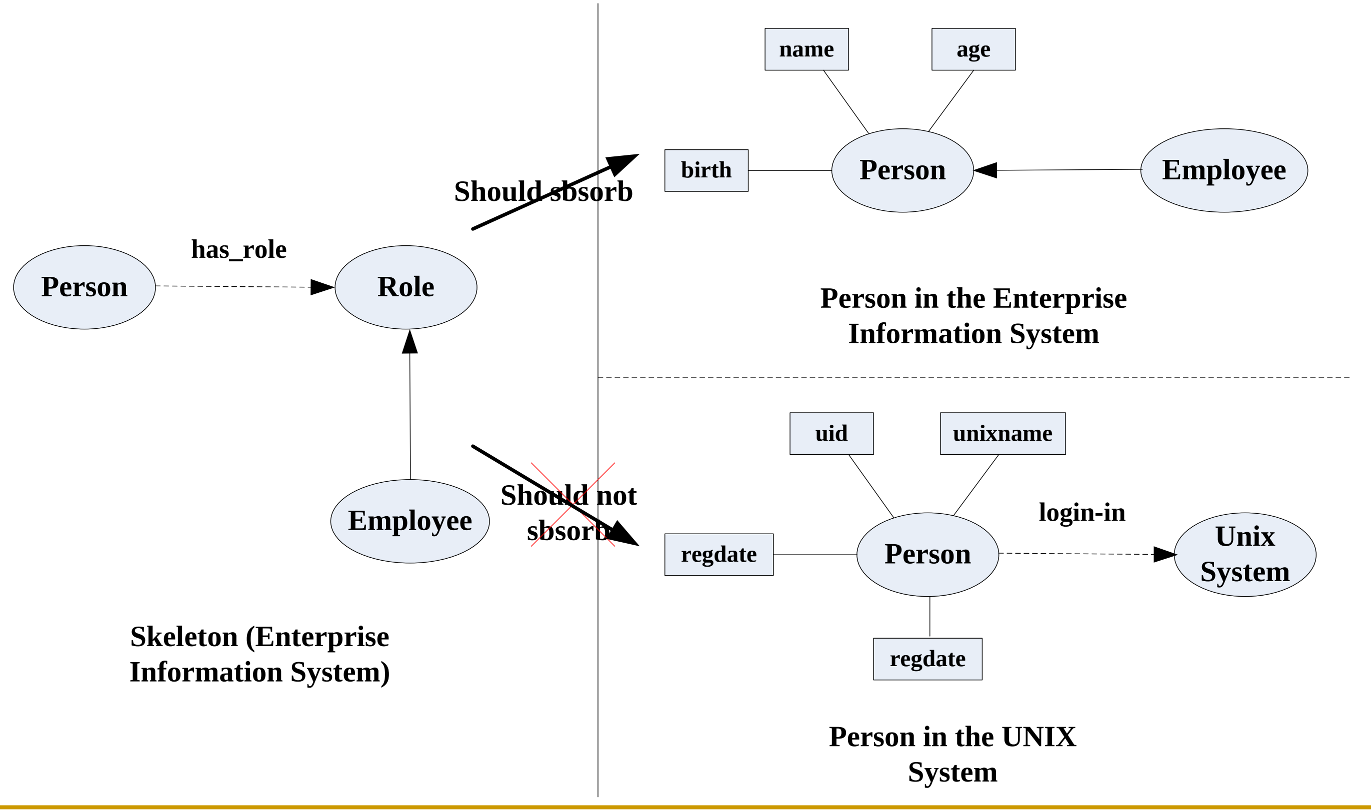
- DODO provides a dynamic iterative process for constructing a domain ontology.
  - Constructing the semantic skeleton
  - Capturing the knowledge
  - Absorbing the knowledge
  - Verifying the knowledge



Absorbing
Verifying
Capturing
Maintain
Construct
Domain
Experts
Semantic Skeleton
Learning
Ontologies in
other Entities

Step1: constructing the semantic skeleton

- Goal
  - To provide enough background knowledge to eliminate the possible ambiguities.



Definition of the semantic skeleton

- An semantic skeleton is a tuple S:= (C, P, Hc, Hp), where:
  - C: The set of key concepts of a given application domain.
  - P: The set of optional properties of a given application domain.
  - Hc: The hierarchy of predefined key concepts.
  - Hp: The hierarchy of predefined key properties.

Differences between semantic skeleton and ontology

- The semantic skeleton is much simpler
  - Doesn't have the complex constraints including data types.
  - Easy to construct.
- Semantic skeleton is incomplete
  - Important concepts
  - Important properties
- Semantic skeleton is steady
  - Changes to the semantic skeleton should be handled only by domain experts.

Step2 : Capturing the knowledge

- Goal
  - Capture new knowledge
  - Capture different knowledge
- Method
  - Matching the semantic descriptions of the published resources with current constructed ontology to find their similarities and differences.

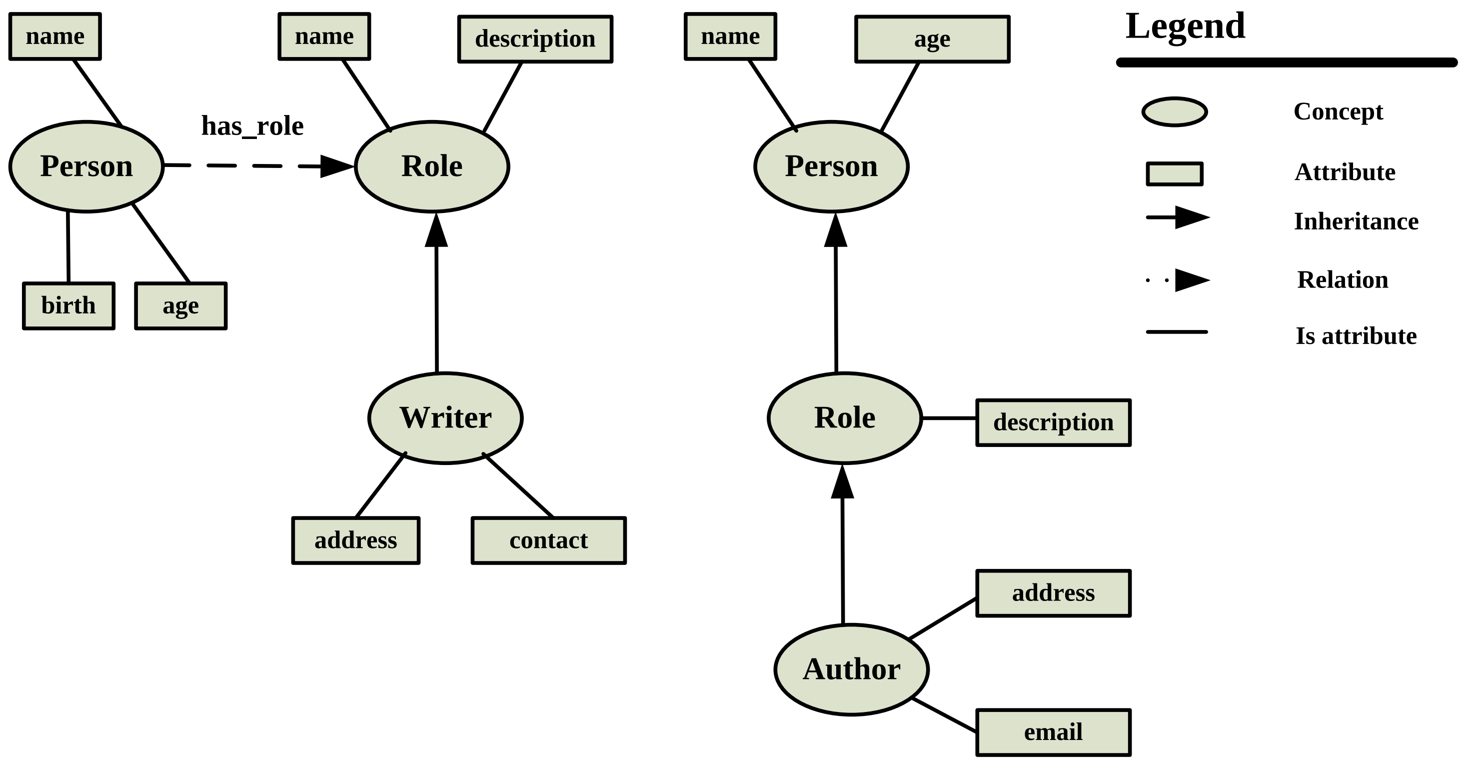
Matching algorithm

- H-Match algorithm [16-17]
  - Name Affinity :
    - To quantify the level of matching two concepts based on their names.
    - Refer to WordNet [15]
  - Structure affinity
    - To quantify the level of matching two concepts based on their attributes
    - Proportional to the number of similar attributes
  - Context affinity
    - To quantify the level of matching two concepts based on their contexts.
    - Proportional to the affinity of directly adjacent concepts
  - Global affinity: the linear combination of the above affinities

Matching results

- New knowledge
  - New concept
  - New attribute
  - New relation
- Update knowledge
  - Label change
  - Relation change
  - Domain change
  - Range change

An example



Constructing Ontology

Resource Provider A

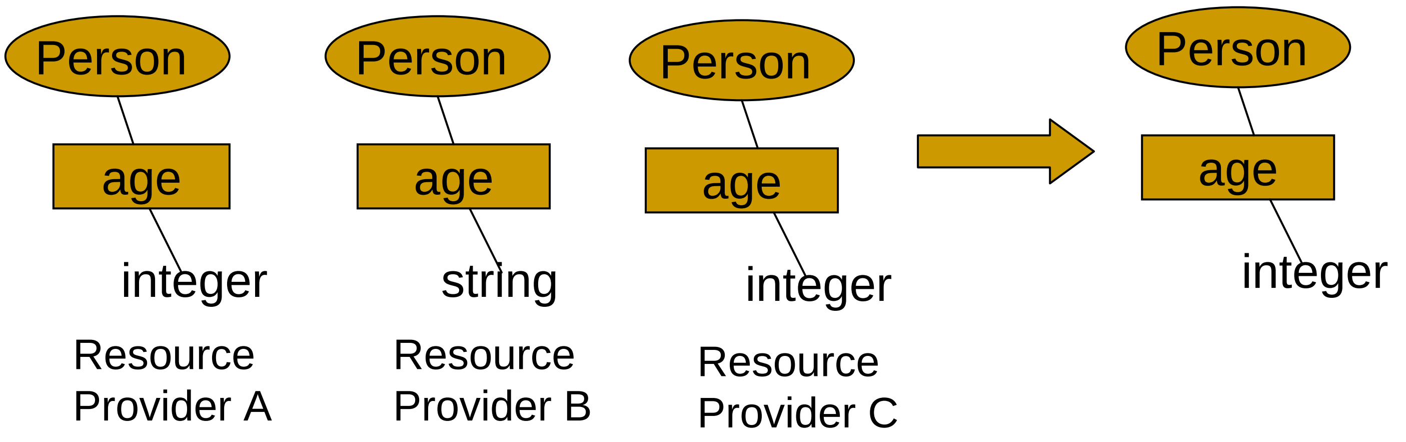
- Fragment of matching result
  - New attribute: Author.email
  - Label change: Writer->Author

Step3: Absorbing the knowledge

- Goal
  - To absorb captured knowledge under the supervising of the domain experts.
- Problems
  - What to absorb?
    - In the running of a VO, knowledge about even the same kind of resource may have multiple editions.
  - How to absorb?

The principle to absorb

- Consensus-driven principle: if a piece of knowledge is provided by most of resource providers, then it should be recommended in high priority.



The way to absorb

- Algorithm
  - For each kind of matching result, a sub-algorithm is developed to solve the ontology change problem.
  - Traverse all the matching result to implement the absorption.

Matching Result	Sub-algorithm
New concept	1 Try to find the super-concept in the current constructing ontology 2 If found, derive new concept from existed concepts 3 Otherwise, add a completely new concept
New attribute	1 add a completely new attribute
New relation	1 Try to find the super-relation in the current constructing ontology 2 If found, derive new concept from existed relations 3 Otherwise, add a completely new relation
Label change	1 remove the original label 2 add the new label
Relation change	1 remove the original relation 2 add the new relation
Domain change	1 remove the original domain 2 add the new domain
Range change	1 remove the original range 2 add the new range

Step4: Verifying the constructing ontology

- Consistency : means no contradictory conclusions can be drawn from valid input definitions.
- Completeness : means all that is supposed to be in the ontology is explicitly represented in it or can be inferred using other definitions.
- Non-redundancy : means that the constructed ontology should avoid the errors such as redefining expressions that were already explicitly defined or that can be inferred using other definitions.

Application

A real scenario

- Enterprise dynamic supplying chains
  - Goal : realizing resource sharing among a huge electronic manufacturing enterprise and its upriver and downriver partners
  - Scale : more than 20 enterprises and almost 100 kinds of resources are involved
  - Application
    - A VO is constructed to form the alliance among these organizations.
    - DODO mechanism is exploited to construct a flexible and scalable semantic support environment.

Conclusion

- DODO
  - Goal: Dynamically constructing a domain ontology for a VO
  - Characteristics
    - Runtime construction
      - The construction of ontology doesn’t need to be finished before the application.
      - Stresses the consummation of ontology in the runtime.
    - Lightweight
      - It avoids integrating two ontologies. On the contrary, it focus only on the semantic descriptions of sharable resources.

Reference

- 1 Ian Foster,Carl Kesselman,Steven Tuecke. The Anatomy of the Grid Enabling Scalable Virtual Organizations.
- 2 Wolff, M. (1995). Ki-Net - New Organizational Structures for Engineering Design, Web page.
- 3 Chen, L., Fanke, C., and Yanbo, H. SATOR: A Scalable Resource Registration Mechanism Enabling Virtual Organizations of Enterprise Applications. *In Proceedings of 4th International Conference on Grid and Cooperative Computing (GCC 2005).* LNCS, Volume 3795 / 2005, Springer-Verlag GmbH, 2005, 744 – 749.
- 4 Jones, D.M., Bench-Capon, T.J.M. and Visser, P.R.S. Methodologies for Ontology Development. *In Proceedings of IT&KNOWS Conference.* XV IFIP World Computer Congress, Budapest, August.
- 5 Bruijn, J. d., et al. State-of-the-art Survey on Ontology Merging and Aligning V1. SEKT-project report D4.2.1 (WP4), IST-2003-506826, 2003.
- 6 Yang, Y., Shanping, L., and Ming, G. A Comparative Study on Three Ontology Mapping Methods in the Integration of Product Knowledge. *In Proceedings of the 5th World Congress on Intelligent Control and Automation.* June 15-10,2004,Yangzhou,P.R,China
- 7 Uschold, M., King, M., Moralee, S., and Zorgios, Y. The Enterprise Ontology. The Knowledge Engineering Review, 13(1), 31-89.

- 8 Grüninger, M. and Fox, M. S. Methodology for the design and evaluation of ontologies. *In proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing.* Held in conjunction with IJCAI-95, Montreal, Canada.
- 9 Fernandez, M., Gomez-Perez, A. and Juristo, N. METHONTOLOGY: From Ontological Art Towards Ontological Engineering. AAAI-97 Spring Symposium on Ontological Engineering. Stanford University, March 24-26th.
- 10 Pinto, H.S., Gomez-Perez, A., and Martins, J.P. Some issues on ontology integration. *In Proceedings of the IJCAI-99 Workshop on Ontologies and Problem-Solving methods (KRR5).* Stockholm, Sweden.
- 11 Borst, P., Akkermans, H. and Top, J. Using Explicit Ontologies in KBS Development. Engineering Ontologies. International Journal of Human Computer Studies special issue, 46(2/3), 365–406.
- 12 Gruber, T. and Olsen, G. R. An Ontology for Engineering Mathematics. In E. S. J. Doyle and P. Torasso, eds, ‘KR94 Proceedings’, Morgan Kaufmann, 258–269.
- 13 Mitra, P. and Wiederhold, G. An algebra for semantic interoperability of information sources. In IEEE International Conference on Bioinformatics and Biomedical Egineering, 2001, 174–182.
- 14 Bergamaschi, S., Castano, S., Vincini, M. and Beneventano, D. Semantic integration of heterogeneous information sources. Special Issue on Intelligent Information Integration, Data & Knowledge Engineering, 2001, 215–249.
- 15 Miller, A.G. WordNet: A lexical database for english. Communications of the ACM, 1995, 38(11), 39–41.
- 16 Castano, S., Ferrara, A., Montanelli, S., and Zucchelli, D. HELIOS: a general framework for ontology-based knowledge sharing and evolution in P2P systems. In Proceedings of DEXA’03 2nd Web Semantics Workshop. Prague, Czech Republic, 2003.
- 17 Castano, S., Ferrara, A. and Montanelli, S. H-MATCH: an algorithm for dynamically matching ontologies in peer-based systems. In Proceedings of the First Workshop on Semantic Web and Databases (SWDB-03), co-located at VLDB 03. Berlin, Germany, 2003.

Thanks!

For more information about our group, please visit:

http://sigsit.ict.ac.cn

Or email to me:

- liuchen@software.ict.ac.cn