

---

# Social Infobox: Collaborative Knowledge Construction by Social Property Tagging

**Masahiro Hamasaki**

National Institute of Advanced  
Industrial Science and Technology  
/JST, CREST.  
1-18-13 Sotokanda  
Chiyoda-ku, Tokyo JAPAN.  
hamasaki@ni.aist.go.jp

**Masataka Goto**

National Institute of Advanced  
Industrial Science and Technology  
1-1-1 Umezono  
Tsukuba, Ibaraki JAPAN.  
m.goto@aist.go.jp

**Hideaki Takeda**

National Institute of Informatics  
2-1-2, Hitotsubashi  
Chiyoda-ku, Tokyo JAPAN.  
takeda@nii.ac.jp

---

Copyright is held by the author/owner(s).  
CSCW 2011, March 19–23, 2011, Hangzhou, China.  
ACM 978-1-4503-0556-3/11/03.

**Abstract**

We propose a novel style of social tagging to construct knowledge collaboratively called Social Property Tagging and introduce the prototype system Social Infobox. Structured data is useful for computer system, however defining structure of knowledge for representing data semantics is usually a costly and time consuming task. In general, data structures are constructed by experts of knowledge engineering. Our method aims to construct not only structured data but also structure of data collaboratively by simple user input.

**Keywords**

Collaborative tagging, information management, semantic web

**ACM Classification Keywords**

H.5.3 Information Interfaces and Presentation: Group and Organization Interfaces – web-based interaction, collaborative computing.

**General Terms**

Algorithms, Human Factors, Design

## **Introduction**

Defining structure of knowledge for representing data semantics is usually a costly and time consuming task. Ontology is one of powerful structures of knowledge and there are many methodologies and tools for construction. The most popular approach is engineering-oriented, i.e., a small number of individuals carefully constructs the representation of the domain of discourse, and provides them at some point in time to a wider member of community.

Gruber defined ontologies are explicit conceptualizations of a domain of discourse [1]. In context of semantic web, the representation of ontology should have a formal semantics because it leads machine readable and supports machine reasoning. However, it is important to note that ontologies are not just formal representations of a domain, but much more community contracts about such formal representations [2]. Since a discourse is a dynamic social process and such a community contract cannot be static, it must be able to reflect the community consensus at any point in time.

Wikipedia's Infobox is famous structure data constructed by CGM. However, it takes the popular approach, editors use pre-defined template. Some tools for collaborative construction of ontologies are proposed [3], but most of all them aim collaboration among experts or semi-experts who know set of vocabularies to represent structure of knowledge.

How does wisdom of crowds construct structure of knowledge? Social tagging has the potential to produce socially constructed information organization schemes [4, 5]. However, such the produced structure is simple

and it is hard to describe in detail. So in this paper, we propose novel style of social tagging for construct structured data.

## **How does wisdom of crowds construct structure of knowledge?**

It is hard for even a person who has knowledge to understand structure of knowledge. However it does not mean that there is no structure for knowledge in their mind. It seems that each person has a personal and weak structure for knowledge and it is not externalized and shared.

What method can externalize and share such a personal and weak structure of knowledge? We found that many users can externalize and share short contents which is immediately available for recall e.g., tags or comments from the success of social bookmarking. However, how does social tagging construct structure of knowledge?

Before that, we reconsider base structure of knowledge: class and property. Property indicates a characteristic of resource. People add a property to a resource in order to categorize or identify a resource from others. For example, we can guess tags for Tim B. Lee are "World Wide Web", "W3C", "Sir", etc. When we convert from these tags to attribute-value pairs, these tags are "creator: World Wide Web", "affiliation: W3C", and "title: Sir". At that time, we do not guess tags e.g., his blood type or stronger hand. It means that these characteristics are less important than the above when we categorize or identify Tim B. Lee. On the other hand, stronger hand will be tagged when we tag baseball players.

Form this scenario, we can find existence of a group of resources which we need to compare each other or identify one from them. When we suppose such an implicit group is a prototype of a class, we can find quasi class from popular set of properties. It means that wisdom of crowds constructs structure of knowledge.

### Proposed Method

In this research, we proposed **Social Property Tagging** as a method to construct knowledge collaboratively. It allows users to input information without formal definition of classes and properties and assists users to construct knowledge collaboratively. It consists of two components.

- **Property Tagging:** Users add attribute-value pair as a tag of a resource. It facilitates input data for users because the size of input becomes small and simple. It allows users to add new properties which they need. Then users can input data based on own style of structure of information.
- **Property Suggest:** It suggests properties and their values to users using input data. It helps data to construct structure and reduces burden of users to input data.

We develop the prototype system called **Social Infobox**<sup>1</sup> for collaborative knowledge construction by social property tagging (Fig.1). The proposed system allows users to read, create, delete, and update data like a Wiki. The data consists of Resource, Property, and Value. A resource has some properties and their

<sup>1</sup> <http://tinyurl.com/socialinfobox>

values (attribute-value pairs). This style of representation is similar to Wikipedia's Infobox. Users can add not only resources but also properties and their values freely without definition of a class or a property. So users can add properties to a resource easily like a social tagging. Wikipedia allows all users to write description but it is a hard work so a few people edit them in spite of the huge number of users. Adding a property and its value or filling a value to a property is not so hard for users so we expect that many users add metadata like a social tagging.

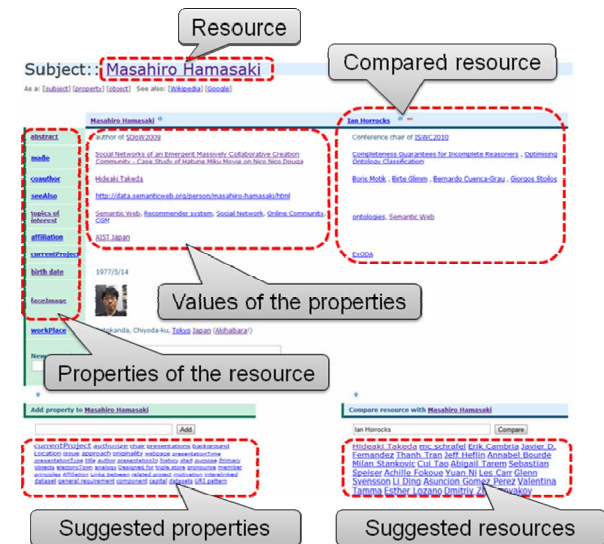


figure 1. Screenshot of the prototype system.

However, it merely increases variety of property only if users add properties freely. It cannot progress toward structuration of whole data. Then, the system suggests

properties using existing data. It aims to not only reduce user's load but also become structured. Viewed from interface aspect, Property suggest has 2 types.

**Direct Suggest** is the system suggests properties and values based on co-occurrence of properties because popular set of properties indicates quasi class.

**Comparison suggest** is the system suggests resources to compare based on similarity of properties and their values Because role of a property is describing characteristic to identify from others A user can find properties and values from comparison of a target and others

### User Study

In this section, we report user studies using our prototype system. Users of the prototype system are only few members include the developer and we have not find verified results yet. However, we think this report of user studies imply a possibility of the proposed system to help users to construct knowledge without explicit definition of and deep consideration for a class.

Firstly, we found a group of resources which shared some properties. For example, "Tim B. Lee" and "Masahiro Hamasaki" shared properties "affiliation" and "expertise". It seems that it is a (weak) internal definition of a class. Secondly, we found a group of resources which become a value of same properties. For example, "WWW", "IJCAI", and "AAAI" are values of property "International Conference". It seems that it is a (weak) external definition of a class and a name of the property is a name of that class. Thirdly, we found properties whose values are few varieties. "Type of Service" property is one of examples in this case. It has values, e.g., "Social Networking" or "Video Sharing". It

seems that these are (weak) external definition of classes and values are name of these classes.

### Conclusion

In this paper, we propose Social Property Tagging and introduce the system called Social Infobox for collaborative knowledge construction. The proposed system makes users to add resources and properties easily like social tagging, and suggests properties to construct set of vocabularies collaboratively that cover the classes and properties users need. It does not require explicit pre-defined structure and allows all potential users to describe information or knowledge using vocabularies which they deem necessary.

### Acknowledgements

This research was supported by JST, CREST.

### Citations

- [1] Thomas R. Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human-Computer Studies* 43 (1993), 907-928.
- [2] Martin Hepp, Daniel Bachlechner, Katharina Siorpaes. OntoWiki: Community-driven Ontology Engineering and Ontology Usage based on Wikis. *Proc. The 2006 International Symposium on Wikis* (2006).
- [3] Gianluca Correndo, Harith Alani. Survey of Tools for Collaborative Knowledge Construction and Sharing. *Proc. Workshop on CISW 2007* (2007), 2-5.
- [4] Peter Mika. Ontologies are us: A unified model of social networks and semantics. *Journal Web Semantics* 5 (2007), 1570-8268.
- [5] Emilee Rader, Rick Wash. Influences on tag choices in del.icio.us. *Proc CSCW '08*, ACM Press (2008), 239-248.