

Document annotation with a graphical software tool JGat

JGat is a Graphical annotation tool for indexing documents by using pre-defined classification systems like *thesauruses*, *taxonomies*, and *ontologies*. The classification systems define *concepts*, their *relationships*, and a *vocabulary* associated with the concepts. The vocabulary may be *multilingual* so that each concept and relationship has names in more than one language. Well-known examples of such systems are MeSH (Medical Subject Headings), HPMULTI (European multilingual thesaurus on health promotion), and SNOMED CT (Systematized Nomenclature of Medicine–Clinical Terms).

JGat allows any number of classification systems to be plugged in. The user of the program can easily change the conceptualization environment even when indexing a single document. This applies also to the language of the vocabulary. If the installation of the classification system includes term translations, then the user can change the language while indexing the document, and look up concepts in different languages.

JGat represents concepts graphically as *interactive hyperbolic trees* (see picture). The program uses two kinds of tree views, which are called here *hierarchic taxonomical* and *ontological views*. The taxonomical tree shows generalization relationships between concepts. This type is also called 'is a kind of' -relationship, and is commonly used in classification. Most of the ontologies include it as one of the predicate types they support. Taxonomical trees are also called *generalization trees*. The picture below describes the taxonomical view.

If an installed classification system has other types of relationships registered, then JGat can also show an ontological view for a concept. The common usage of this approach is that the user selects a concept from a taxonomical tree by clicking the mouse on a node. Then the program shows the conceptual context of the selected topic. The resulting graphical representation is a hyperbolic tree as well, but the nature of the tree is different from the generalization tree. The tree nodes represent both the concepts and the types of relationships between those. The relationship types are called *predicates*, or *properties*. The concept leading to the predicate is called a *subject*. The arcs that leave the property node point to concept nodes representing *objects* of the relationship. The tree thus elaborates the original concept of interest by using a collection of subject–predicate–object -relationships connected to each other so that the object in one relationship plays the role of the subject in the next one.

MeSH, as an example, defines generalization relationships between its concepts called *descriptors*. But each descriptor can be characterized also by other predicates such that 'can be qualified by', 'is related to disciplines', 'has pharmacological actions', 'has sub concepts', 'sub concepts are called by terms', and 'sub concepts has semantic types'. If the user clicks his/her mouse on a node in a hierarchical MeSH tree, he/she can then ask for a characterization of the descriptor with the concepts like these. The program displays them in a separate hyperbolic tree representing the ontological map of the object concepts like 'qualifiers', 'disciplines', and so on.

When the user decides to index a document by a concept displayed, he/she clicks the mouse on the

corresponding node and selects the corresponding entry from the list of possible actions shown in a popup menu. But he/she can also use the ontological views during the annotation sessions. Thus, when using MeSH, one can further specify the aspect how the document handles the topic. In MeSH, the aspect concepts are called *qualifiers*.

Classification systems can be very large, for example MeSH enumerates more than 23000 concepts and SNOMED CT many more. The systems try to describe the concept as well as they can do by *terms* of different languages. Each term may have *synonyms*. By the way, the predicate leading from the term to its synonym can be characterized by the statement 'means something like'. Anyway, the concept the user wants to use when indexing a document, is searched by using terms and their synonyms. The user enters a string of letters, and JGat searches from the catalog of all those terms starting with the string, and displays them. The user can then select a term matching with the search condition, after which JGat shows the taxonomical neighborhood of the selected concept. He/she can then browse by dragging the tree and select the concept itself, or some of the narrower or broader concepts nearby. The user can expand the view shown to include more detailed terms, for example. Or, he/she may want to go towards the root of the taxonomical tree in order to see sibling or cousin branches related to a node displayed in the original view.

References

- Gruber T (2001): [What is an Ontology?](#)
- HPMULTI <http://www.hpmulti.net/>
- MeSH <http://www.nlm.nih.gov/mesh/meshhome.html>
- SNOMED CT <http://www.snomed.org/>
- Lamping J, Rao R, Pirolli P (1995): [A focus+context technique based on hyperbolic geometry for visualizing large hierarchies.](#)

Applet Viewer: fi.ktl.julia.htanno.AnnotationApplet

Applet

Public Health Practice

In English suomi

Document URL
http://www.ktl.fi/portal/suomi/julkaisut/kansanterveyslehti/lehdet_2005/7_2005/tupakka_ja_alkoholi_pikkulasten_elokuvale

KTL Pub/MeSH 2002

Search Term
Bio

Biological Sciences

Health Education
Public Health Practice

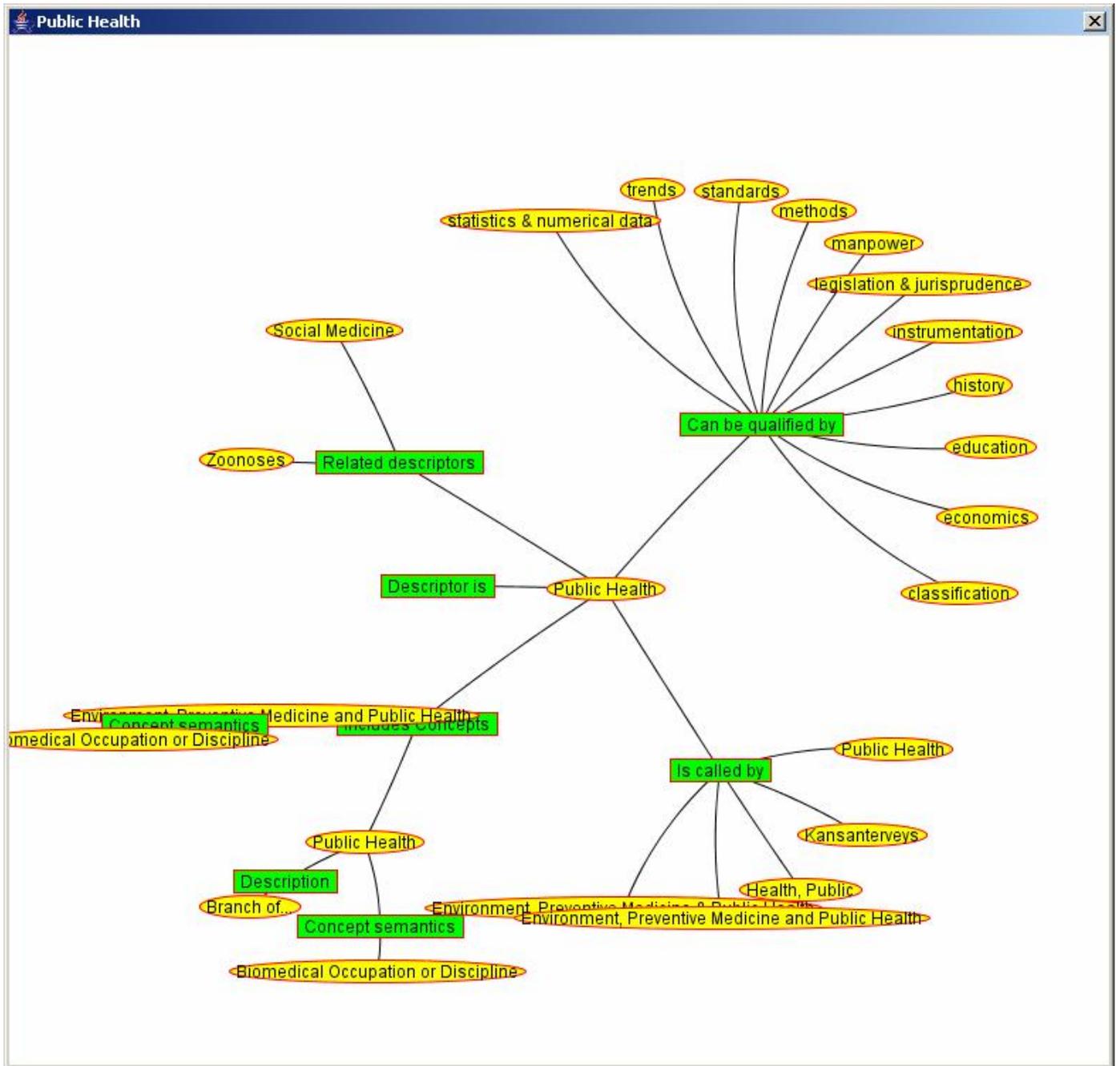
Save Clear

Applet started.

Public Health

- Disease Transmission, Patient-to-Professional
- Food Microbiology
- Disease Transmission
- Health Transition
- Disease Reservoirs
- Equipment Contamination
- Chemical Case Outbreaks
- Pesticide Residues
- Environmental Pollution
- Epidemiologic Factors
- Health Education
- Epidemiologic Methods
- Consumer Product Safety
- Sanitation
- Carrier State
- Public Health Practice
- Epidemiologic Methods
- Endemic Diseases
- Environment
- Environmental Health
- Environment and Public Health
- Biological Phenomena
- Cell Physiology
- Biological Phenomena, Cell Physiology, and Immunity
- Environmental Health
- Health Occupations
- Nutrition
- Nursing
- Specialties
- Medicine
- Biological Sciences
- Genetics
- Specialties and Pharmacologic Phenomena
- Fresh Water
- Cities later
- Environment, Controlled
- Disasters
- Soil
- Conservation of Natural Resources
- Energy-Generating Resources
- Seasons
- Atmosphere
- Meteorological Factors
- Ecology
- Immune Vaccination
- Primary Prevention
- Preventive Medicine
- Dental Health Surveys
- Public Health
- Health Education, Dental
- Wound Healing
- Biological Phenomena
- Cell Physiology
- Biological Phenomena, Cell Physiology, and Immunity
- Environmental Health
- Health Occupations
- Nutrition
- Nursing
- Specialties
- Medicine
- Biological Sciences
- Genetics
- Specialties and Pharmacologic Phenomena

A screenshot was taken from the display of the current JGat version 0.1. The taxonomy displayed presents the subpart of the MeSH tree 'Biological Sciences', on which researches in KTL had been published before 2002.



A screenshot describing an ontological view.