A wealth of information on plant anatomy and morphology is available in the current and historical literature, and molecular biologists are producing massive amounts of transcriptome and genome data that can be used to gain better insights into the development, evolution, ecology, and physiological function of plant anatomical attributes. Integrating anatomical and molecular data sets is of major importance to the field of wood science, but this is often hampered by the lack of a standardized, controlled vocabulary that allows for cross-referencing among disparate data types. One approach to overcome this obstacle is through the annotation of data using a common controlled vocabulary or “ontology” (Ashburner et al. 2000; Smith et al. 2007). An ontology is a way of representing knowledge in a given domain that includes a set of terms to describe the classes in that domain, as well as the relationships among terms. Each term can be associated with an array of data such as names, definitions, identification numbers, and genes involved. Ontologies are fundamental for unifying diverse terminologies and are increasingly used by scientists, philosophers, the military and online web search engines. In an ontology, terms are carefully defined, allowing a wide array of researchers to (1) use terms consistently in scientific publications or standardized handbooks on quality/trait evaluations, and (2) search for and integrate data linked to these terms in anatomical, genetic, genomic, and other types of biological databases.

The Plant Ontology (PO, www.plantontology.org) is a structured vocabulary and database resource that links plant anatomy and development to gene expression and phenotypic datasets from all areas of plant biology (Jaiswal et al. 2005; Avraham et al. 2008). The Wood Ontology project (http://wiki.plantontology.org/index.php/Wood_anatomy_ontology_meeting__2012_at_NYBG__agenda), which was recently initiated during the Wood Ontology Workshop at the New York Botanical Garden on 1) Netherlands Centre for Biodiversity Naturalis, P.O. Box 9514, 2300 RA Leiden, The Netherlands. 2) Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97331, USA. 3) L.H. Bailey Hortorium, Department of Plant Biology, Cornell University, Ithaca, NY 14853, USA. 4) Institute of Forest Genetics, US Forest Service, Davis, CA 95618, USA. 5) Department of Forest Ecosystems & Society, Oregon State University, Corvallis, OR 97331, USA. 6) Department of Botany, Connecticut College, New London, CT 06320, USA. 7) Clemson University Genomics Institute, Clemson University, 51 New Cherry Street, Clemson, SC 29634, USA. 8) New York Botanical Garden, Bronx, New York 10458, USA. 9) Department of Plant Sciences, Davis, CA 95616, USA. *) Corresponding author: Frederic Lens [E-mail: lens@nhn.leidenuniv.nl].
Figure 1. A simple ontology diagram showing a subset of wood anatomy terms from the Plant Ontology. This diagram illustrates the two most important relationships in the ontology (is_a and part_of) and shows how data annotations are associated with ontology terms. For example, a tracheid (PO:0000301) is_a type of tracheary element (PO:0000290) and all tracheary elements are part_of some xylem tissue (PO:0005352). Using the relationships specified in this ontology, a computer or human could infer that any tracheid is part_of some xylem tissue. Similarly, it can be inferred that a tracheid is a plant cell (PO:0009002), that xylem is a type of plant tissue, and that both are plant structures (PO:0009011).

The ontology also facilitates genomics studies through data annotations, such as genes expressed in plant structures represented by terms in the Plant Ontology. For example, px1, a class III peroxidase-encoding gene from Picea abies, is associated to tracheid via its effect on lignin biosynthesis during tracheid development (Marjamaa et al. 2006). Since a tracheid is part_of xylem, xylem also exhibits the phenotypic qualities of px1. Similarly, PgMYB2, a transcription factor that acts as regulator of lignin and phenylpropanoid metabolism during wood formation in Picea glauca, is associated to the general term xylem, since the experiment did not specify the exact cell type it was expressed in (Bedon et al. 2007). If a user searches the Plant Ontology for annotations on xylem and its subtypes, both px1 and PgMYB2 will be retrieved.

If the homologs of these genes are known for other tree species, a researcher can query the mutant and gene expression data for those species to look for genes regulating wood quality and development.