

Plant data visualisation using network graphs

Afrina Adlyna Mohamad-Matrol¹, Siow-Wee Chang¹ and Arpah Abu^{1,2}

¹Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia

²Centre of Research for Computational Sciences and Informatics for Biology, Bioindustry, Environment, Agriculture and Healthcare, University of Malaya, Kuala Lumpur, Malaysia

ABSTRACT

Background: The amount of plant data such as taxonomical classification, morphological characteristics, ecological attributes and geological distribution in textual and image forms has increased rapidly due to emerging research and technologies. Therefore, it is crucial for experts as well as the public to discern meaningful relationships from this vast amount of data using appropriate methods. The data are often presented in lengthy texts and tables, which make gaining new insights difficult. The study proposes a visual-based representation to display data to users in a meaningful way. This method emphasises the relationships between different data sets.

Method: This study involves four main steps which translate text-based results from Extensible Markup Language (XML) serialisation format into graphs. The four steps include: (1) conversion of ontological dataset as graph model data; (2) query from graph model data; (3) transformation of text-based results in XML serialisation format into a graphical form; and (4) display of results to the user via a graphical user interface (GUI). Ontological data for plants and samples of trees and shrubs were used as the dataset to demonstrate how plant-based data could be integrated into the proposed data visualisation.

Results: A visualisation system named plant visualisation system was developed. This system provides a GUI that enables users to perform the query process, as well as a graphical viewer to display the results of the query in the form of a network graph. The efficiency of the developed visualisation system was measured by performing two types of user evaluations: a usability heuristics evaluation, and a query and visualisation evaluation.

Discussion: The relationships between the data were visualised, enabling the users to easily infer the knowledge and correlations between data. The results from the user evaluation show that the proposed visualisation system is suitable for both expert and novice users, with or without computer skills. This technique demonstrates the practicability of using a computer assisted-tool by providing cognitive analysis for understanding relationships between data. Therefore, the results benefit not only botanists, but also novice users, especially those that are interested to know more about plants.

Submitted 30 March 2018

Accepted 12 August 2018

Published 31 August 2018

Corresponding author

Arpah Abu, arpah@um.edu.my

Academic editor

Massimiliano Zanin

Additional Information and
Declarations can be found on
page 22

DOI [10.7717/peerj.5579](https://doi.org/10.7717/peerj.5579)

© Copyright

2018 Mohamad-Matrol et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Bioinformatics, Human-Computer Interaction, Computational Science

Keywords Network graph, Data visualisation, Ontology, PlantViz, Plant knowledge

INTRODUCTION

Plants play an important role and benefit all forms of life. In total, 452 vascular plant families and ~381,910 species have been identified by botanists across the world, as reported by [Willis \(2017\)](#). The biological field is full of diversified data inputs ([Marx, 2013](#)). Plant data range from complete genome sequences to geographical information of plant species distribution ([Armstead et al., 2009](#); [Hughes, 2006](#)). In addition, advancement in research technology and methodology has caused plant data to increase rapidly. However, shortage of skill and time to analyse these data points continue to be significant obstacles. Relationships between data are often left out, which eventually produce less valuable information. In addition, any constraints in textually inspecting the massive amount of data collected may cause valuable points to be discarded ([Keim, 2002](#)). Associating data, that is, related to one another, however, will point out unique points of the data, which will provide new information on various knowledge domains. Hence, it is important for researchers to interpret valuable data and present it in an engaging format so that it is easy for the community to understand. This is because viewing data in plain form such as in texts and tables will not be sufficient to provide clear explanation.

Data visualisation is a comprehensive field involving the crossover between mathematics, computer science, cognitive and perception science and engineering ([Telea, 2014](#)). Data visualisation is defined as the representation of data using a visual or artistic approach rather than the traditional reporting method ([Yuk & Diamond, 2014](#)). Data visualisation plays an important role in many fields such as business ([Tegarden, 1999](#)), geography ([Groenendyk, 2013](#)) and biology ([Chen et al., 2014](#); [Jensen & Papin, 2014](#); [Sedova, Jaroszewski & Godzik, 2016](#)). Visualised data better conveys the unique properties of the data it represents. In particular, visualising biological data helps researchers to view the data from a different angle, leading to new insights. This is because data shown via graphical representation makes it easier for humans to conduct data analysis, as it provides more cognitive support ([Tory & Moller, 2004](#)).

The key to effective data visualisation is the selection of the right type of visualisation to match the type of data used such as charts, network graphs, Sankey diagrams and tree maps. Most data visualisations are interactive, enabling users to manipulate and explore the visualisation instead of perusing only a set of fixed diagrams. For instance, IHME Viz Hub ([IHME, 2017](#)) compiles a number of datasets related to health problems around the world in the form of interactive maps. Consensus PathDB ([Kamburov et al., 2011](#)) is a database that integrates different types of functional interactions between physical entities in the cell and uses network graphs to show the interactions between proteins. Apart from that, EcoCyc ([Keseler et al., 2013](#)) is a comprehensive database outlining the genome sequence of *Escherichia coli*, where molecular data is displayed in a detailed timeline diagram. In addition, EcoCyc also allows users to customise the display to suit their interests, thus enabling them to move from one region of a genome to another.

Another key to effective data visualisation is the combination of functioning visualisation tools. A good visual library that consists of a set of programming languages

helps in the design of any kind of visualisation needed. D3.js ([Teller, 2013](#)) is an example of a JavaScript-based library, that is, used for creating dynamic and interactive data visualisations in a web browser. It heavily utilises Cascading Style Sheets, Hypertext Markup Language and Scalable Vector Graphics (SVG) standards, which provide controls over the final result to users. D3.js is applied in visualising different types of data such as the speed and direction of the wind ([Cook, 2017](#)), obesity rate of adults ([Map, 2017](#)) and property prices ([Jackson, 2015](#)).

A few botanic-related visualisation systems have been developed previously based on different motivations. For instance, Gramene ([Tello-Ruiz et al., 2016](#)) is an open source database that focuses on genomics in crops and a number of model plant species. This database is curated with controlled vocabularies from a set of ontologies. Besides that, the Botany Array Resource ([Toufighi et al., 2005](#)) provides web-based tools of microarray data specifically for plant species and some animal species. Another example is Ensemble Plants ([Bolser et al., 2016](#)), which is an online database that contains genome-level information for 39 plant species and functional tools for genomic alignments, functional annotation and other purposes. Apart from that, data visualisation is also used for modelling a virtual three-dimensional plant model in order to simulate the growth process of crops such as the tomato plant modelling system ([Lu, Deng & Fei, 2015](#)). However, most available botanical visualisation systems lack interactive elements that allow users to manipulate data, as the systems focus only at the genomic level.

From the studies discussed above, it can be concluded that the huge number of plant data is due to the wide range of plant types and species in existence, together with advancements in research technology. This vast amount of data requires an effective way of presentation for both researchers and members of the community, particularly in a more structured and interactive form. Presenting data in visual form helps to convey a deeper meaning of the data to users, thus encouraging knowledge inference among them. Two important factors that need to be considered in data visualisation are the types of visualisation that best suit the data and the types of visualisation tools for development.

This study looks into how data visualisation can be applied in deducing relationships between plant data, for example, the relationship between taxonomical data, between samples (i.e. physical sample of the plant) and between taxonomical data and samples. Thus, a visual representation for plant data is proposed. The objectives of this study are: (1) to integrate plant-based data, which consists of taxonomical data (in textual format) and samples (in textual format and images); (2) to transform text-based results in Extensible Markup Language (XML) serialisation format into graphical form; and (3) to develop a visualisation system for plant data. In addition, a user evaluation for expert and novice users is conducted to evaluate the proposed visual representation of plant data.

MATERIALS AND METHODS

Visual-based representation framework

The framework used in this study illustrates the flow of data representation from the database, that is, transformed into visualisation. [Figure 1](#) shows a generic flowchart

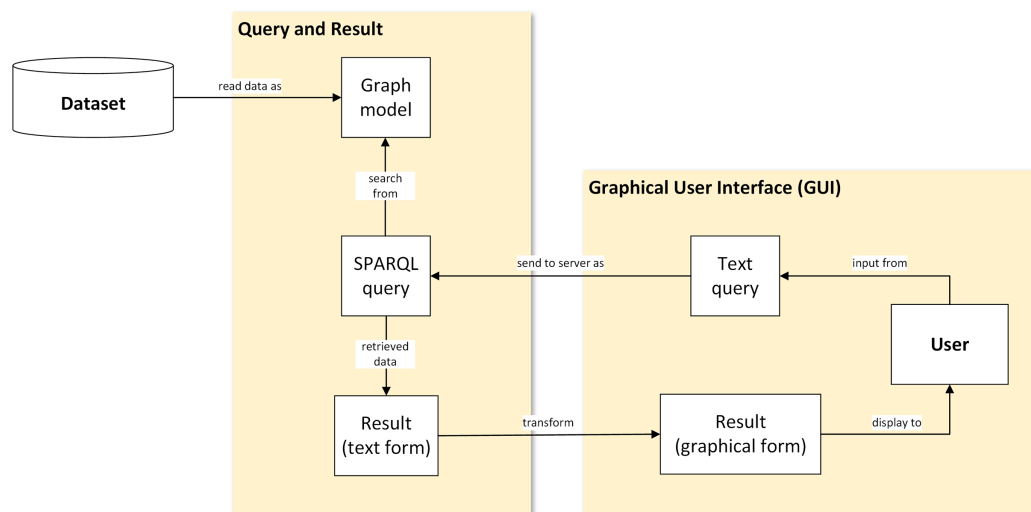


Figure 1 A proposed visual-based representation framework. Generic flowchart of plant data visualisation. [Full-size !\[\]\(5f471a71b78d7676bc356df190b88ab4_img.jpg\) DOI: 10.7717/peerj.5579/fig-1](#)

of the proposed visual representation for plant data. There are four main steps involved: (1) conversion of the ontological dataset into graph data model; (2) query from graph data model; (3) transformation of text-based results in XML serialisation format into graphical form; and (4) display of the result to the user via a graphical user interface (GUI).

Dataset

The dataset used in this study was obtained from Plant Ontology UM (POUM), an ontology that contains tree and shrub data collected from University of Malaya (UM), Kuala Lumpur, Malaysia. The following steps show the preparation of the dataset used in this study.

Step 1: Define the plant data description

Figure 2 shows the plant data description in this study. Taxonomical classification describes the rank of a plant from species to kingdom level along with the common name, authorship and year. Plant morphological characteristics describe the features of the bark, leaf, fruit, flower and the plant as a whole tree, at the species level. Ecological attributes describe the level of water usage for plant growth, as well as the type of soil and habitat. Geological distribution describes the location and GPS coordinates of the plant sampling area.

In addition, the plant samples for each species were described and images of the leaf, flower, bark and fruit of the whole tree were collected as well. The morphological characteristics, ecological attributes and geological distribution of plant samples were described individually due to the fact that the difference of habitat environments at plant sampling areas (see Fig. 3) affect the appearances of the plant samples.

Step 2: Build the proposed ontology

Plant data description as shown in Fig. 2 was then translated into the proposed ontology schema as shown in Fig. 4. The concepts (the entity), the concepts' properties (the data

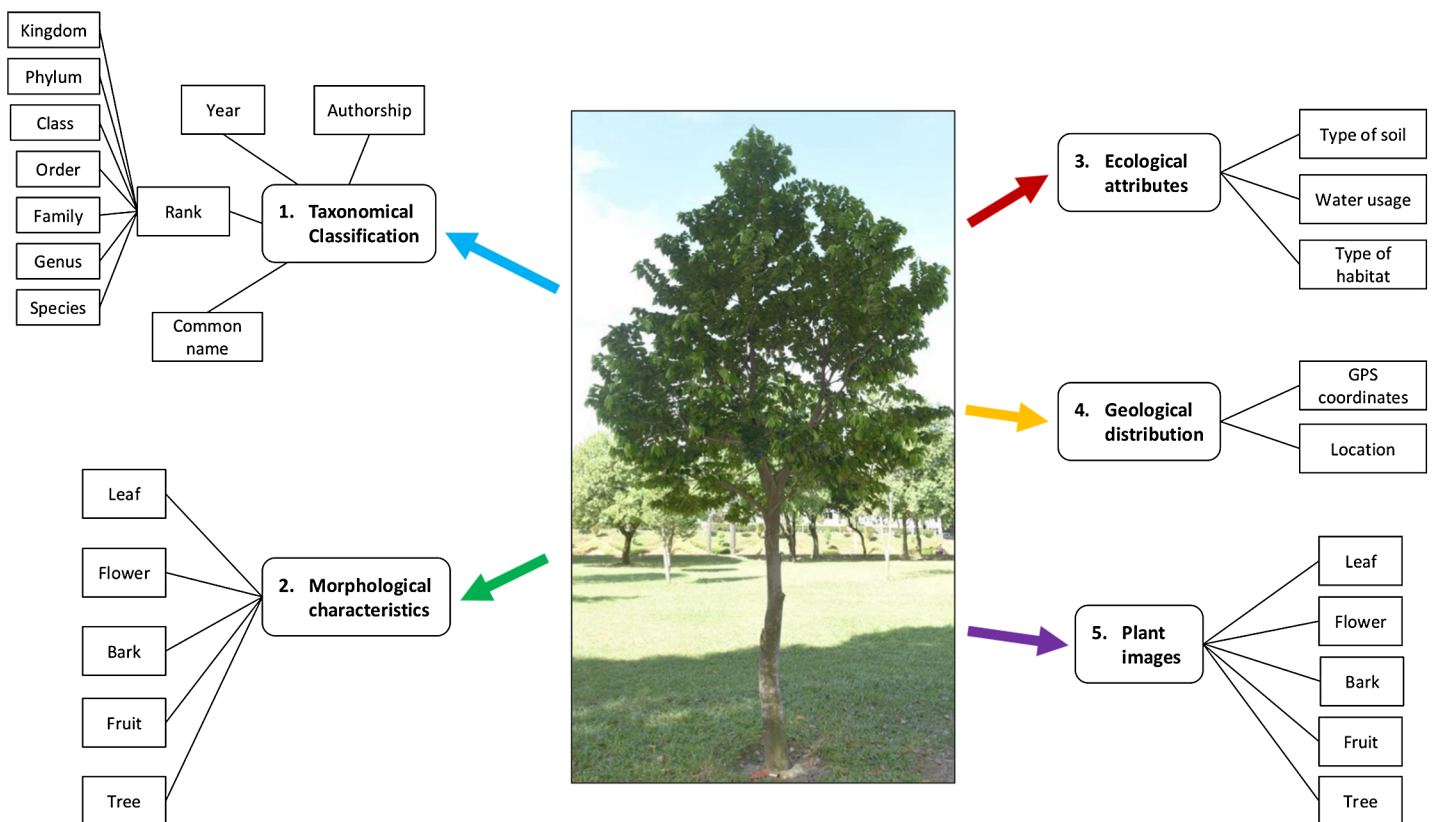


Figure 2 Plant data description. A plant is described with the taxonomical classification, morphological characteristics, ecological attributes, geological distribution and the plant images. Photo credit: the photo archive at UM Plant Knowledge (<http://103.18.1.10:8080/plantviz/>). Full-size DOI: 10.7717/peerj.5579/fig-2

property) and the relationships between concepts (the object property) of the domain were determined in this schema.

To prevent any misunderstanding in the terms used to describe the plant, a standardised set of vocabularies was designed. These vocabularies were adapted from the existing schema, namely the Biodiversity Information Standards (TDWG, 2018) which consists of a number of specific biodiversity data standards such as Life Sciences Identifier, Darwin Core and TDWG Access Protocol for Information Retrieval. Besides that, few vocabularies were also newly defined because these vocabularies are unavailable in any existing schema.

This ontology schema was then converted into an ontological form, a machine-readable formal specification (in *owl* file format) which proceeded to the reasoning process to complete the process of plant data annotation. Protégé 5.2 was used for these purposes. Figure 5 shows a partial of POUM using OntoGraph plug-in in Protégé 5.2.

There are more than 200 images of plants in the image database. These images are annotated with vocabularies from the POUM, consisting of eight main classes, five sub-classes, 17 object properties and 39 data properties (Table S1). There are 43 species of 42 genera for trees and 31 species of 28 genera for shrubs (Tables S2 and S3) with a total of 222 samples. As shown in Fig. 6, the image of a plant sample is described using sample information such as attributes of the sample's bark, leaf, flower and

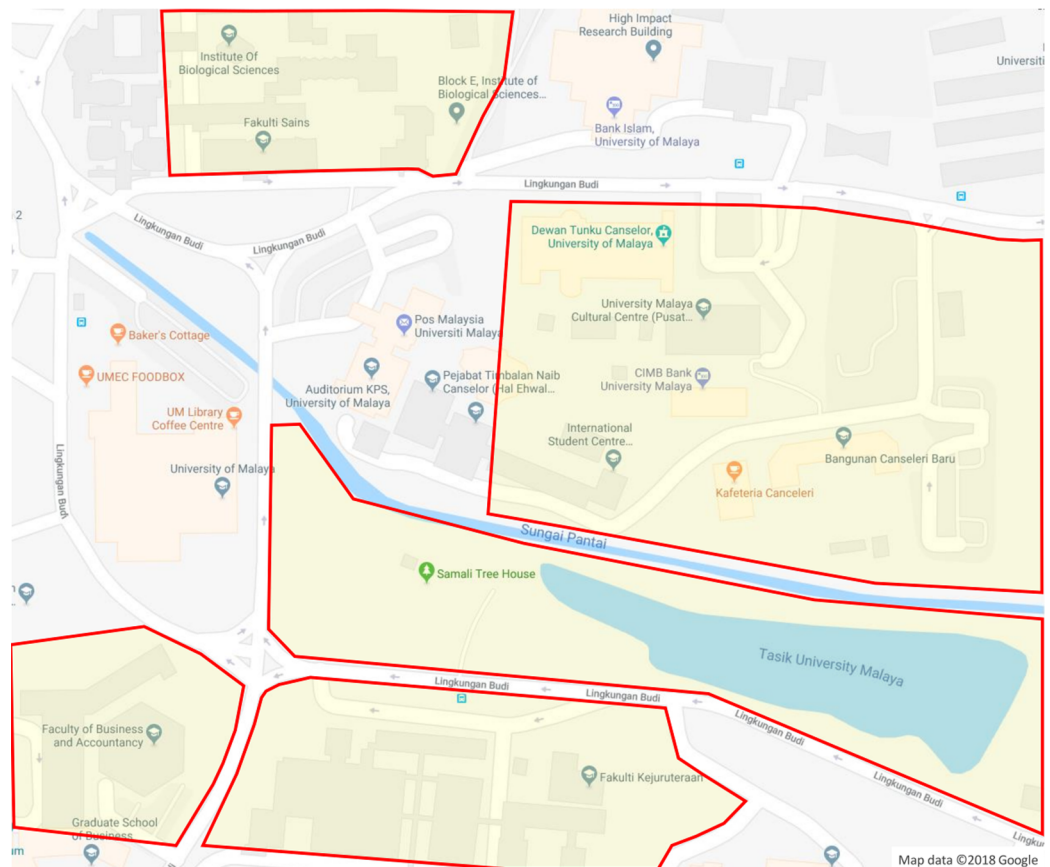


Figure 3 Plant sampling areas in UM. There are five main locations: DTC, Fakulti Kejuruteraan, Fakulti Perniagaan dan Perakaunan, Fakulti Sains and Tasik Varsiti. Map data © 2018 Google.

[Full-size](#) [DOI: 10.7717/peerj.5579/fig-3](https://doi.org/10.7717/peerj.5579/fig-3)

fruit, object of the image, location of where the sample was collected, and taxonomic data that focuses on taxonomical ranks and plant characteristics used in describing the plant species.

Workflow

Figure 7 shows the workflow of passing ontological data into the creation of a network graph.

Step 1: Conversion of ontological dataset into a graph data model

Ontological data of the POUM in *owl* file format is commonly encoded in a Resource Description Framework (RDF) data model with XML syntax. *Jena (2017)* introduced the serialisation of ontological data into an RDF graph data model before it was queried.

Step 2: Query from graph data model

The query languages used were Simple Protocol and RDF Query Language (SPARQL). RDF query language retrieves and manipulates any data stored in RDF format (*Harris, Seaborne & Prud'hommeaux, 2013*). Once users submit a text query, an SPARQL

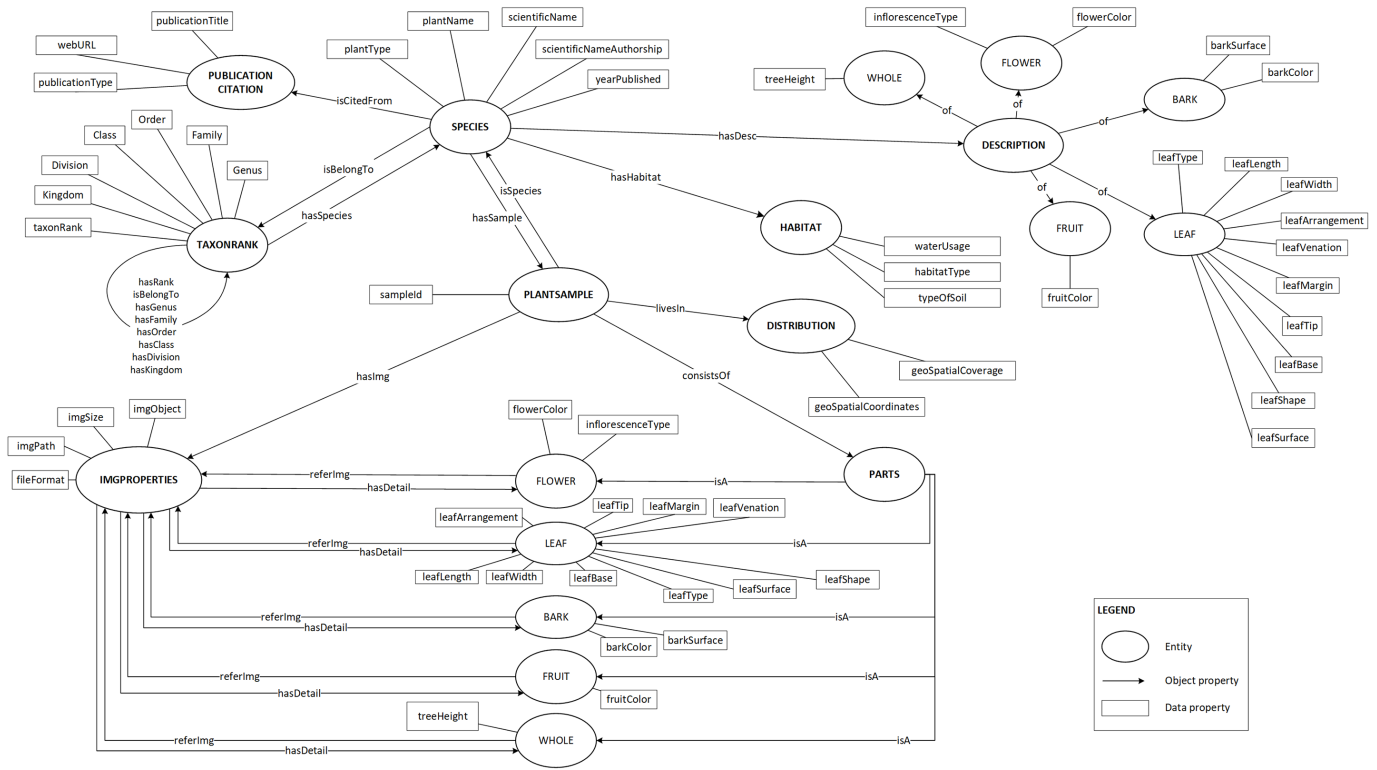


Figure 4 Proposed ontology schema. The plant data description translated into ontology in a graph format.

Full-size DOI: 10.7717/peerj.5579/fig-4

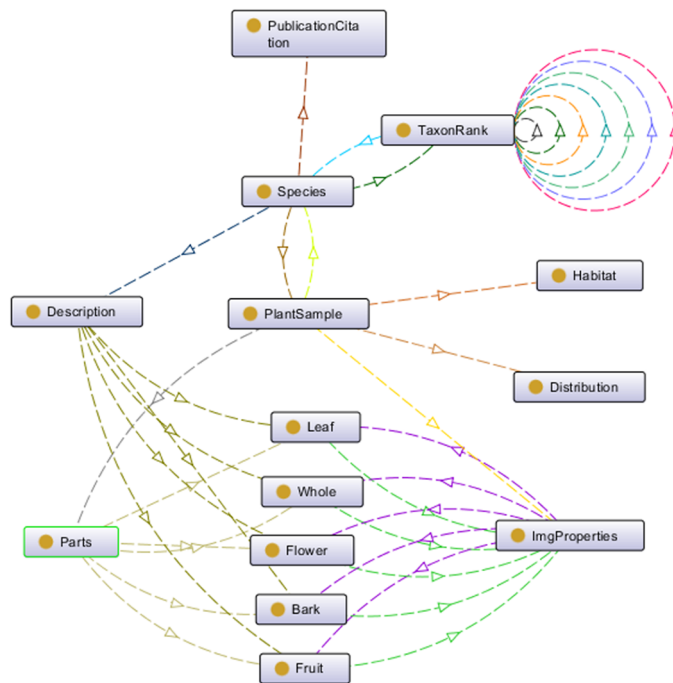


Figure 5 Top-level entities in POUM ontology. A partial of POUM using OntoGraph plug-in in Protégé 5.2.

Full-size DOI: 10.7717/peerj.5579/fig-5

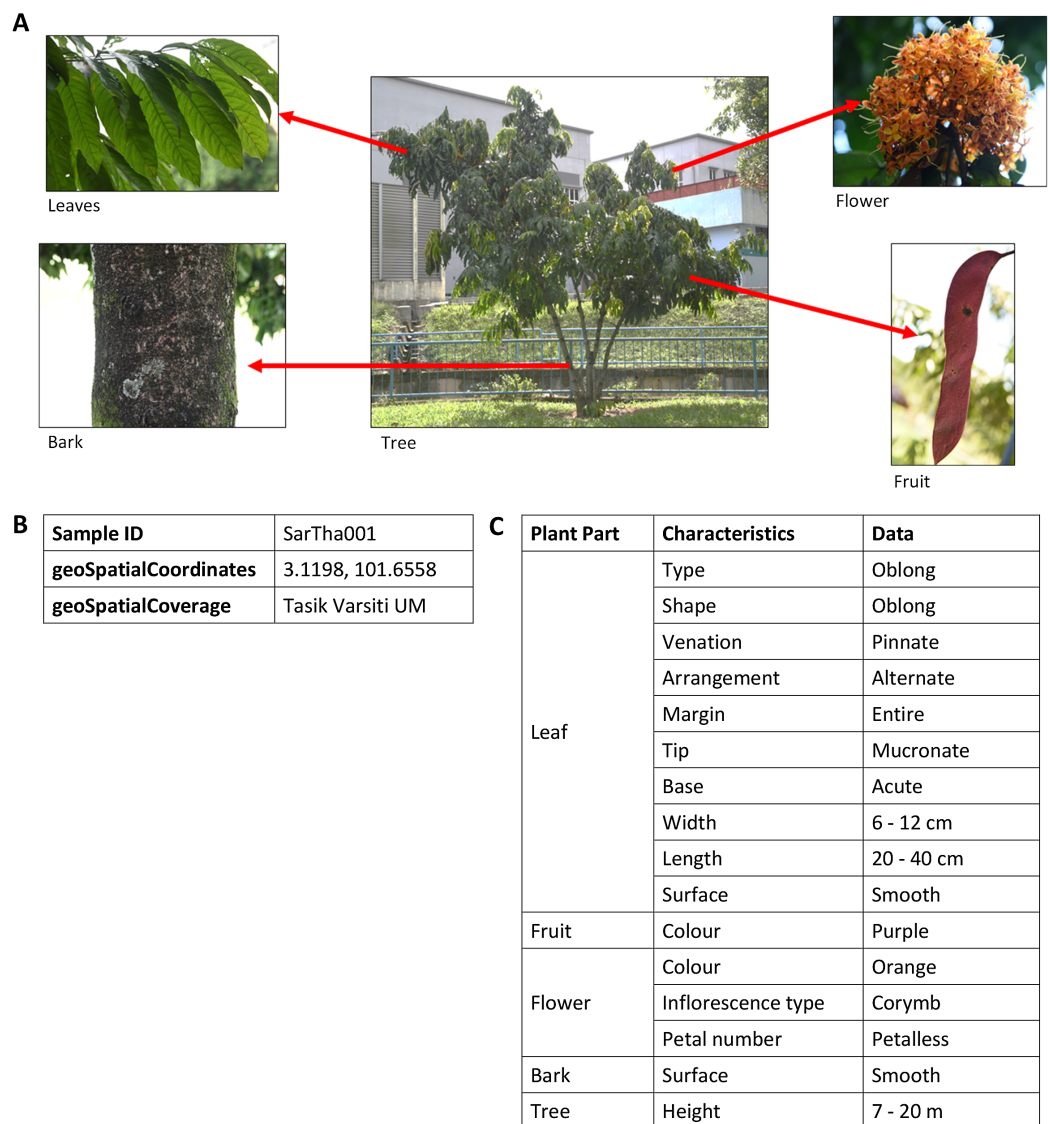


Figure 6 An example of images and description of *Saraca thaipingensis* sample. (A) Digital images representing the *S. thaipingensis* sample by a whole tree and its parts, namely leaves, bark, flower and fruit. (B) Unique ID and location data of *S. thaipingensis* sample. (C) Description of characteristics of *S. thaipingensis* sample. Photo credit: the photo archive at UM Plant Knowledge (<http://103.18.1.10:8080/plantviz/>). Full-size DOI: 10.7717/peerj.5579/fig-6

query is sent to the server, where querying is executed to the RDF graph data model of the POUM. The results of the query are in the form of XML syntax and then further structured into a JSON format text to be used by the D3.js library.

Step 3: Transformation of text-based results in XML serialisation format into graphical form

The relationship between data is crucial for highlighting data association in the knowledge domain. Three types of relationships are taken into consideration:

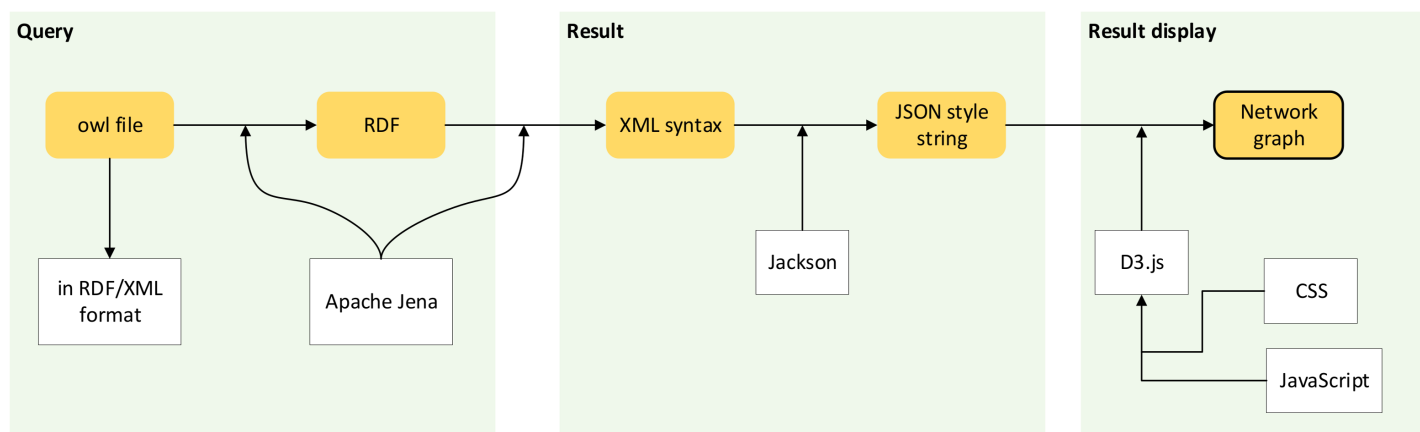


Figure 7 Workflow of network graph development. Processes involve passing ontological data to the creation of a network graph.

Full-size DOI: [10.7717/peerj.5579/fig-7](https://doi.org/10.7717/peerj.5579/fig-7)

1. The relationship between one taxon to another taxon. A taxon is linked to another by its family name. For example, *Delonix regia* and *Acacia auriculiformis* are linked to one another as both are in the same family, Fabaceae.
2. The relationship between a taxa and its sample. Each taxon has three samples and each sample has a unique identifier. For instance, *Murraya paniculata* has three samples, namely 'SMurPan001', 'SMurPan002' and 'SMurPan003'.
3. The relationship between samples. Samples are related when they are from the same taxon or are obtained from the same location. For instance, samples of *Lagerstroemia indica*, *Manihot esculenta* and *Terminalia catappa* are collected from the same location, which is 'DTC UM'.

These relationships can be clearly shown when illustrated to users in a visual form. In order to transform texts into a graphical form, the D3.js visualisation library was chosen, as this utilises different types of programming languages in the data visualisation design. The library aims to shorten the development process while maintaining the quality of the system. There are many types of visualisations that can be designed using D3.js; thus, it is important to choose a suitable type of visualisation.

For the type of visualisation in this study, a network graph was chosen. This graph enables the visualisation of plant data, as the relationship between one data to another can be illustrated clearly. The network graph is a type of graph that highlights the relationship between entities and consists of 'nodes' as entities and 'links' as lines to link between entities. There are two types of nodes, which are the 'parent' (PN) and 'children' (CN) nodes. PN is a type of node that has one or more CN. Meanwhile, CN is a type of node that has a PN ancestor.

Before generating the graphic representation, a canvas was first set up. The D3.js library uses a layout controlled by force and SVG as the container for the visualisation. A few parameters were defined such as distance, gravity and size from the function `d3.layout.force()`. The result in JSON format is called a variable, as it was further arranged using D3.js to create a network graph. Besides that, some interactive features such as view

node label, highlight node's links, expand or shrink group of nodes, page of sample information and thumbnail images of plant sample were added into the network graph in order to develop a user-friendly system. This is to enable the interaction between user and data. Hence, users were able to manipulate the content of the network graph with the implemented features.

Step 4: Display result to the user in a graphical user interface

The generated network graph was then displayed in the GUI to provide the full experience to users, as this was embedded in the interface.

Data and source code for the plant visualisation system (PlantViz) development can be accessed and retrieved from <https://github.com/afrinaad/PlantViz/>.


User testing

There are a variety of data visualisation tools available to help developers achieve the objective of data visualisation. Therefore, it is desirable to determine whether or not the developed visualisation is successful in achieving user needs. In order to measure the performance of the developed plant data visualisation, a user evaluation was carried out. User evaluation consists of two tests, namely: (1) usability heuristics and (2) query and visualisation evaluation. User evaluation involved expert users with research background or experience in the botanical field including users that might have little skill in information technology (IT). Meanwhile, novice users who have little or no research background or experience in the botanical field might have skills in IT.


The usability heuristic evaluation was adapted from Nielsen's 10 usability heuristics for user interface (UI) design (Nielsen, 1992). These are the general principles for an interactive UI design. In this study, 10 usability features were adapted to match the developed visualisation system. A sample of the questionnaire used in this study is shown in Fig. 8. Users were given five minutes to explore the GUI before the evaluation. This step was performed to observe the users' first impression of the visualisation system.

A query and visualisation evaluation was conducted to assess the efficiency of the visualisation system in delivering visualised content to users (Amri, Ltifi & Ayed, 2015; Hearst, Laskowski & Silva, 2016). This helps measure the performance of query sending from the GUI to the server and translates the results into a visual format. In addition, this evaluation assists in the observation of the ability of users to interact with data successfully using features of the visualisation system. Figure 9 shows a sample of the questionnaire for this purpose. Users were given guidelines and instructions for using the developed data visualisation system before performing user evaluation. Both the guidelines and instructions were given to four different cases based on specific search parameters, namely *scientific name* (Case 1), *family name* (Case 2), *location* (Case 3) and *water usage* (Case 4). Query and visualisation were evaluated via a rating of '1' to '5', where '1'-Poor, '2'-Fair, '3'-Average, '4'-Good and '5'-Excellent. Comments from the users was taken into consideration in order to improve the developed visualisation system.

	Usability Features	Yes	No
E1	Status of the page is displayed (e.g. 'updates', 'news')		
E2	Easy to understand the contents written in the website		
E3	No irrelevant information in the website		
E4	Query page is user-friendly		
E5	Query page is consistent to use		
E6	Guidelines are given to prevent any errors in querying		
E7	Querying is easy to use even though without reading the manual page		
E8	Instructions are visible when using the visualization system		
E9	The error message is easy to understand		
E10	'Help' guideline on the website/visualization system is adequate		

Figure 8 Sample of questionnaire for usability heuristics evaluation. A total of 10 features were adapted from Nielsen's 10 usability heuristics for UI. [Full-size](#)  DOI: 10.7717/peerj.5579/fig-8

Criteria		Rating score				
Query evaluation		1	2	3	4	5
Q1	Access to a query page					
Q2	Perform a query					
Q3	Guidance on how to perform a query is adequate					
Q4	Number of parameter that can be queried at a time					
Q5	Time needed to complete a query is less than 1 minute					
Visualization evaluation		1	2	3	4	5
V1	Quality of visualization based on the query result					
V2	Speediness of the system to load the network graph					
V3	Visualization of the retrieved information is easy to understand					
V4	Able to infer knowledge from the relationship of information: •Sample to sample •Taxon to taxon •Sample to taxon					
V5	Relevant images are shown when hovering the cursor on specific nodes					
V6	Information shown in visualization is sufficient					
V7	Access to get a complete information of plant sample					
V8	Error message is shown if there is no retrieved result					

Figure 9 Sample of questionnaire for query and visualisation evaluation. The questionnaire consists of five criteria of query evaluation and eight criteria of visualisation evaluation for the visualisation efficiency assessment. [Full-size](#)  DOI: 10.7717/peerj.5579/fig-9

Next, Fisher's Exact test and *t*-test were performed to analyse the outcome of the evaluation. Based on the usability heuristic evaluation, Fisher's Exact was conducted to check whether or not the GUI of the system is dependent on the users' knowledge in botanic and IT. Meanwhile, based on the query and visualisation evaluation, a *t*-test

was conducted to check whether or not there was any significant difference between the evaluation done by expert and novice users on the visualisation system.

RESULTS

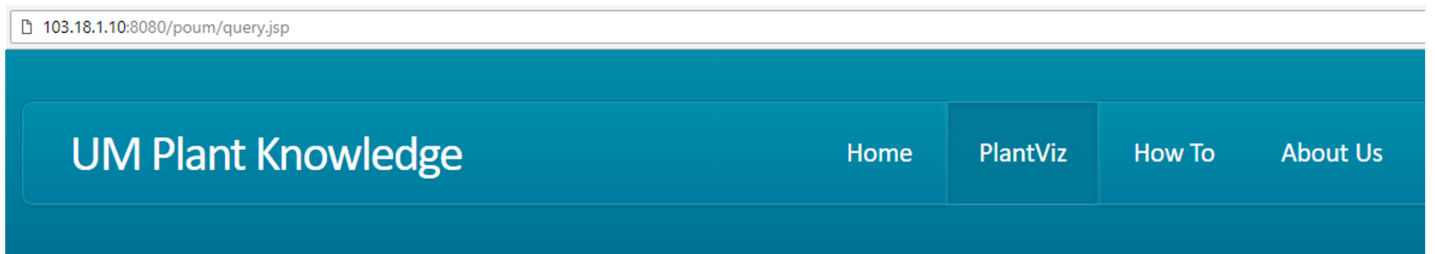
Plant visualisation system

Based on the proposed visual-based representation for plant species data, a prototype of the web-based plant data visualisation system was developed called PlantViz, as shown in Fig. 10. PlantViz consists of a query tool to search for data in POUM and a graphical viewer to display the retrieved results to users. There are four parameters in the Query page, which are based on *scientific name*, *family name*, *location* and *water usage*.

The *scientific name* and *family name* parameters are commonly used as search parameters in many public databases (NRCS, 2017; PFAF, 2017; UCONN, 2017). Besides that, the *location* parameter is for the location where the plant sample was collected. This was chosen as one of the query parameters because the plant samples in this study were mainly collected from various areas in UM. In addition, PlantViz's target users are members of the university, who are familiar with these locations such as 'DTC UM', 'Tasik Varsiti UM' and 'Fakulti Sains UM', all locations in UM. The *water usage* parameter defines the level of water needed by a plant species for its growth. Since parameters such as *scientific name* and *family name* were used to represent the taxonomical aspect and *location* was used to represent the geographical information of the plant. The *water usage* parameter was chosen to represent the morphology attribute of the plant.

The proposed methodology for transforming text-based results into visualisation form was implemented in PlantViz. Figure 11 shows a fragment of the ontological data serialised into an RDF graph data model that contains a collection of RDF nodes attached to each other by annotated relations. This process was done using a *ModelFactory* class from the Apache Jena library.

The RDF graph data model was then queried from when the users used the query tool in PlantViz. Textual query from users was sent as a SPARQL query using *QueryFactory* and *QueryExecutionFactory* classes from Apache Jena. The result of the query was in XML format, which was then restructured into JSON format. The flexibility of JSON enables users to organise results according to any arrangement, as JSON can contain any number, Unicode character, Boolean operators, array, object or null value (Bassett, 2015). Figure 12 shows an example of the query results in both XML and JSON formats. The example shown is a shortened part of the result to show the difference in terms of structure for the query result before and after being transformed into JSON. Figure 12A illustrates the results generated by Apache Jena, which is in XML format. This was then rearranged into JSON format and later organised into different arrays based on the relationship of the data, as shown in Fig. 12B. For example, general information on the plant such as its common name, water usage and soil type were organised into the same array. Results of the query in JSON format were then passed to the D3.js library to form a network graph as a visualisation of the query result. For instance, the network graph shown in Fig. 13 is the result of querying the *Magnolia figo* plant species.



QUERY

In PlantViz's query tool, users are able to search on any of given parameters available in the drop-down menu. Choose any parameter, key in a search keyword or choose from the options given, and click Search to perform the query.

Search for : [Reset](#)

- Choose option
- Choose option
- Scientific Name
- Family Name
- Location
- Water Usage

Figure 10 GUI of PlantViz. Main interface of PlantViz that consists of a query tool and a graphical viewer.

Full-size DOI: 10.7717/peerj.5579/fig-10

```
[http://103.18.1.10:8080/plantdb/ontology/plantont#MagFig_Bark, http://103.18.1.10:8080/plantdb/ontology/plantont#BarkColor, "Brown"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#MagFig_Bark, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://103.18.1.10:8080/plantdb/ontology/plantont#Description]
[http://103.18.1.10:8080/plantdb/ontology/plantont#MagFig_Bark, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
[http://103.18.1.10:8080/plantdb/ontology/plantont#TriWhi_Fruit, http://103.18.1.10:8080/plantdb/ontology/plantont#fruitColor, "Brown, Green"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#TriWhi_Fruit, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://103.18.1.10:8080/plantdb/ontology/plantont#Description]
[http://103.18.1.10:8080/plantdb/ontology/plantont#TriWhi_Fruit, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DipGra_Bark, http://103.18.1.10:8080/plantdb/ontology/plantont#BarkSurface, "Fissured"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DipGra_Bark, http://103.18.1.10:8080/plantdb/ontology/plantont#BarkColor, "Light-grey"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DipGra_Bark, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://103.18.1.10:8080/plantdb/ontology/plantont#Description]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DipGra_Bark, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HopOdo_Leaf, http://103.18.1.10:8080/plantdb/ontology/plantont#leafBase, "Rounded"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HopOdo_Leaf, http://103.18.1.10:8080/plantdb/ontology/plantont#leafType, "Simple"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HopOdo_Leaf, http://103.18.1.10:8080/plantdb/ontology/plantont#leafMargin, "Entire"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HopOdo_Leaf, http://103.18.1.10:8080/plantdb/ontology/plantont#leafArrangement, "Alternate"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HYMCOUO-B001, http://103.18.1.10:8080/plantdb/ontology/plantont#imgSize, "1.2 MB"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HYMCOUO-B001, http://103.18.1.10:8080/plantdb/ontology/plantont#imgPath, "/HYMCOUO-B001.jpg"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HYMCOUO-B001, http://103.18.1.10:8080/plantdb/ontology/plantont#imgObject, "Bark"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HYMCOUO-B001, http://103.18.1.10:8080/plantdb/ontology/plantont#fileFormat, ".jpg"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HYMCOUO-B001, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://103.18.1.10:8080/plantdb/ontology/plantont#ImgProperties]
[http://103.18.1.10:8080/plantdb/ontology/plantont#HYMCOUO-B001, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DurEre_Bark, http://103.18.1.10:8080/plantdb/ontology/plantont#BarkSurface, "No information"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DurEre_Bark, http://103.18.1.10:8080/plantdb/ontology/plantont#BarkColor, "Brown"^^http://www.w3.org/2001/XMLSchema#string]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DurEre_Bark, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://103.18.1.10:8080/plantdb/ontology/plantont#Description]
[http://103.18.1.10:8080/plantdb/ontology/plantont#DurEre_Bark, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
[http://103.18.1.10:8080/plantdb/ontology/plantont#SAllCat001, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
[http://103.18.1.10:8080/plantdb/ontology/plantont#SAllCat001, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://103.18.1.10:8080/plantdb/ontology/plantont#PlantSample]
[http://103.18.1.10:8080/plantdb/ontology/plantont#SStrCri001_Whole, http://www.w3.org/1999/02/22-rdf-syntax-ns#type, http://www.w3.org/2002/07/owl#NamedIndividual]
```

Figure 11 Ontological data in the form of a RDF graph data model. A fragment of the ontological data in RDF serialisation format.

Full-size DOI: 10.7717/peerj.5579/fig-11

Three types of data relationships, as mentioned earlier, are clearly shown from the results of each chosen query parameter and depicted in Fig. 14. Moreover, unlike plant ontology (PO; *Jaiswal et al., 2005*), which provides a static graphical view, the graphical viewer in PlantViz provides interactive elements, which allow users to explore the result (*Lohmann et al., 2015*), as illustrated in Fig. 15. Table 1 lists the description of visualisation features, functions and its conditions.

PlantViz is freely accessible at <http://103.18.1.10:8080/plantviz/>. The detailed manual on how to use the PlantViz is provided as [Supplemental Article S1](#).

A

```
<?xml version="1.0"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">
  <head>
    <variable name="id"/>
    <variable name="plantname"/>
    <variable name="scname"/>
    ...
  </head>
  <results>
    <result>
      <binding name="id">
        <literal datatype="http://www.w3.org/2001/XMLSchema#string">
SMagFig001
</literal>
      </binding>
      <binding name="plantname">
        <literal datatype="http://www.w3.org/2001/XMLSchema#string">
Cempaka telur
</literal>
      </binding>
      <binding name="scname">
        <literal datatype="http://www.w3.org/2001/XMLSchema#string">
Magnolia figo
</literal>
      </binding>
      ...
    </result>
  </results>
</sparql>
```

B

```
{
  "id": "1",
  "name": "Magnolia figo",
  "con": "Scientific Name",
  "size": 2000,
  "children": [
    {
      "id": "2",
      "name": "General Information",
      "con": "General Information",
      "size": 2000,
      "children": [
        {
          "id": "3",
          "name": "Cempaka telur",
          "con": "Common Name",
          "size": 2000
        },
        {
          "id": "4",
          "name": "Fertile loamy, well-drained soils",
          "con": "Soil",
          "size": 2000
        },
        {
          "id": "5",
          "name": "Moderate",
          "con": "Water Usage",
          "size": 2000
        }
      ]
    }
  ]
},
...
]
```

Array

Figure 12 An example of query result in XML and JSON formats. (A) The results generated by Jena is in XML format. (B) The results in JSON format that show the relationship between the data. [Full-size !\[\]\(5f471a71b78d7676bc356df190b88ab4_img.jpg\) DOI: 10.7717/peerj.5579/fig-12](https://doi.org/10.7717/peerj.5579/fig-12)

RESULT

Result for Scientific Name : *Magnolia figo*

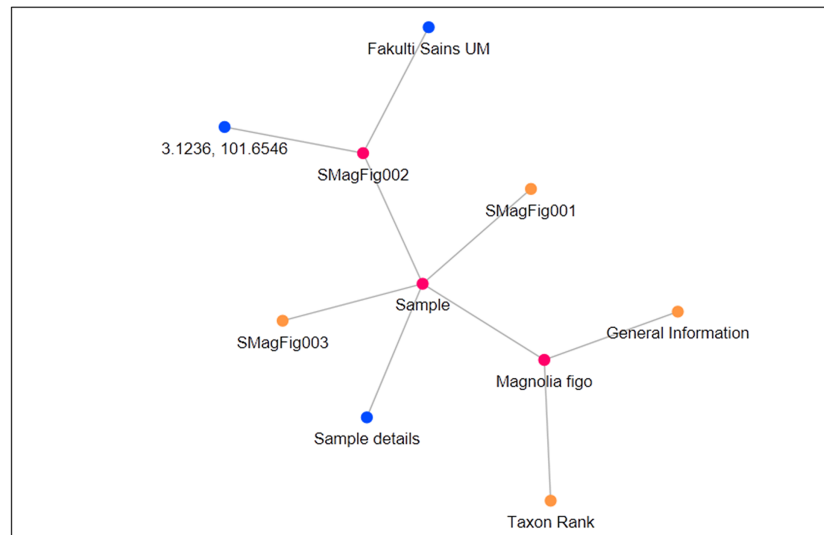
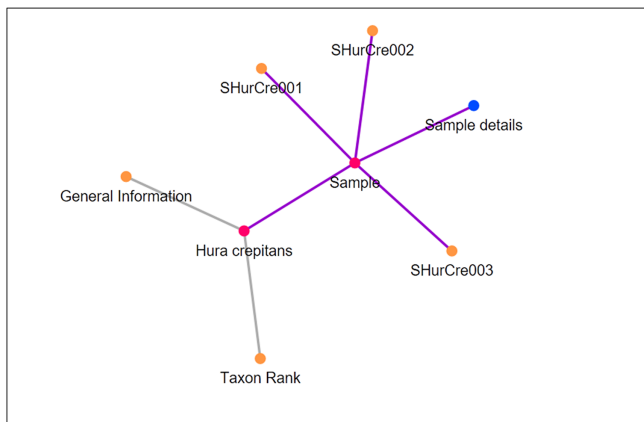
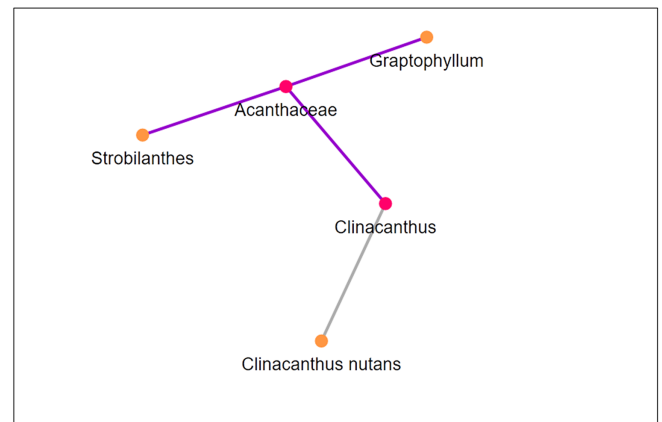


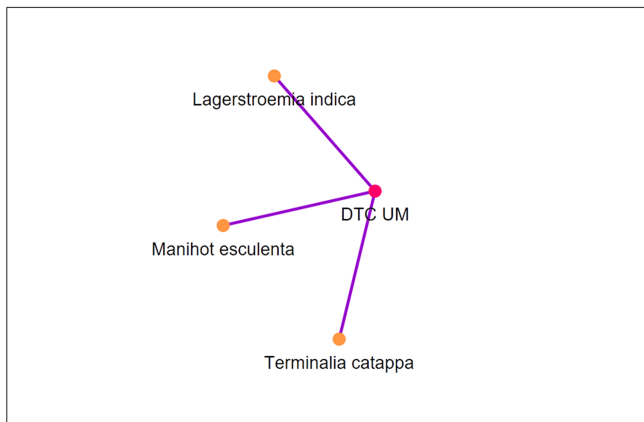
Figure 13 PlantViz data visualisation. An example of visualisation generated in PlantViz using the scientific name parameter, *Magnolia figo*. [Full-size !\[\]\(e18234c3c79d4e9aedd7f39240ab238d_img.jpg\) DOI: 10.7717/peerj.5579/fig-13](https://doi.org/10.7717/peerj.5579/fig-13)

A RESULTResult for Scientific Name : *Hura crepitans***B** RESULT

Result for Family Name : Acanthaceae

**C** RESULT

Result for Location : DTC UM

**D** RESULT

Result for Water Usage : Low to moderate

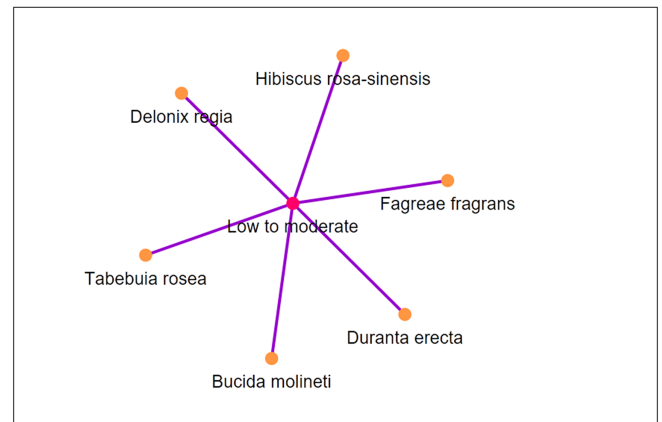


Figure 14 Examples of data visualisation for the four query parameters. (A) Query parameter *scientific name*. (B) Query parameter *family name*. (C) Query parameter *location*. (D) Query parameter *water usage*. [Full-size !\[\]\(e3ea84e00090d181251f63e2c7b0c977_img.jpg\) DOI: 10.7717/peerj.5579/fig-14](https://doi.org/10.7717/peerj.5579/fig-14)

User evaluation

User evaluation was performed on the querying and visualisation of PlantViz. Sixty users including 30 expert users and 30 novice users participated in this user evaluation. The expert users are botanists and researchers in biodiversity with little IT knowledge while the novice users are undergraduate students from UM.

Figure 16 presents an analysis of the usability heuristics evaluation by both expert and novice users. This shows that the majority of users rated ‘Yes’ for most of the features. As shown in Fig. 16A, E6 has the highest number for the rating ‘Yes’ (all 30 expert users voted ‘Yes’), while for novice users, as shown in Fig. 16B, E2, E6 and E9 have the highest number for the rating ‘Yes’ (all 30 novice users voted ‘Yes’). Meanwhile, for both types of users, E1 has the highest number for rating ‘No’ (16 out of 30 expert users and 14 out of 30 novice users voted ‘No’). This is consistent with prototype development. The guidelines for using the system are available with no status for the system being shown.

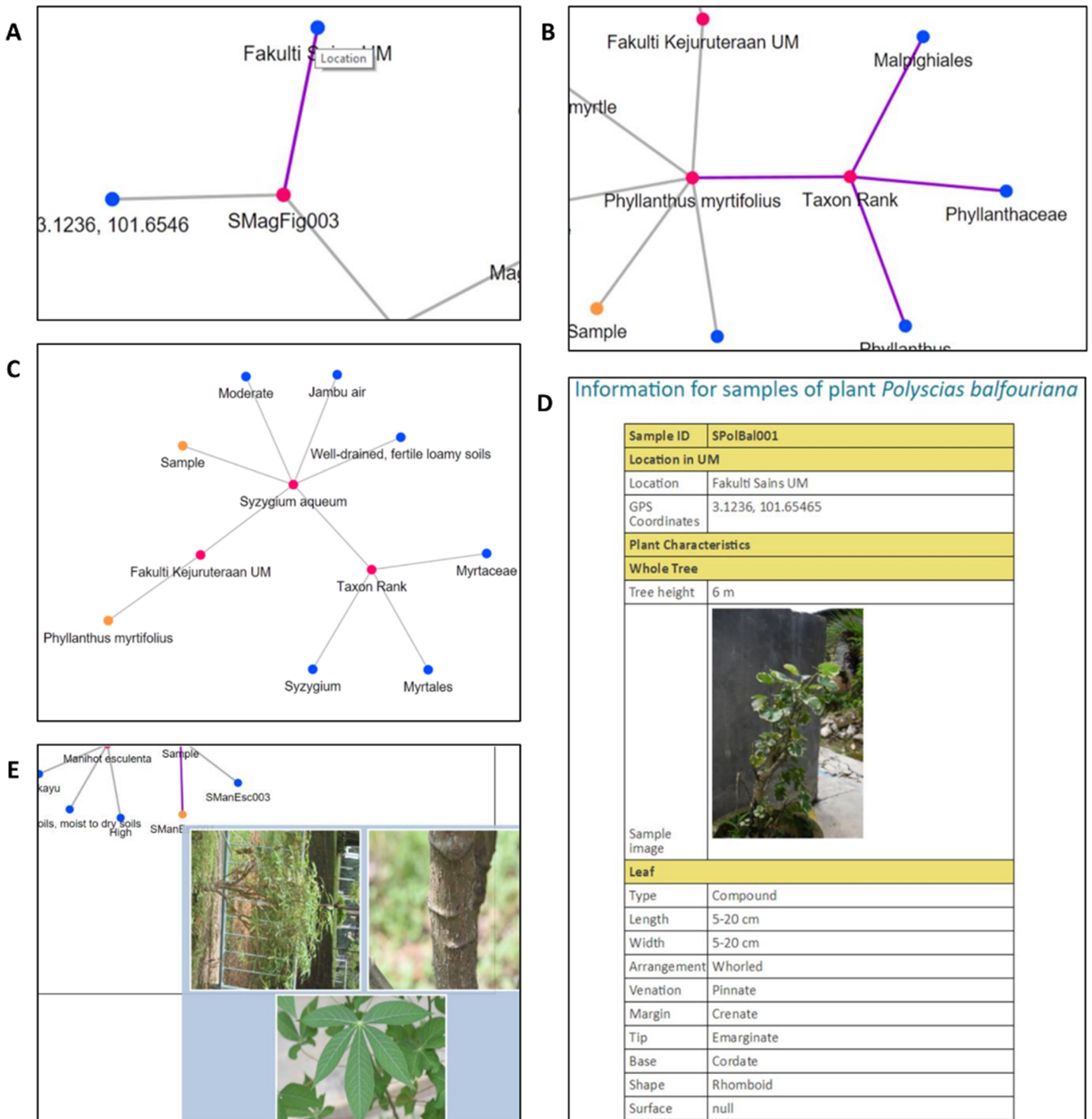


Figure 15 Interactive elements in PlantViz's graphical viewer. (A) View node label, (B) highlight node links, (C) expand or shrink group of nodes, (D) sample information page and (E) thumbnail images of plant sample. Photo credit: the photo archive at UM Plant Knowledge (<http://103.18.1.10:8080/plantviz/>).

Full-size DOI: 10.7717/peerj.5579/fig-15

Table 1 List of visualisation features in PlantViz.

Visualisation features	Functions	Conditions
View node label (Fig. 15A)	To show the label of a node	<ul style="list-style-type: none"> Visible when the cursor is hovered
Highlight node's links (Fig. 15B)	To show relationship between nodes	<ul style="list-style-type: none"> Visible when the cursor is hovered Only highlights links connected to the node
Expand or shrink group of nodes (Fig. 15C)	To allow users to expand or shrink a group of nodes	<ul style="list-style-type: none"> When users click on node of 'parent' type Only node of 'parent' type can be expanded or shrunk
Sample information page (Fig. 15D)	To redirect users to a new page containing information of plant samples	<ul style="list-style-type: none"> When node with text 'More detail' is clicked
Thumbnail images of plant samples (Fig. 15E)	To show thumbnail-sized images of plant samples	<ul style="list-style-type: none"> Visible when the cursor is hovered on nodes with label 'Sample ID'

Note:

The features' description, function and its conditions.

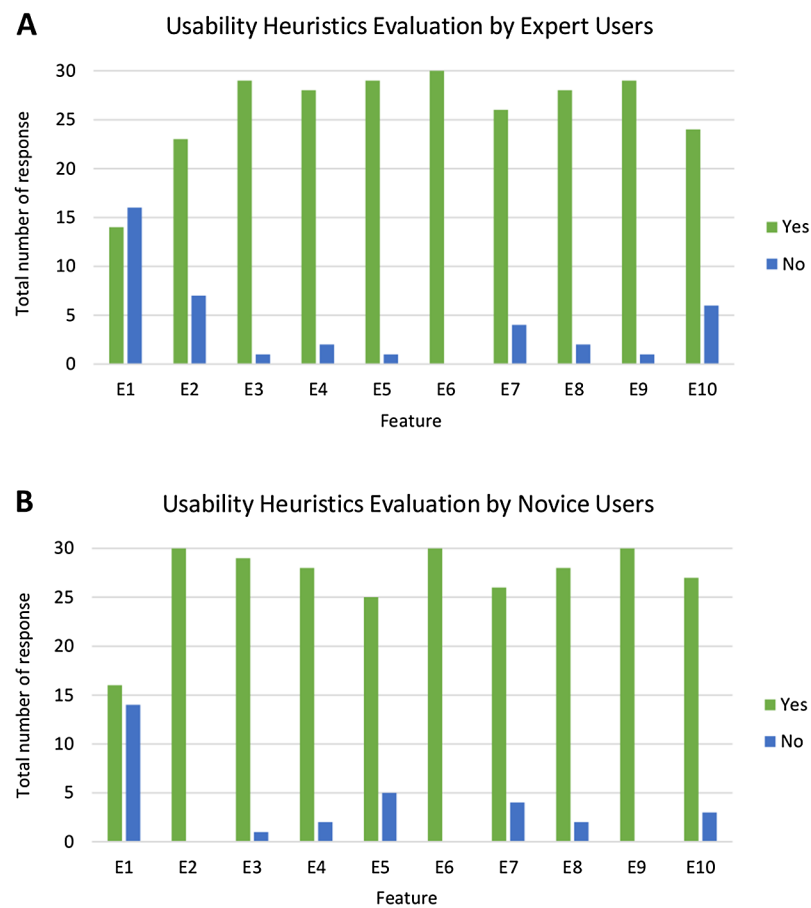


Figure 16 Analysis of usability heuristics evaluation by expert and novice users. The bar chart shows the total number of responses that rated 'Yes' and 'No' for each feature in the usability heuristics evaluation, as per the (A) expert users and (B) novice users.

Full-size  DOI: [10.7717/peerj.5579/fig-16](https://doi.org/10.7717/peerj.5579/fig-16)

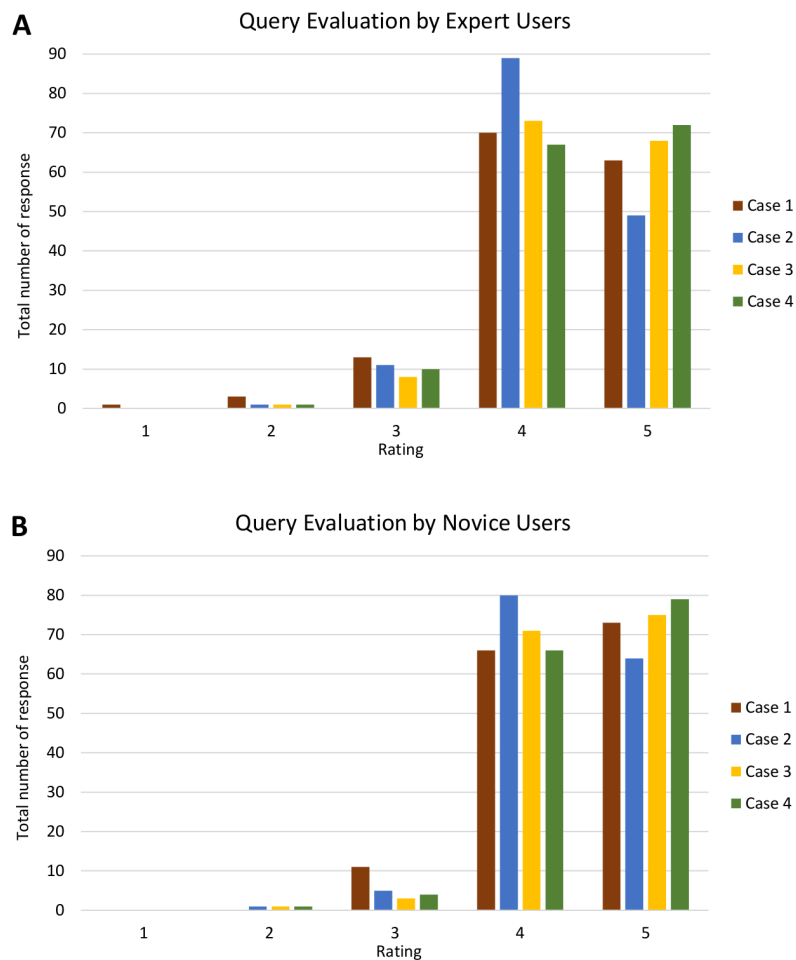


Figure 17 Analysis of query evaluation by expert and novice users. The bar chart shows the total number of responses for each rating of all four cases by (A) expert users and (B) novice users. [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02_img.jpg\) DOI: 10.7717/peerj.5579/fig-17](https://doi.org/10.7717/peerj.5579/fig-17)

Fisher's Exact test was conducted to check whether or not the user experience in using PlantViz is dependent on their expertise level. The null hypothesis H_0 is that there is no difference between the usability heuristics evaluation performed between expert users and novice users. The two-tailed probability (p) value of Fisher's Exact test on usability heuristics evaluation is 0.312, ($p < 0.05$), which means that there is no significant difference in the usability heuristic evaluation between expert and novice users. This also indicates that PlantViz's UI is adequate for all types of users regardless of their IT knowledge.

Analyses of query and visualisation evaluation for all four cases by both expert and novice users are shown in [Figs. S1](#) and [S2](#), respectively. [Figure 17](#) shows the analysis of query evaluation by expert and novice users. The total number of responses for each case is plotted against the evaluation rating. The query evaluation by both types of users is shown in [Figs. 17A](#) and [17B](#). These show similar results whereby the majority of expert and novice users gave a '4' or '5' rating for most of the query criteria. There is one

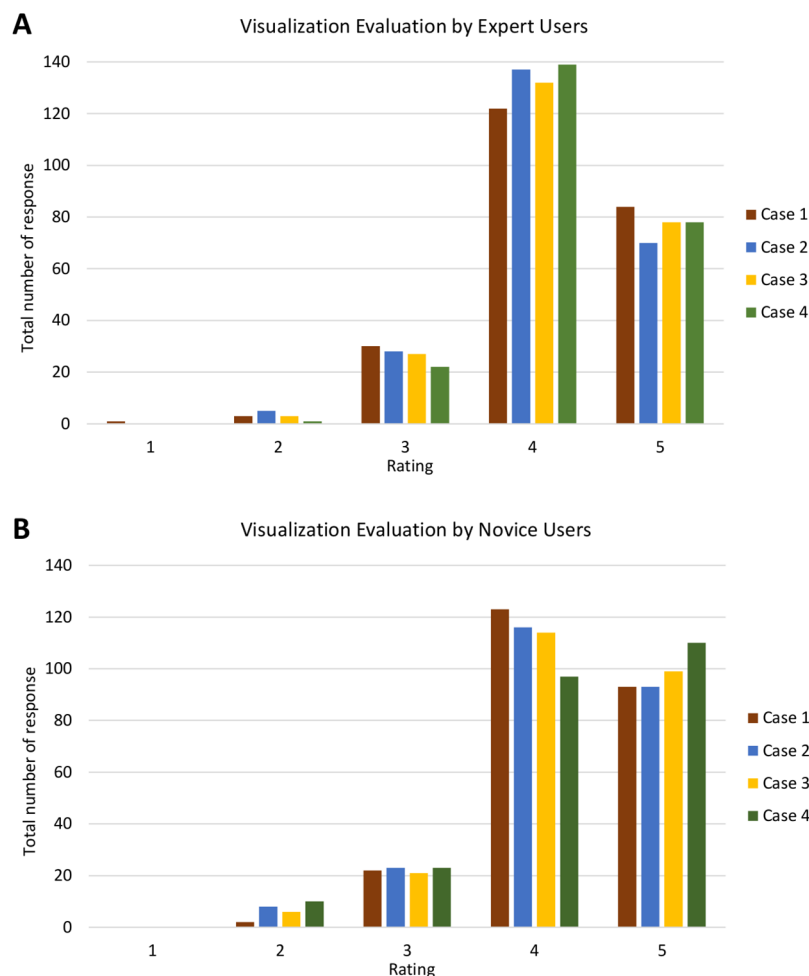


Figure 18 Analysis of visualisation evaluation by expert and novice users. The bar chart shows the total number of responses for each rating for all four cases by (A) expert users and (B) novice users.

Full-size  DOI: [10.7717/peerj.5579/fig-18](https://doi.org/10.7717/peerj.5579/fig-18)

response that gave a rating of ‘1’ (one expert user) and a total of nine responses that gave a rating of ‘2’ (six expert and three novice users). A low rating was given for Q4, as shown in Fig. S1.

Furthermore, Fig. 18 shows the analysis of visualisation evaluation by expert and novice users. The total number of responses for each case is plotted against the evaluation rating as well. Figures 18A and 18B show that both types of users rated ‘4’ for most of the visualisation criteria. There is one response with a rating of ‘1’ and 35 responses with a rating of ‘2’ (see Fig. S2). V7 was given a rating of ‘1’ (one expert user rated ‘1’) and had the highest number for a rating of ‘2’ (four experts and nine novice users rated ‘2’). It can therefore be concluded that the PlantViz interactive feature for the sample information page should be improved and be more easily accessible. Besides that, V6 had seven responses with a rating of ‘2’ (two experts and five novice users rated ‘2’), which shows that users did not agree that the information shown in the graphical viewer was sufficient. More information could be added to the network graph such as type of life cycle,

Table 2 *t*-test results.

Cases	<i>p</i> -Value	
	Query evaluation	Visualisation evaluation
Case 1	0.082	0.133
Case 2	0.105	0.165
Case 3	0.177	0.172
Case 4	0.225	0.409

Note:

Results of independent sample *t*-test for query and visualisation evaluation.

propagation method and characteristics of other parts of plants. Linking the data to other external databases such as the accession term in Trait Ontology ([Walls et al., 2012](#)) and PO would extend the information scope to users.

Independent sample *t*-tests were conducted to test the significance of the evaluation done by both types of users. The null hypothesis H_0 declares that there is no difference between the evaluation of expert and novice users. *T*-tests were performed on each case for both evaluations with results as shown in [Table 2](#). The *p*-values for all cases are higher than the significance level, thus there is a lack of evidence to reject H_0 . Hence, it can be concluded that there is no statistically significant difference between the query and visualisation evaluation performed by expert and novice users. This signifies that PlantViz can be used by users of all levels, with or without knowledge in the botanical or IT fields.

DISCUSSION

Many plant species share resemblance in appearance, yet there are many characteristics such as genomic data, morphological attributes and geographical attributes that can distinguish one plant species from the other. More often than not, the data or information in textual form found on many online databases are presented using lengthy and wordy descriptions. Consequently, this hinders researchers from deducing new knowledge from the presented plant data. In this paper, we report an alternative approach for presenting plant data, which is via network graph. Thus, by visualising the retrieved results from the database, the problem of lengthy retrieved texts as part of the results can be eliminated. In addition, this method emphasises the relationship between data and presents this information in a visualised form for users. Hence, from the visualised data, the relationship between data can be easily analysed and inferred.

As mentioned earlier, PlantViz consists of a query tool and a graphical viewer. Users can use PlantViz via a user-friendly GUI. Thus, two types of tests were conducted to assess the performance efficiency of PlantViz, namely usability heuristics evaluation and query and visualisation evaluations.

From the observation, the *p*-values of the query and visualisation evaluation ([Table 2](#)) share the same pattern, with Case 1 having the lowest *p*-value. It can be presumed that in Case 1, users were not yet familiar with the graphical viewer of the PlantViz system. This is because the graphical viewer is not a common tool in many public plant-based databases such as The Plant List ([List, 2013](#)) and NParks Flora & Fauna

(<https://florafaunaweb.nparks.gov.sg/Home.aspx>). However, all other cases, namely Case 2, Case 3 and Case 4, have p -values higher than Case 1 for both evaluations; hence, the previous assumption is valid as users had just started to become acquainted with PlantViz. In addition, the GUI design of both the query tool and graphical viewer in PlantViz is simple yet still appropriate for both types of users. As for Case 4, both evaluations had the highest p -value. Therefore, it can be concluded that users were easily accustomed to the PlantViz system. This also verifies that the GUI for PlantViz is consistent throughout all cases, whereby each case uses different search parameters.

Yet, filtering feature could be added to the graphical viewer to grant users the ability to filter search results. For instance, the *location* parameter was used as the search parameter for Case 3 and this could generate a cluttered network graph, as many plant species were collected from the same location. Thus, as mentioned by *Cline et al. (2007)*, users could filter the results by selecting only a certain type as shown in the graphical viewer, and this could help in retrieving the most relevant one. In addition, users should be allowed to filter other types of search parameters such as family name, number of plant sample collected or type of plant.

The query method described here is simple—a typical Boolean search is used. Thus, a more robust and efficient query could be achieved by implementing other query and search methods such as the ranking algorithm (*Tran, Tsuji & Masuda, 2009; Zhiguo & Zhengjie, 2010*) and natural language query processing (*Damljanovic, Agatonovic & Cunningham, 2010; Paredes-Valverde et al., 2015; Tablan, Damljanovic & Bontcheva, 2008; Varga et al., 2014; Wang et al., 2007*). It is vital for users to query the system without being attached to a fixed one and to select pre-defined parameters. Based on the low-rated Q4, as shown in *Fig. S1*, users were dissatisfied with the number of search parameters that can be used to query at a time. This limitation will be enhanced in the future by allowing users to add more than one search parameter to narrow down the retrieved results, so that more relevant results could be retrieved.

The POUM currently consists of plant descriptions and their samples in textual and image forms. PlantViz was able to visualise the POUM dataset by showing the relationships between: (1) 'taxon-taxon'; (2) 'taxon-sample' and (3) 'sample-sample'. This dataset will be extended to include other plant types along with their descriptions and plant systematics, behaviour, ecology and diversity. Thus, POUM could be linked (*Hebeler et al., 2009; Smith et al., 2007*) as well to other existing ontologies (*Tello-Ruiz et al., 2016*). For this reason, PlantViz will be further enhanced to support additional features such as: (1) querying from different datasets at a time and show their relationships and (2) providing data analytics by inferring the visualised relationships.

Furthermore, the visualisation feature will also be improved in order to enhance the usability of the proposed visual-based representation for plant data. Other features such as enabling users to choose the types of visualisation to be generated and choosing a colour scheme to differentiate the relationships between data could be considered as well.

In addition, it is possible to apply the visual-based representation for large dataset such as PlantCLEF (*Goeau, Bonnet & Joly, 2017*) and National Phenology Database

(*Marsh, 2017*) as these two examples have the database backend and provide the visualisation tool. Definitely, the data will be retrieved; however, the visualised data will be cluttered and interactivity might be inefficient. Hence, an extra task pre-processing the retrieved data such as through dimensionality reduction and data pre-fetch may be needed before the data could be visualised. This pre-processing would retrieve a lesser amount of data, whereby dimensionality reduction helps in reducing the computational load (*Kaski & Peltonen, 2011*), while data pre-fetch helps in improving the response time (*Battle, Chang & Stonebraker, 2016*). Besides that, this method could be improved by employing the hierarchical exploration method which is commonly used to visualise large and high-dimensional data as proposed by *Lin et al. (2013)*, *Yang, Ward & Rundensteiner (2003)* and *Zinsmaier et al. (2012)*. This method is practical because it allows data visualisation in different levels of details (*Bikakis, 2018*).

CONCLUSION

The field of biology generates thousands of data every day with the advancements in modern tools and technologies. Hence, it is important to implement a proper methodology for retrieving data and making it accessible to users in an effective way. This study focuses on presenting retrieved data from an ontology to users in a visualisation form. Hence, a visual representation of plant data, PlantViz, was proposed. Data from POUM database was converted as a graph data model and queried using SPARQL. Results from the query were then structured into JSON format before being transformed into a visualisation form and then presented to users in GUI form. User evaluation analysis and results show that PlantViz can be used by users of different levels, either expert users from the botanical field, students or laymen with an interest in botany. PlantViz eliminates lengthy texts of information as results for user queries. In addition, users were able to interact with the data directly, as the visualisation is dynamic. Most importantly, this technique demonstrates the practicability of using computer-assisted tools by providing cognitive analysis to understand the relationship between data. Moreover, this study also facilitates users in inferring and gaining new data insights.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This project was supported by the University of Malaya under UMRG grant (RP038B-15AET). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
University of Malaya under UMRG grant: RP038B-15AET.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Afrina Adlyna Mohamad-Matrol conceived and designed the experiments, performed the experiments, analysed the data, prepared figures and/or tables, authored or reviewed drafts of the paper.
- Siow-Wee Chang conceived and designed the experiments, analysed the data, authored or reviewed drafts of the paper.
- Arpah Abu conceived and designed the experiments, performed the experiments, analysed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:

GitHub: <https://github.com/afrinaad/PlantViz/>

PlantViz is freely accessible at <http://103.18.1.10:8080/plantviz/>.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.5579#supplemental-information>.

REFERENCES

- Amri S, Ltifi H, Ayed MB. 2015.** Towards an intelligent evaluation method of medical data visualizations. In: *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*. Piscataway: IEEE DOI [10.1109/ISDA.2015.7489198](https://doi.org/10.1109/ISDA.2015.7489198).
- Armstead I, Huang L, Ravagnani A, Robson P, Ougham H. 2009.** Bioinformatics in the orphan crops. *Briefings in Bioinformatics* **10(6)**:645–653 DOI [10.1093/bib/bbp036](https://doi.org/10.1093/bib/bbp036).
- Bassett L. 2015.** *Introduction to JavaScript Object Notation: A To-the-Point Guide to JSON*. Sebastopol: O'Reilly Media.
- Battle L, Chang R, Stonebraker M. 2016.** Dynamic prefetching of data tiles for interactive visualization. In: *Proceedings of the 2016 International Conference on Management of Data*. San Francisco: ACM, 1363–1375.
- Bikakis N. 2018.** *Big Data Visualization Tools*. New York: Springer, 11.
- Bolser D, Staines DM, Pritchard E, Kersey P. 2016.** Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. In: Edwards D, ed. *Plant Bioinformatics: Methods and Protocols*. New York: Springer, 115–140.
- Chen CM, Lai KJ, Pai TW, Chang HT. 2014.** Transcriptome data visualization in pathways with application to zebrafish embryo datasets. In: *2014 Eighth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS), Birmingham*. Piscataway: IEEE, 515–518 DOI [10.1109/CISIS.2014.74](https://doi.org/10.1109/CISIS.2014.74).
- Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang PL, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD. 2007.** Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2(10)**:2366–2382 DOI [10.1038/nprot.2007.324](https://doi.org/10.1038/nprot.2007.324).
- Cook P. 2017.** UK wind chart. Available at <http://charts.animateddata.co.uk/ukwind/> (accessed 28 October 2017).

- Damljanovic D, Agatonovic M, Cunningham H. 2010.** *Natural Language Interfaces to Ontologies: Combining Syntactic Analysis and Ontology-Based Lookup through the User Interaction*. Berlin, Heidelberg: Springer, 106–120.
- Groenendyk M. 2013.** Emerging data visualization technologies for map and geography libraries: 3-D printing, holographic imaging, 3-D city models, and 3-D model-based animations. *Journal of Map & Geography Libraries* **9(3)**:220–238 DOI [10.1080/15420353.2013.821436](https://doi.org/10.1080/15420353.2013.821436).
- Goeau H, Bonnet P, Joly A. 2017.** Plant identification based on noisy web data: the amazing performance of deep learning (LifeCLEF 2017). In: *CLEF 2017-Conference and Labs of the Evaluation Forum*. Dublin, Ireland: LifeCLEF, 1–13.
- Harris S, Seaborne A, Prud'hommeaux E. 2013.** SPARQL 1.1 query language. Available at <https://www.w3.org/TR/rdf-sparql-query/> (accessed 21 October 2017).
- Hearst MA, Laskowski P, Silva L. 2016.** Evaluating information visualization via the interplay of heuristic evaluation and question-based scoring. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose: ACM, 5028–5033 DOI [10.1145/2858036.2858280](https://doi.org/10.1145/2858036.2858280).
- Hebeler J, Fisher M, Blace R, Perez-Lopez A. 2009.** *Semantic Web Programming*. Hoboken: Wiley Publishing.
- Hughes S. 2006.** Opinion piece: genomics and crop plant science in Europe. *Plant Biotechnology Journal* **4(1)**:3–5 DOI [10.1111/j.1467-7652.2005.00164.x](https://doi.org/10.1111/j.1467-7652.2005.00164.x).
- IHME. 2017.** Life expectancy and probability of death | IHME Viz Hub. Available at <http://vizhub.healthdata.org/le/> (accessed 3 March 2017).
- Jackson H. 2015.** E&W property price map. Available at http://helenjacksonanalytic.co.uk/EW_prop_price.html (accessed 28 October 2017).
- Jaiswal P, Avraham S, Ilic K, Kellogg EA, McCouch S, Pujar A, Reiser L, Rhee SY, Sachs MM, Schaeffer M, Stein L, Stevens P, Vincent L, Ware D, Zapata F. 2005.** Plant ontology (PO): a controlled vocabulary of plant structures and growth stages. *Comparative and Functional Genomics* **6(7–8)**:388–397 DOI [10.1002/cfg.496](https://doi.org/10.1002/cfg.496).
- Jena A. 2017.** A free and open source Java framework for building Semantic Web and Linked Data applications. Available at <https://jena.apache.org/> (accessed 28 October 2017).
- Jensen PA, Papin JA. 2014.** MetDraw: automated visualization of genome-scale metabolic network reconstructions and high-throughput data. *Bioinformatics* **30(9)**:1327–1328 DOI [10.1093/bioinformatics/btt758](https://doi.org/10.1093/bioinformatics/btt758).
- Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. 2011.** ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Research* **39(suppl_1)**:D712–D717 DOI [10.1093/nar/gkq1156](https://doi.org/10.1093/nar/gkq1156).
- Kaski S, Peltonen J. 2011.** Dimensionality reduction for data visualization [applications corner]. *IEEE Signal Processing Magazine* **28(2)**:100–104 DOI [10.1109/MSP.2010.940003](https://doi.org/10.1109/MSP.2010.940003).
- Keim DA. 2002.** Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics* **8(1)**:1–8 DOI [10.1109/2945.981847](https://doi.org/10.1109/2945.981847).
- Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Schroder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, Karp PD. 2013.** EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Research* **41(D1)**:D605–D612 DOI [10.1093/nar/gks1027](https://doi.org/10.1093/nar/gks1027).

- Lin Z, Cao N, Tong H, Wang F, Kang U, Chau DHP. 2013.** Demonstrating interactive multi-resolution large graph exploration. In: *2013 IEEE 13th International Conference on Data Mining Workshops, Dallas*. Piscataway: IEEE, 1097–1100 DOI 10.1109/ICDMW.2013.124.
- List TP. 2013.** The plant list – a working list of all plant species. Available at <http://www.theplantlist.org/> (accessed 7 February 2018).
- Lohmann S, Link V, Marbach E, Negru S. 2015.** *WebVOWL: Web-Based Visualization of Ontologies*. Cham: Springer International Publishing, 154–158.
- Lu C, Deng L, Fei M. 2015.** An improved visualization modelling method of greenhouse tomato plants based on L-system. *Chinese Automation Congress (CAC), Wuhan*. Piscataway: IEEE, 480–485 DOI 10.1109/CAC.2015.7382548.
- Map V. 2017.** US map states – choropleth plus bar. Available at <https://vida.io/documents/4vZ9mRGyepoyQxFcK> (accessed 28 October 2017).
- Marsh L. 2017.** The USA National Phenology Network: a framework for delivery of phenology data products on multiple spatiotemporal scales. *Biodiversity Information Science and Standards* 1:e20188 DOI 10.3897/tdwgproceedings.1.20188.
- Marx V. 2013.** Biology: the big challenges of big data. *Nature* 498(7453):255–260 DOI 10.1038/498255a.
- Nielsen J. 1992.** Finding usability problems through heuristic evaluation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Monterey*. New York: ACM, 373–380 DOI 10.1145/142750.142834.
- NRCS U. 2017.** *The PLANTS Database*. Greensboro: National Plant Data Team. Available at <https://plants.usda.gov>.
- Paredes-Valverde MA, Rodríguez-García MÁ, Ruiz-Martínez A, Valencia-García R, Alor-Hernández G. 2015.** ONLI: an ontology-based system for querying DBpedia using natural language paradigm. *Expert Systems with Applications* 42(12):5163–5176 DOI 10.1016/j.eswa.2015.02.034.
- PFAF. 2017.** *The PFAF Database (Plant for a Future)*. Available at <http://www.pfaf.org/> (accessed 19 October 2017).
- Sedova M, Jaroszewski L, Godzik A. 2016.** Protael: protein data visualization library for the web. *Bioinformatics* 32(4):602–604 DOI 10.1093/bioinformatics/btv605.
- Smith B, Ashburner M, Rosse C, Bard J, Bug W, Ceusters W, Goldberg LJ, Eilbeck K, Ireland A, Mungall CJ, Leontis N, Rocca-Serra P, Ruttenberg A, Sansone S-A, Scheuermann RH, Shah N, Whetzel PL, Lewis S. 2007.** The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology* 25(11):1251–1255 DOI 10.1038/nbt1346.
- Tablan V, Damljanovic D, Bontcheva K. 2008.** A natural language query interface to structured information. In: Bechhofer S, Hauswirth M, Hoffmann J, Koubarakis M, eds. *The Semantic Web: Research and Applications: 5th European Semantic Web Conference, ESWC 2008, Tenerife, Canary Islands, Spain, June 1–5, 2008 Proceedings*. Berlin: Springer, 361–375.
- TDWG BIS. 2018.** Biodiversity information standards (TDWG). Available at <https://github.com/tdwg> (accessed 25 March 2018).
- Tegarden DP. 1999.** Business information visualization. *Communications of the AIS* 1:4.
- Telea AC. 2014.** *Data Visualization: Principles and Practice*. Second Edition. Wellesley: A. K. Peters, Ltd.
- Teller S. 2013.** *Data Visualization with d3.js*. Birmingham: Packt Publishing.

- Tello-Ruiz MK, Stein J, Wei S, Preece J, Olson A, Naithani S, Amarasinghe V, Dharmawardhana P, Jiao Y, Mulvaney J, Kumari S, Chougule K, Elser J, Wang B, Thomason J, Bolser DM, Kerhornou A, Walts B, Fonseca NA, Huerta L, Keays M, Tang YA, Parkinson H, Fabregat A, McKay S, Weiser J, D'Eustachio P, Stein L, Petryszak R, Kersey PJ, Jaiswal P, Ware D. 2016. Gramene 2016: comparative plant genomics and pathway resources. *Nucleic Acids Research* 44(D1):D1133–D1140 DOI 10.1093/nar/gkv1179.
- Tory M, Moller T. 2004. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics* 10(1):72–84 DOI 10.1109/TVCG.2004.1260759.
- Toufighi K, Brady SM, Austin R, Ly E, Provart NJ. 2005. The Botany Array Resource: e-Northern, Expression Angling, and promoter analyses. *Plant Journal* 43(1):153–163 DOI 10.1111/j.1365-313X.2005.02437.x.
- Tran VX, Tsuji H, Masuda R. 2009. A new QoS ontology and its QoS-based ranking algorithm for Web services. *Simulation Modelling Practice and Theory* 17(8):1378–1398 DOI 10.1016/j.simpat.2009.06.010.
- UCONN. 2017. UCONN Plant Database. University of Connecticut. Available at <http://hort.uconn.edu/index.php> (accessed 15 October 2017).
- Varga B, Trambitas-Miron AD, Roth A, Marginean A, Slavescu RR, Groza A. 2014. LELA—a natural language processing system for Romanian tourism. In: 2014 Federated Conference on Computer Science and Information Systems. Piscataway: IEEE DOI 10.15439/2014F323.
- Walls RL, Athreya B, Cooper L, Elser J, Gandolfo MA, Jaiswal P, Mungall CJ, Preece J, Rensing S, Smith B, Stevenson DW. 2012. Ontologies as integrative tools for plant science. *American Journal of Botany* 99(8):1263–1275 DOI 10.3732/ajb.1200222.
- Wang C, Xiong M, Zhou Q, Yu Y. 2007. PANTO: a portable natural language interface to ontologies. In: Franconi E, Kifer M, May W, eds. *The Semantic Web: Research and Applications: 4th European Semantic Web Conference, ESWC 2007, Innsbruck, Austria, June 3–7, 2007 Proceedings*. Berlin: Springer, 473–487.
- Willis KJ. 2017. State of the World's Plants 2017. Report Royal Botanic Gardens, Kew. Available at https://stateoftheworldsplants.com/2017/report/SOTWP_2017.pdf.
- Yang J, Ward MO, Rundensteiner EA. 2003. Hierarchical exploration of large multivariate data sets. In: Post FH, Nielson GM, Bonneau G-P, eds. *Data Visualization: The State of the Art*. Boston: Springer, 201–212.
- Yuk M, Diamond S. 2014. *Data Visualization for Dummies*. Hoboken: John Wiley & Sons, Inc.
- Zhiguo D, Zhengjie D. 2010. Improved ontology ranking algorithm based on semantic web. In: 2010 3rd IEEE International Conference on Ubi-Media Computing. Piscataway: IEEE DOI 10.1109/UMEDIA.2010.5543916.
- Zinsmaier M, Brandes U, Deussen O, Strobel H. 2012. Interactive level-of-detail rendering of large graphs. *IEEE Transactions on Visualization and Computer Graphics* 18(12):2486–2495 DOI 10.1109/TVCG.2012.238.