Geochemistry, Geophysics, Geosystems

Key Points:
- The Petrology Workspace and Database (PWD) offers an online interactive data repository for the management and visualization of multilevel data sets
- The *Workspace* tool allows to organize data into tree-like structures, useful for the spatial context of data at different scales. Plotting tools are also available
- The PWD aims to provide an interactive data repository that facilitates data transparency, effective data communication, and sharing

Supporting Information:
- Supporting Information S1

Correspondence to:
F. Costa,
fcosta@ntu.edu.sg

PWD: A Petrological Workspace and Database tool

Dayana Schonwalder-Angel[1], Helena Albert[1,2], Lam Ngoc Tuan[3], Nguyen Xuan Phi[3], Sri Budhi Utami[4], Weiran Li[1], Phung Minh Khanh[3], Christina Widiwijayanti[1], and Fidel Costa[1,4]

[1]Earth Observatory of Singapore, Nanyang Technological University, Singapore, [2]Central Geophysical Observatory, Spanish Geographic Institute (IGN), Madrid, Spain, [3]School of Computer Science and Engineering, Nanyang Technological University, Singapore, [4]Asian School of the Environment, Nanyang Technological University, Singapore

Abstract The Petrology Workspace and Database (PWD; https://petro.wovodat.org) is a web-based data repository and interface that allows researchers to access, share, store, and manage petrological, mineralogical, and whole rock data in a contextualized manner. The uniqueness of the PWD is that it links images to different types of images and to compositional data providing a powerful visualization and framework of information at a wide range of scales, from the meter-sized outcrop to a few micrometers of a thin rock section. The PWD archives various data types and formats, and it includes multilevel data sets with an interactive online interface. The database is linked with other databases for volcanic eruptions (Smithsonian' Global Volcanism Program (GVP)) and for whole rock chemistry (EarthChem). The tool has four main features: (1) storage and management of spatial-referenced data (e.g., from fieldwork notes to lab geochemical analysis); (2) a hierarchical relationship between different types of data using the *Workspace* interactive tool; (3) graph plots to visualize the data; and (4) the possibility of data sharing in a database structure that is managed by authorization levels. The PWD is a practical and efficient database management system that facilitates effective contextualized data preservation and sharing among scientists. It also can serve as a Data Management Plan and provide a framework for auditing research integrity both of which are becoming the new standards of most funding agencies and journals.

Plain Language Summary The Petrology Workspace and Database is an online and standalone interactive data repository that allows management and visualization of a wide range of petrological and geochemical rock information. Users can organize many different data types at different scales into tree-like structures. The system allows creation of links reflecting the spatial context between data that is critical for correct interpretation of the information. The Petrology Workspace and Database can serve as a Data Management Plan and provides a framework for auditing of research integrity.

1. Introduction

Petrological and geochemical studies of rocks use a wide range of data types at different spatial and georeferenced levels, from field observations to in situ quantitative analysis at the microscopic scale. The wide variety of data and petrological and geochemical analytical techniques allow researchers to collect a large amount of information that is not only different in nature and format (e.g., field photographs, instrumental imaging, quantitative analysis, and metadata) but also has complex spatial relationships. For example, a single electron microprobe analysis from a mineral at a micrometer scale needs to be related to a thin rock section that is several centimeters in size, which in turn is made from a hand specimen several decimeters in size, that was sampled from an outcrop of several meters. For proper interpretation, data preservation, and auditing of scientific procedures, it is imperative that published data sets can be contextualized and traced back between the various scales. However, there is currently no software or application tool that allows this, and thus many petrological data sets are not being properly managed at various levels. It is the purpose of this paper to alleviate this situation by providing a new online interactive tool, the Petrology Workspace and Database (PWD).

Geochemical and petrological studies typically follow a workflow involving several steps, starting with macroscopic sample collection in the field (specimens of several decimeters in size) and detailed description of the rocks. The information gained from this step can have already a variety of observations that are spatially variable (e.g., hand samples may be heterogenous as in banded pumices), and it is the basis for further detailed observational steps, including interpretation of microscopic textures at thin section

scale, identification of suitable phases for further analyses (e.g., minerals, glass, inclusions), compositional phase characterization, calculations of various types from the acquired data (e.g., geothermobarometry), and finally interpretation of underlying geological processes. Each of these "steps" creates its own specialized data set, and a comprehensive interpretation often relies on the complementary nature of their spatial relationships. For example, in volcanic rocks, the significance of compositional point analyses in zoned minerals depends on the location of the measurement within the crystal (e.g., core versus rim) (Blundy & Cashman , 2008). Likewise, in the case of metamorphic rocks the compositions and textures of neighboring crystals and the rock matrix are interrelated, providing information about the degree of recrystallization of the rock and the fluids involved during metamorphism (Vernon, 2004). Since petrological studies rarely focus on a single crystal or a single rock, the growing amount of data creates significant challenges to efficiently organize and archive the information systematically (Figure 1). Moreover, the lack of standardization results in ineffective data visualization, tracking, reconstructing, and sharing among collaborators. It is therefore important for researchers to be able to document how these distinct types of data at each step are interconnected and avoid losing the opportunity to develop meaningful connections between the different types of data that can be tracked for years to come or reconstructed by researchers who were not part of the original data collection process. These points above are the main aims of PWD, which provides a platform to archive and relate data sets, enabling data to be fully presented and preserved within a spatial context.

There are currently various efforts to develop databases that store and organize data related to the Earth and Environmental Sciences, from a Geographic Information System (GIS) database for the study of continental collisions (Khan et al., 2006) to comprehensive compilations of crystal structures and experimental magmatic phase equilibria data (Downs & Hall-Wallace , 2003; Hirschmann et al., 2008). Global-to local-scale catalogs of geochemical data sets, such as PetDB' EarthChem (www.earthchem.org/petdb), NAVDAT (Walker et al., 2006), GEOROC (http://georoc.mpch-mainz.gwdg.de) (Downs & Hall-Wallace, 2003; Lehnert et al., 2000), and GeoREM (Jochum et al., 2005; Jochum & Nohl, 2008), as well as volcano information populated from the Smithsonian GVP' Volcanoes of the World database (Global Volcanism Program, 2013), have centralized large amount of data that are readily available for researchers. Other online interfaces like WOVOdat (https://wovodat.org) (Newhall et al., 2017; Costa et al., 2019) and V:hub (https://vhub.org) (Palma et al., 2014) have provided interactive spaces where collaborators can share data about volcanic unrest alongside modeling and analyses tools. However, researchers still lack an interactive tool that allows organizing multilevel petrology and geochemical data sets in a systematic manner, where the spatial relation and interrelation between files are clearly established. This is especially significant in large projects where several researchers, sometimes from multiple institutions, are involved. This is also especially relevant when researchers want to study the same samples that were studied in the past with new analytical tools. In these situations, data organization and sharing becomes challenging owing to personal preferences in data archiving and/or the turnover of researchers within a given project.

From another perspective, the large quantity of data generated by the scientific community has led funding agencies, publishers, and governmental agencies to require data management plans that not only guarantee proper archival of the data but also long-term maintenance and auditing for research integrity (Stall et al., 2019; Wilkinson et al., 2016). Ideally, data publication and preservation should follow the fundamental principles of findability, accessibility, interoperability, and reusability to ensure transparency, reproducibility, and reusability (Stall et al., 2019; Wilkinson et al., 2016). Although numerous general-purpose repositories have been created to make data available (Roche et al., 2015; Wilkinson et al., 2016), improvement is still needed for practices relating the repositories management to ensure data reproducibility (Roche et al., 2015; Sandve et al., 2013; White et al., 2013). Digital exchange and transfer of scientific knowledge in a collaborative environment is not a trivial task (Bechhofer et al., 2010). Failure in describing methods, incomplete data sets, insufficient metadata and/or inadequate file formats, reluctance to share data due to competition, and perception that the process is technically difficult and time consuming (Palmer et al., 2004; Parr & Cummings , 2005; Roche et al., 2015) are some of the problems that can hinder data evaluation. However, data sharing initiatives that allow access limitations and are aided by user-friendly data archives can ameliorate these issues (White et al., 2013). Overall, data need to be understandable, easy to analyze, and readily available; metadata, including information on how the data were collected, units of
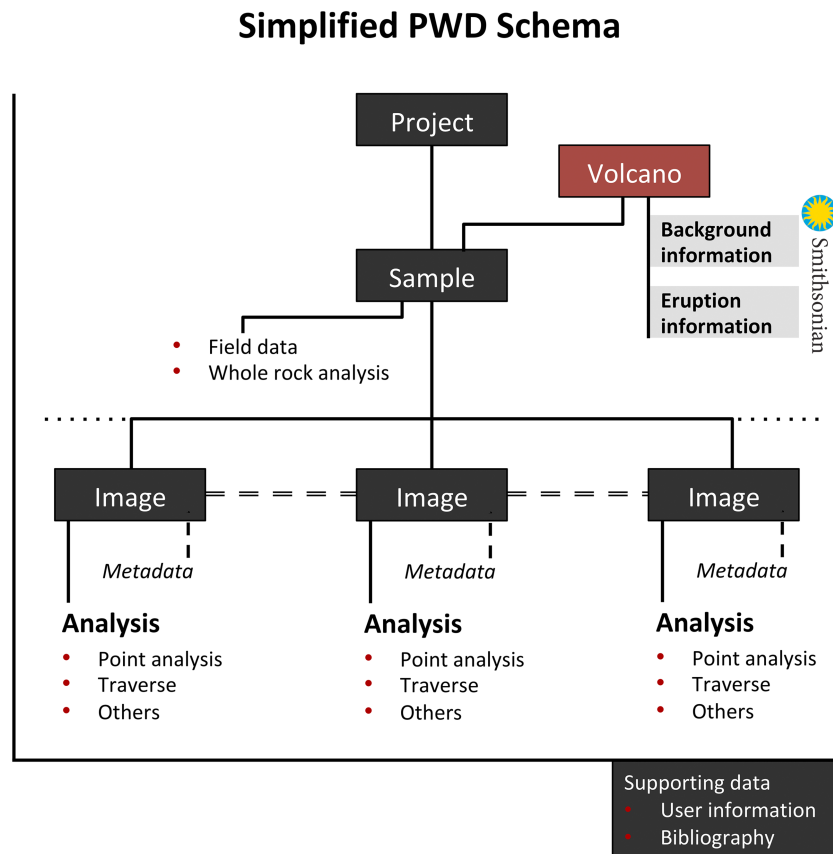
AGU
100
ADVANCING EARTH
AND SPACE SCIENCE

Geochemistry, Geophysics, Geosystems

10.1029/2019GC008710

Figure 1. Simplified schema of Petrology Workspace Database (PWD) showing hierarchy and link between the storage tables. The *Project* is the parent in the PWD hierarchical data, which can have multiple *Samples*. Volcanic samples can be linked to its *Volcano* and *Eruption*. Dotted lines express that a *Sample* can have many *Images*. It is possible to create many-to-many relations (double-dashed lines) between images. Any field data and description will be archive in *Sample* table. User Information table will be linked to all tables, and it defines the user's authorization level to access/edit data. *Bibliography* table is a supporting information that can be linked to other tables. Detailed schema and structure of the database can be seen in the supporting information; (Schonwalder Angel, 2019).

measurements, and descriptions should be logically organized, complete, and sufficiently clear to enable interpretation and make it easier to combine different data sets (Michener et al., 1997; White et al., 2013; Zimmerman, 2007).

In this paper we present the PWD, an online data management system that standardizes how data are systematically archived and organized using interactive tools (Figure 1), to facilitate data visualization, reconstruction, and sharing, as the links between files are interactively displayed. The PWD provides useful tools such as the creation of simple data plots (e.g., scatter diagrams of geochemical data) and archiving publications and references related to specific projects. These features allow the centralization of all the information related to a project from field data to analytical results, facilitating data query.

2. Database Platform, Structure, and Schema

The PWD was originally created to organize volcanic rock samples, but it is based on *Project* folders that allow also igneous, metamorphic, or sedimentary rocks (Figure 1). The MongoDB nonrelational database (https://www.mongodb.com/) was chosen to archive various petrological data types (e.g., image, photo, spreadsheet, and text). Any data uploaded in the database will be then stored in a table, a tabulated document with a specific format following the PWD standard format and structure (supporting information; (Schonwalder Angel, 2019)). The framework of the PWD consist of several tables with each *Sample* linked to a *Project*. For volcanic samples we have also added the *Eruption* and *Volcano* table, the latter being the center point of the data

structure from which other data can be linked. The *Eruption* and *Volcano* tables have been populated with information from the Smithsonian GVP (Global Volcanism Program*, 2013*). If a rock sample has been already archived in EarthChem it can be linked through the use of the International Geo Sample Numbers. For a given *Sample*, the link between files (e.g., images and metadata) and the sample is created by uploading files to the *Sample* information page. The creation of links with spatial context (i.e., image to image and image to compositional data) are done, interactively, through a *Workspace* tool and the *Image* information page. In addition, there is a *Bibliographic* table with references; if one or more references are related with a particular sample, the link can be created at the *Sample* information page. More detailed information about the tables, schema, and structure can be found in the supporting information (schema and structure and back-end) (Schonwalder Angel, 2019).

3. PWD Interface

The PWD is a web application currently hosted by a server at Nanyang Technological University, and it is open publicly to allow real-time interactive management. For optimal functionality, we recommend using Google Chrome® as a browser. The open source package is also provided freely to individuals and institutions who wish to use the system on their own server. The package can be downloaded through our website (https://petro.wovodat.org/#/open_source). The step-by-step settings and installation documentation and the user manual can be found in the supporting information (Schonwalder Angel, 2019). Here we briefly describe the main functionality of the PWD. The images from the user manual that are references in this document will appear as "Figure SD_N", where N is the figure sequence in the user manual.

The main PWD page has five menus (Figure 2): Input, Access, Tutorial, Source, and Account. The *Input* menu allows users to populate the PWD with new data organized in five main categories as described below: project, sample, volcano, eruption, and reference (Figures SD_1–SD_9). In the *Access* menu, users can see a volcano world map (Global Volcanism Program*, 2013*) with projects and samples registered in the system (Figures SD_10–SD_13). The volcano icons and the names are hyperlinks to their respective information pages. In the *Tutorial* and *Source* menus, users can find detailed documentations and download the standalone package and installation guide. The *Account* menu shows the user information and list of all projects and samples created by them or that have been shared by other users, with read-only or read-write authorization level.

3.1. Data Input

To start a project or add new data to an existing project, users are required to log in. For this user need to register/create an account by clicking the "Login/Sign Up" menu, where they select their role as a researcher and/or principal investigators (PI).

3.1.1. New Project

To create a new project, users provide a unique name and brief description and select within the list of registered users the name of the project's PI (Figure SD_2). Users can also set the accessibility of the data with four authorization levels: private (restricted to researcher and PI), read-only, public, and customized access (i.e., specific users). At any time, the PI and system administrator can modify the level of data accessibility or user authorization to prevent data redundancies/duplicity.

3.1.2. New Sample

For a new sample there are two options: Users can complete an online form for an individual sample (Figures SD_3 and SD_4), or for multiple samples, users can upload a spreadsheet following the template provided in the webpage (Figures SD_5 and SD_6). Each new sample needs to be linked to a project. Geographic coordinates are required, while other attributes (e.g., rock composition and deposit type) are optional and can be added later. The system automatically creates an information page for each sample, which displays its location on an interactive Google®-based map (Figures SD_16–SD_19). Here users can modify information by making changes in the page (i.e., edit option) or add new information by uploading any relevant files such as compositional data tables (e.g., whole-rock analysis), the age of the eruption unit and the geochronology technique being used, and images (e.g., photomicrographs). The system archives the information and keeps the data readily available. For volcanic rocks we encourage users to link the sample to their source volcano and/or eruption. This generates more connectivity between samples within the system, which assists its query within the database.
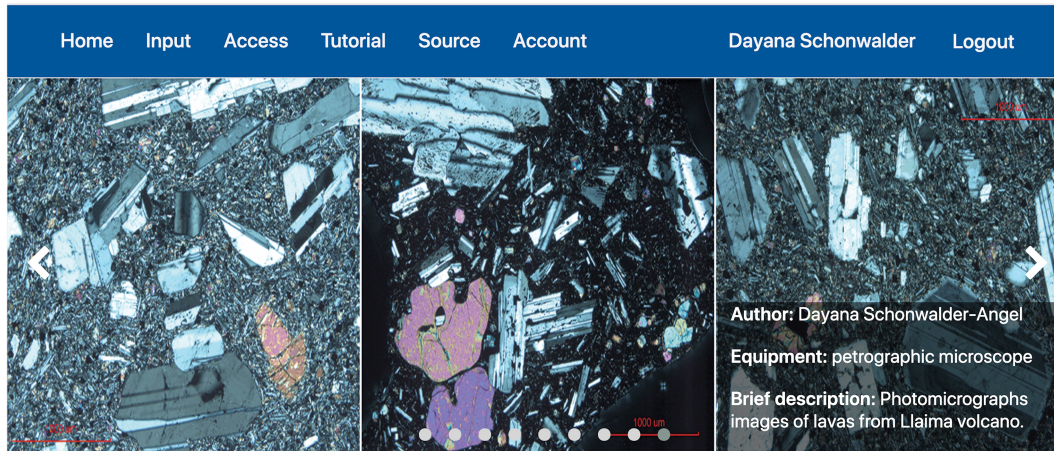
Figure 2. Screenshot of the PWD homepage with the blue bar showing its main menus.

3.1.3. Volcano, Eruption, and Bibliography

Users need to complete an online form for adding volcanoes, eruptions, and references that are not preregistered within the PWD (Figures SD_7–SD_9). The system automatically creates a volcano page (Figure 3) that displays an interactive Google®-based map and general information such as type of volcano and type of eruptions. For eruptions, projects, and samples that are linked to the volcano, they will be listed in this page and the location of the samples are visible on the map; these are hyperlinked to their respective information pages. Users can also link papers to specific samples and/or volcanoes, by selecting a particular publication from a dropdown menu that appears in each sample, volcano, and eruption page. We encourage users to provide more complete information regarding each new data entry, with the goal of creating a robust database.

Figure 3. Example of a volcano page at the PWD. The image is a screenshot of the information page for Piton de la Fournise, Réunion Island. An interactive Google® map shows the locations of the samples that had been linked to this volcano. By clicking the "rock" icon on the map, a pop-up window appears showing the sample name and its geographical coordinates.
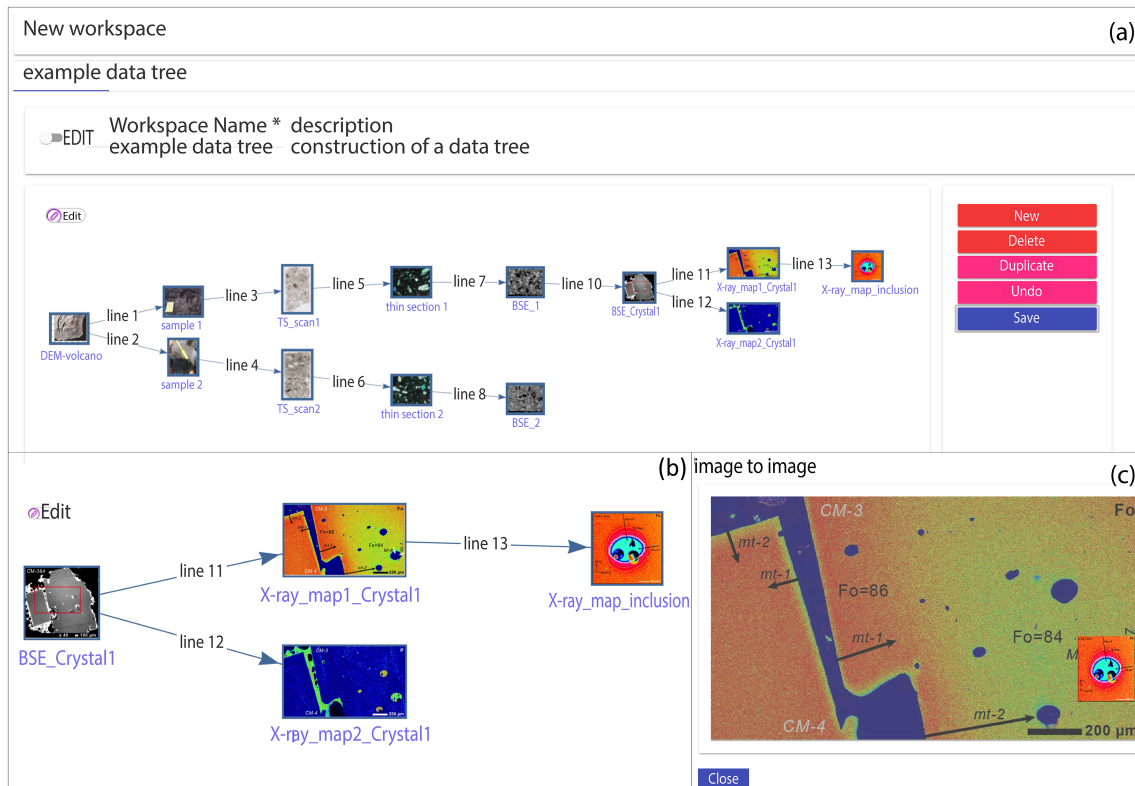
Figure 4. Example of a data tree in the PWD workspace. (a) Screenshot of a workspace with a data tree created to establish a hierarchy between images at different spatial scales: from the regional-scale map of a volcano down to the micrometers-scale image of a glass inclusion. (b) Screenshot of a magnified view of the data tree shown in panel a. (c) Screenshot of the pop-up window that shows the scale and location of a "child" image on its "parent" image. The digital elevation model (DEM), crystal and inclusion images extracted from Albert et al. (2019).

3.2. Data Management

This section refers to how the users can interact with the data in the PWD. Data management is not a menu per se but rather a set of tools so users can link their data (e.g., image-to-image, image-to-geochemical analysis, and image-to-file) using the interactive *Workspace* and image analysis tools.

3.2.1. Workspace

The *Workspace* is analogous to a virtual canvas that provides users a way to organize and visualize the hierarchy between images for a specific sample (Figure 4). In the *Workspace*, users can build "parent-child" links between images to generate data trees (Figures SD_20–SD_32), which are useful to visualize spatial relations between images at different scales or between images acquired with different instruments or techniques. The *Workspace* feature is available at the *Sample* page, where users can create, modify, duplicate, or delete as many *Workspaces* as needed. Upon creating a new *Workspace*, users will be presented with a blank space, where the "node" and the "add edge" options are available. The "node" allows users to call any image that may be uploaded to the sample page, whereas the "add edge" allows users to create links between the images in a "parent" "child" structure. One-to-many and many-to-many links can be established. At the *Workspace*, each image is also a hyperlink to its information page where links to compositional data can be established.

3.2.2. Image

The PWD creates an information page for any uploaded image, where users find all the image metadata (at the visibility tab, Figure 5a) and can also create links between images and any compositional data (at the point analysis and traverse tabs, Figure 5a). The functionality of this link option is similar to the GIS proposed by Linzmeier et al. (2018). However, with PWD users are able to create links between images and other data types (e.g., metadata, publication, and spreadsheet) interactively on the web interface without the need to install a GIS software.

3.2.2.1.   Point Analysis

At the point analysis tab, in the *Image* page (Figure 5), users are able to create links between the image and in situ compositional data, and they can also create diverse plot charts (e.g., line, bar, pie, ternary, and scatter) to visualize such data. To input in situ compositional data, users upload an Excel® spreadsheet following a given template. Users can include major and trace elements, as well as isotopic and geochronologic data. The location (spatial context) of each in situ measurement with respect to the image can be indicated in the spreadsheet, if coordinates are known (e.g., from metadata generated by the instrument used in the analyses), or manually by clicking on the image (Figures SD_3–SD_41). In the spreadsheet users can also assign colors to the data points (Figures 5a and SD_33–SD_44). Once the data have been uploaded, the system displays an interactive table that contains an identification of each point, their colors and coordinates, and two "empty" element columns, which can display any data that have been uploaded in the Excel® file. The users select the element they want to inspect (e.g., SiO_2), and the system will automatically populate the table (Figures 5a and SD_41–SD_44). The user can also choose from a range of plot charts to visualize the data (Figures 5b and SD_44–SD_63).

3.2.2.2.   Traverse

The Traverse tab (Figure 6a) provides a similar functionality as the Point Analysis tab, where users can upload the compositional data measured along a traverse and define its location on the image by selecting the position of the first and last points. A line marking the traverse position will be automatically displayed on the image, and a variation diagram of concentration versus distance will be created (Figure 6b). The diagram is interactive, and it displays each element data uploaded in different colors. Generally, element concentrations ranged in several orders of magnitude (e.g., from weight percent to ppm), so, when plotted together, their variation along the traverse might not be easily observable, due to the difference in the scale. To solve this, users can select or deselect (on the diagram) the elements they want to highlight and compare; the diagram will automatically rescale to show the units of the elements chosen. The diagram is displayed below the image, so users can have a spatial context of the variation of a particular element(s) along the traverse and interactively compare the characteristics of each element with respect to the image and to other elements.

4. Case Studies: Application and Examples

The chemical and physical characteristics of minerals are fundamental for understanding the processes that formed the rock (e.g., Blundy & Cashman, 2008). After acquiring various types of compositional and mineralogical information for one or more crystal(s), it is necessary to combine them to gain a more comprehensive insight of the geological context, which could span a range of spatial scales and/or timescales. However, this is not easy to do, and below we provide a few examples to illustrate how the PWD facilitates such comparisons by constructing data tree-like structure in the *Workspace* and creating plot charts and compositional traverse charts that are linked to particular images.

4.1. Simple Comparison of Petrologic Information Using Data Trees (*Workspace* Tool)

We created a data tree (Figure 4) with a hierarchy of images at various scales. We start with the spatial relation between a main "parent" image (e.g., digital elevation model of a volcano) and several "child" images. In this case, the "child" images include pictures of two hand samples collected from different locations of the volcano, the scanned images of the petrographic thin rock sections, the backscattered electron images acquired with a scanning electron microscope from various areas within these thin sections, and the X-ray compositional maps of different elements obtained with an electron microprobe in individual crystals and their melt inclusions (Figure 4). With the links between images made by the arrows, it is easy to reconstruct the connection between images. This is critical information to have if for example another researcher wishes to know exactly where all these images where acquired for the purpose of further analysis several years later or for research integrity purposes.

Due to the interactive nature of the *Workspace*, users can zoom on the interface and observe particular images and their relationships within the data tree. Figure 4b shows a "branch" of the data tree with a one-to-many relation between a backscattered image of a crystal ("parent") to two X-ray maps ("children") of the same crystal. Each child image contains the spatial distribution of one element. Figure 4b also shows that this crystal hosts several melt inclusions of different sizes. If the compositions of these inclusions have
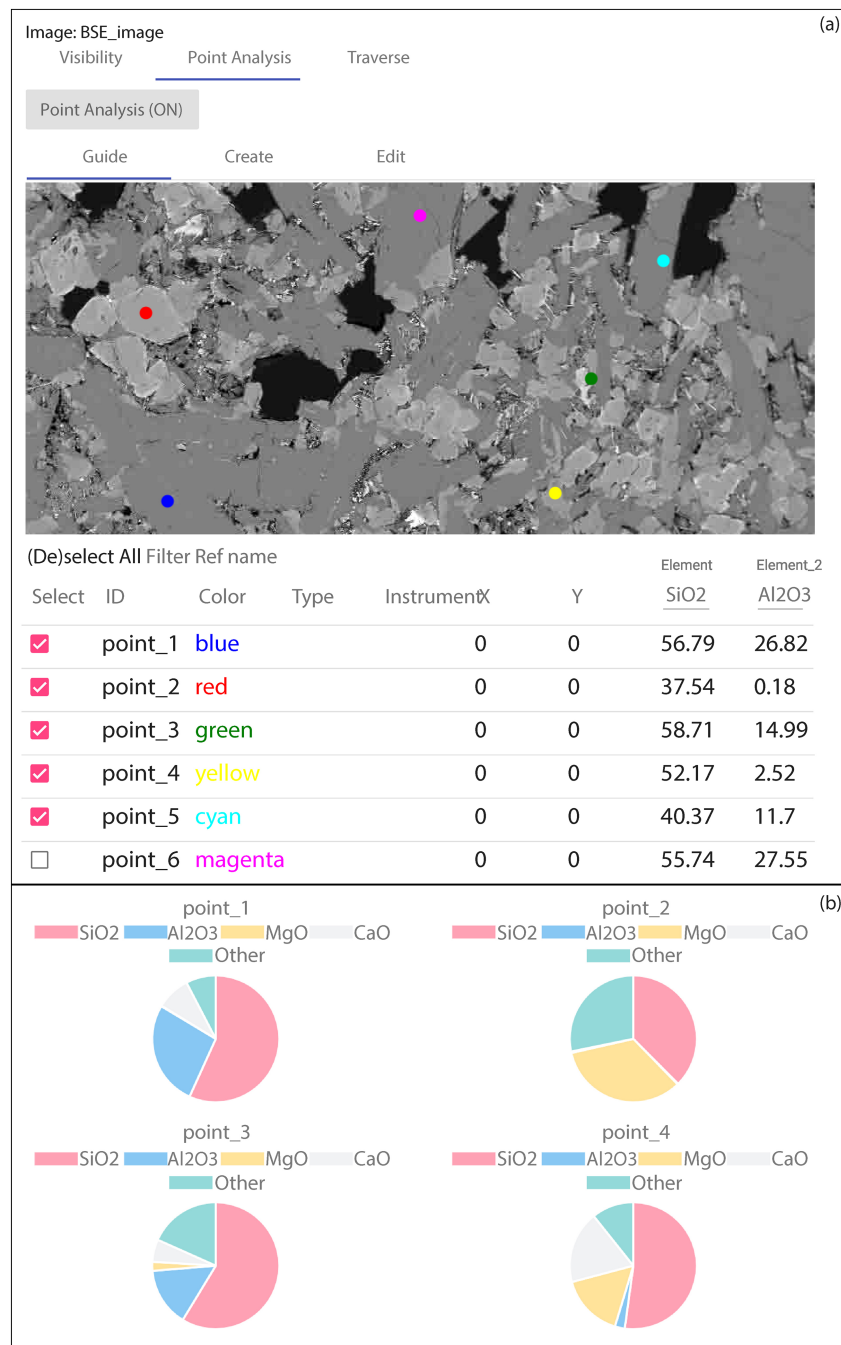
Figure 5. Point analysis example. (a) Screenshot of the image page showing a backscattered electron image of the sample. Location and color of the points was determined when the compositional data were uploaded. The interactive compositional table appears below the image, and it shows the list of data points and their respective colors. The first column on the left allows users to select the points they want to plot; the last two columns, on the right, allow users to type in the element of interest, and the table automatically populates such data. (b) Example of a pie plot for data point selected on the table; each color represents a different element.

been analyzed, like in this example, they can be added to the data tree with the locations of data collection marked on the "parent" image. Figure 4c shows the location of the melt inclusion (i.e., as in the "child" image) in the crystal (i.e., as in the "parent" image), facilitating the identification of this inclusion among the others, and also providing of its location within the crystal. From this example we show how the
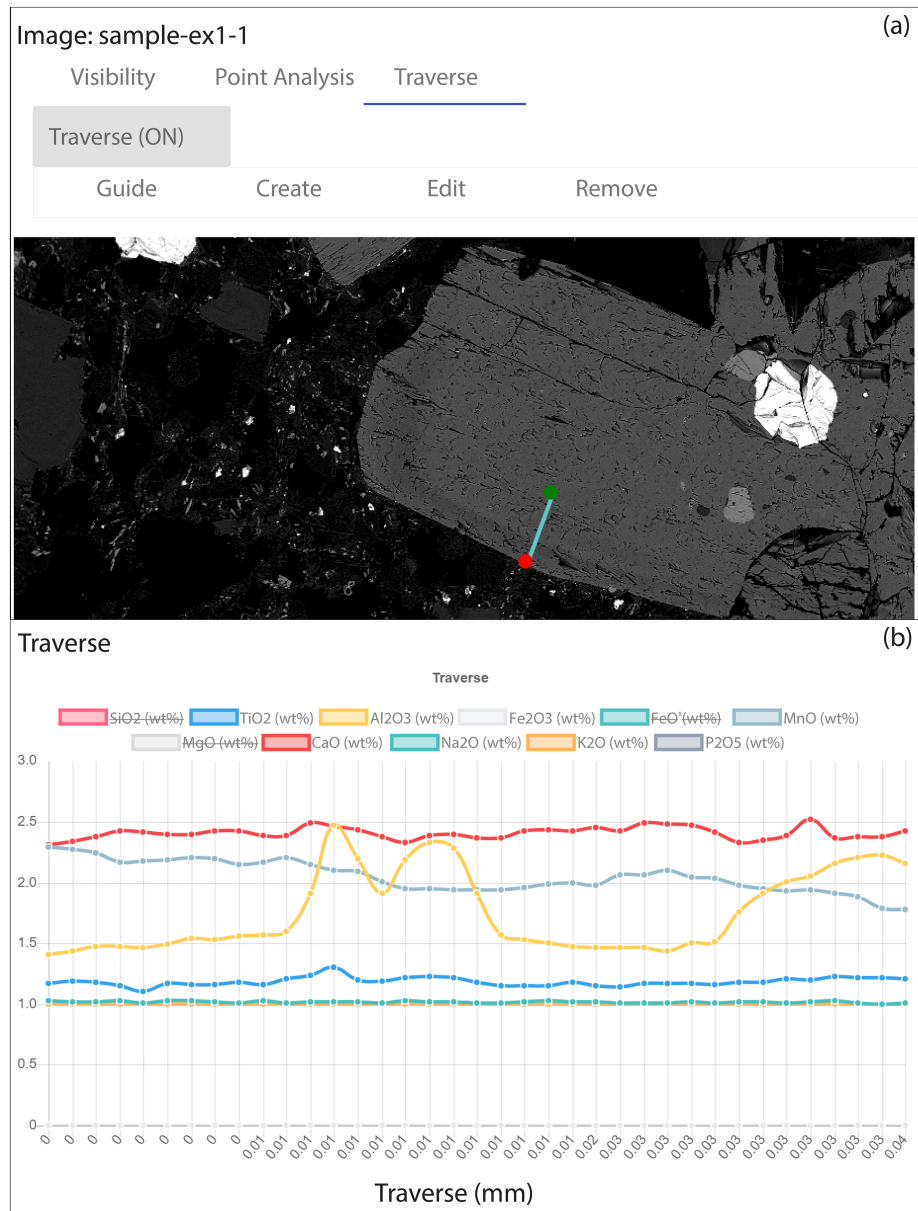
Figure 6. Compositional traverse analyses. (a) Screenshot of the traverse tab at the image information page showing a backscattered image of a zoned crystal. The red dot is the start and the green dot the end of the traverse. (b) Example of the interactive diagram for the traverse where only elements with similar concentrations ranges have been selected for better visualization of the concentration changes.

hierarchy and link between images is managed using the *Workspace* tool, which facilitates not only data reconstruction but also comparison between images with a spatial context. Figures SD_20–SD_32 in the supporting information show the step-by-step instructions for the creation of data trees. Readers can also explore this example online (at https://petro.wovodat.org/#/access/sample/5d5b6679b3b2e736187888f9).

4.2. Creation of Plot Charts for In Situ Analysis and Their Spatial Context (Point Analysis Tool)

Quantitative point analysis (i.e., in situ) of different phases are a cornerstone of petrological studies, and their interpretation requires the knowledge of their spatial context. The PWD allows to directly connect the images of the phases (e.g., through a Back-scattered electrons (BSE)) to the location of the analysis and the concentrations of the elements or isotopes (Figure 5). This allows to interactively visualize the

compositional variations of different crystals (Figure 5a) found at the interactive table (Figure 5b). Moreover, PWD also allows creating a range of plot charts (Figures 5b–5e and SD_45–SD_62) from bars to scatter plots and ternary diagrams. The visualization of the composition of each data analysis is interactive and includes the identification and the concentration of each element. This example illustrates the functionality of the point analysis tab to quickly inspect in situ measurements for one or more elements for multiple data points simultaneously. It also illustrates the easiness to acquire additional data points for the years to come in a manner that analyses from different users and instruments can be combined. Figures SD_33–SD_62 in the supporting information contain step-by-step instructions for point analysis and creation of plot charts. Readers can also explore this interactive function online (at https://petro.wovodat.org/#/access/imagealy/5d5b9bde7166025573dc14dd).

4.3. Visualization of Spatial Variation of Compositional Data (Data Traverse Tool)

The study of the textural characteristics of crystals and in particular the crystal compositional zonation gives clues to the magmatic/metamorphic processes and timescales involved (Costa et al., 2008; Morgan & Blake , 2006). Such information typically requires the acquisition of compositional traverses across crystals, and information on their location within the crystal is critical for proper interpretation. The PWD has a data traverse tool (see section 3.2.2.2., Figure 6) to visualize crystal zonation (e.g., in a BSE image; Figure 6a) and simultaneously see the one-dimensional traverse of compositional data that is linked to the image (Figures 6b and 6c). Because elemental concentrations in crystal vary by orders of magnitude, the plot tool allows choosing the elements with similar concentrations for better visualization. In this example, by deselecting SiO_2 (pink line), FeO (green line), and MgO (light gray line), the diagram expands at the y axis, to show elements with lower concentrations, and the detail of their behavior along the traverse. The exact concentration of the element and its location within the traverse can be readily obtained from the plot. Figures SD_63–SD_68 in the supporting information contains step-by-step instructions for the compositional traverse analysis. Readers can also explore this interactive function online (at https://petro.wovodat.org/#/access/imagealy/5d777dd1cc8f905fc9f1d739).

5. Summary and Future Work

Most petrological and geochemical data are acquired over a wide range of spatial and textural scales (from the field sample to a few micrometers). Proper interpretation of the data requires integration at all scales, and it is critical to preserve the contextual information for future analysis and evaluation of research data and results by third parties. The PWD can assist users to accomplish this and actively interact with their data sets in a way that is not currently possible by other petrologic databases. The integrated functionality of PWD as an online tool and as a data archive with uniform data format, systematic structure, and user-friendly interface makes it ideal as platform to support a Data Management Plan that meets findability, accessibility, interoperability, and reusability standards. We will develop a complementary system to the PWD that will allow users to perform data mining, advance query tools allowing statistical data analysis, for example, comparing patterns of crystal zonation for rock samples formed by eruptions in basaltic systems. This goal will be possible if the database is populated by the community.

References

Albert, H., Costa, F., Di Muro, A., Herrin, J., Métrich, N., & Deloule, E. (2019). Magma interactions, crystal mush formation, timescales, and unrest during caldera collapse and lateral eruption at ocean island basaltic volcanoes (Piton de la Fournaise, La Réunion). *Earth and Planetary Science Letters*, *515*, 187–199. https://doi.org/10.1016/j.epsl.2019.02.035

Bechhofer, S., D. De Roure, M. Gamble, C. Goble, and B. Iain (2010). Research objects: Towards exchange and reuse of digital knowledge. Nature Precedings. https://doi.org/10.1038/npre.2010.4626.1

Blundy, J., & Cashman, K. (2008). Petrologic reconstruction of magmatic system variables and processes. *Reviews in Mineralogy and Geochemistry*, *69*(1), 179–239. https://doi.org/10.2138/rmg.2008.69.6

Costa, F., Dohmen, R., & Chakraborty, S. (2008). Time scales of magmatic processes from modeling the zoning patterns of crystals. *Reviews in Mineralogy and Geochemistry*, *69*(1), 545–594.

Downs, R., & Hall-Wallace, M. (2003). The American Mineralogist crystal structure database. *American Mineralogist*, *88*(1), 247–250.

Costa, F., Widiwijayanti, C., Nang, T., Fajiculay, E., Espinosa-Ortega, T., & Newhall, C. (2019). WOVOdat – the global volcano unrest database aimed at improving eruption forecasts. *Disaster Prevention and Management*, *28*(6), 738–751. https://doi.org/10.1108/DPM-09-2019-0301

Global Volcanism Program (2013). Volcanoes of the world, v. 4.8.3. Venzke, E (ed.). Smithsonian Institution. Downloaded 25 Oct 2019. https://doi.org/10.5479/si. GVP.VOTW4-2013

Hirschmann, M. M., Ghiorso, M. S., Davis, F. A., Gordon, S. M., Mukherjee, S., Grove, T. L., et al. (2008). Library of Experimental Phase Relations (LEPR): A database and Web portal for experimental magmatic phase equilibria data. *Geochemistry, Geophysics, Geosystems*, *9*, Q03011. https://doi.org/10.1029/2007GC001894

Jochum, K. P., & Nohl, U. (2008). Reference materials in geochemistry and environmental research and the GeoReM database. *Chemical Geology*, *253*(1-2), 50–53.

Jochum, K. P., Nohl, U., Herwig, K., Lammel, E., Stoll, B., & Hofmann, A. W. (2005). GeoReM: A new geochemical database for reference materials and isotopic standards. *Geostandards and Geoanalytical Research*, *29*(3), 333–338.

Khan, S. D., Flower, M. F. J., Sultan, M. I., & Sandvol, E. (2006). Introduction to TETHYS—An interdisciplinary GIS database for studying continental collisions. *Journal of Asian Earth Sciences*, *26*(6), 613–625.

Lehnert, K., Su, Y., Langmuir, C. H., Sarbas, B., & Nohl, U. (2000). A global geochemical database structure for rocks. *Geochemistry, Geophysics, Geosystems*, *1*(5). https://doi.org/10.1029/1999GC000026

Linzmeier, B. J., Kitajima, K., Denny, A. C., & Cammack, J. N. (2018). Making maps on a micrometer scale. *Eos*, *99*. https://doi.org/10.1029/2018EO099269

Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B., & Stafford, S. G. (1997). Nongeospatial metadata for the ecological sciences. *Ecological Applications*, *7*(1), 330–342.

Morgan, D. J., & Blake, S. (2006). Magmatic residence times of zoned phenocrysts: Introduction and application of the binary element diffusion modelling (BEDM) technique. *Contributions to Mineralogy and Petrology*, *151*(1), 58–70.

Newhall, C. G., Costa, F., Ratdomopurbo, A., Venezky, D. Y., Widiwijayanti, C., Win, N. T. Z., et al. (2017). WOVOdat—An online, growing library of worldwide volcanic unrest. *Journal of Volcanology and Geothermal Research*, *345*, 184–199.

Palma, J. L., Courtland, L., Charbonnier, S., Tortini, R., & Valentine, G. A. (2014). Vhub: A knowledge management system to facilitate online collaborative volcano modeling and research. *Journal of Applied Volcanology*, *3*(1), 2.

Palmer, M., Bernhardt, E. S., Chornesky, E. A., Collins, S. L., Dobson, A. P., Duke, C. S., et al. (2004). Ecological science and sustainability for a crowded planet: 21st century vision and action plan for the Ecological Society of America, 1251-1252 pp.

Parr, C., & Cummings, M. (2005). Data sharing in ecology and evolution. *Trends in Ecology & Evolution*, *20*(7), 362–363. https://doi.org/10.1016/j.tree.2005.04.023

Roche, D. G., Kruuk, L. E., Lanfear, R., & Binning, S. A. (2015). Public data archiving in ecology and evolution: How well are we doing? *PLoS Biology*, *13*(11), e1002295. https://doi.org/10.1371/journal.pbio.1002295

Sandve, G. K., Nekrutenko, A., Taylor, J., & Hovig, E. (2013). Ten simple rules for reproducible computational research. *PLoS Computational Biology*, *9*(10), e1003285. https://doi.org/10.1371/journal.pcbi.1003285

Schonwalder Angel, D. (2019). Related data for: PWD: A Petrological Workspace & Database tool manuscript, DR-NTU (Data), V1. https://doi.org/10.21979/N9/LMUUAP

Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lenhert, K., Nosek, B., et al. (2019). Make scientific data FAIR. *Nature*, *27-29*.

Vernon, R. (2004). A practical guide to rock microstructure.

Walker, J. D., Bowers, T. D., Black, R. A., Glazner, A. F., Lang Farmer, G., & Carlson, R. W. (2006). A geochemical database for western. *North American volcanic and intrusive rocks (NAVDAT)*, *397*, 61–71.

White, E., Baldridge, E., Brym, Z., Locey, K., McGlinn, D., & Supp, S. (2013). Nine simple ways to make it easier to (re)use your data. *Ideas in Ecology and Evolution*, *6*(2).

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Zimmerman, A. (2007). Not by metadata alone: The use of diverse forms of knowledge to locate data for reuse. *International Journal on Digital Libraries*, *7*(1-2), 5–16. https://doi.org/10.1007/s00799-007-0015-8