Generating metadata for an experiment: using a tablet ELN

Cerys Willoughby, Jeremy G Frey, Simon J Coles, Susanne Coles.
Chemistry, University of Southampton

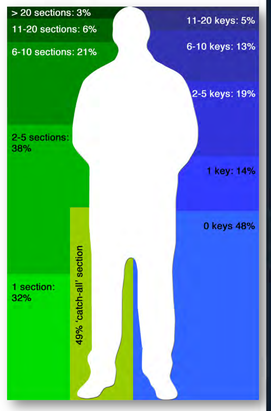
Introduction

The drive towards more transparency in research and open data increases the importance of being able to find information and make links to the data. Metadata is an essential part of this process and for the preservation of knowledge for future exploitation. Metadata is often defined as "data about data" but can better be defined as the information that describes a physical or digital document or object[1]. Metadata provides context to the data and enables relationships between different data to be explored, making the data more usable and reusable, persistent, discoverable, and accessible[2].

Metadata is used in Electronic Laboratory Notebooks to curate experiment data and associated entries with descriptive information and labels that can be used for aggregation and identification, potentially producing a much more effective record than a paper notebook[3]. Machine-generated metadata helps with facilitating metadata exchange and enabling interoperability, but is not necessarily in a form friendly for the humans that also need it.

Metadata use in LabTrove

LabTrove, a researcher-centric web- and cloud-based ELN developed at the University of Southampton, enables users to add their own user-defined metadata. LabTrove provides a flexible metadata framework, enabling the creation of two different kinds of user-defined metadata: Sections used to describe the content of the entries, and key-value pair data useful for describing specific elements of the experiment.



The findings of a survey of metadata use in LabTrove demonstrated that many users do not use metadata effectively and discussions with researchers indicate that in many cases they do not understand what it is or why it is useful.

Capturing experiment metadata

Designing a mobile ELN

Continuing from previous work on developing LabTrove and mobile ELN interfaces at the University of Southampton, we wanted to utilize the new generation of tablet interfaces to produce a flexible tool for capturing the experiment process within the lab. The tool would optionally also provide guidance to the user to help them step through their experiment and record their observations as they go. The development of the mobile ELN, called Notelus, provided the opportunity to include interfaces to capture metadata created both automatically and by the user. The metadata can then be exported to LabTrove with the experiments that the metadata describes. The primary audience for the development of Notelus was Organic Synthesis researchers, and therefore one of the key activities for identifying the user requirements was to observe the current working practices of a group of these researchers. Each of the researchers had different work practices and ways of organizing their experiment-related materials, but some commonalities between them were observed, for example the construction of identifiers used for experiments, samples, and data files; and the information used to search for experiments that had been completed in the past. Together with the common metadata labels added by LabTrove users, these observations fed into the inclusion of metadata elements in Notelus. These included options for describing the project and sub-project, including a place to record experiment conditions, and plan details such as IDs and author information.

Metadata about the experiment

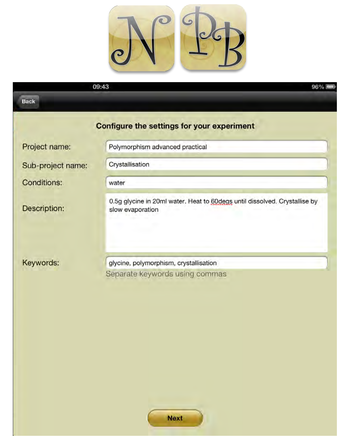
A "Settings" page is provided for each experiment when a new notebook is created that can be completed with information about project and sub-project, but also information about the conditions of the experiment (commonly searched for after an experiment was completed), and keywords so that the user can add metadata that is meaningful for them. Notelus also automatically creates values for the notebook name, the date the notebook was created, and the date it was last modified. When the notebook is exported, the date is captured as the date of the entry in LabTrove.

Creating a plan for the experiment

Part of the decision making process for the development of Notelus was defining structures to represent an experiment and an associated plan to guide the experiment. The design of the plan was based on our experiences of how researchers planned, recorded, and organized their experiment data. The plan includes information about provenance of the plan, such as the ID and Plan author, and the elements of the experiment such as the steps, materials, safety, weights, and equipment. The plan can be generated using XML or another mobile tool called Plan Buddy for Notelus and imported into Notelus. Metadata is then automatically generated based on the information provided by the user about the experiment in the plan.

Exporting to LabTrove

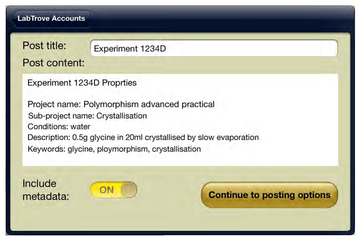
If the user chooses to export their Notelus notebook to LabTrove, they can optionally choose to include the captured metadata in their LabTrove entry.



Experiment Settings page.

```
<material>
    <material-name>Acetic Anhyride</material-name>
    <material-description>A description</material-description>
    <material-safety>Harmful on contact with skin - causes serious burns.
    </material-safety>
    <molecular-weight>102.09</molecular-weight>
    <density>1.08</density>
    <ratio>1</ratio>
    <planned-amount>3.50</planned-amount>
    <planned-amount-unit>mL</planned-amount-unit>
</material>
<material>
    <material-name>Concentrated HZSO4</material-name>
    <material-description>A description</material-description>
    <material-safety>Harmful on contact with skin - causes serious burns.
    </material-safety>
    <molecular-weight>98.10</molecular-weight>
    <density>1.84</density>
    <ratio>-</ratio>
    <planned-amount>0.20</planned-amount>
    <planned-amount-unit>mL</planned-amount-unit>
</material>
<material>
    <material-name>95% Ethanol</material-name>
    <material-description>A description</material-description>
    <material-safety>Toxic, irritating to skin and eyes, highly
        flammable, makes explosive mixtures with air.</material-safety>
    <molecular-weight>46.07</molecular-weight>
    <density>0.79</density>
    <ratio>-</ratio>
    <planned-amount>80.00</planned-amount>
    <planned-amount-unit>mL</planned-amount-unit>
</material>
<material>
    <material-name>Acetylsalicylic Acid</material-name>
    <material-description>A description</material-description>
    <material-safety>Toxic, irritating to skin and eyes.
    </material-safety>
    <molecular-weight>180.16</molecular-weight>
    <density>1.40</density>
    <ratio>1</ratio>
    <planned-amount>0.00</planned-amount>
    <planned-amount-unit>mL</planned-amount-unit>
```

An experiment plan in XML.



Exporting metadata to LabTrove.

Futures

The interface designs and the metadata captured in Notelus is just an example of what could be done, but some of the choices could be applied to LabTrove itself, by providing options to record metadata for materials, equipment, experiment IDs, and other metadata uses identified in the LabTrove study. Alternate mark-up schemes for experiment plans are also under investigation that may lead to additional ways to collect metadata for describing experiments in the lab.

Related to this work we are also investigating whether providing cues can change what metadata is recorded by users.

Conclusions

Surveying patterns of metadata use in an ELN, and the ways that researchers conduct and organize their work, has provided information that is useful for improving the capture of metadata for experiments for others. The results are metadata that represent the kinds of information that researchers use and search for, but we rarely observed in the LabTrove survey.

Literature

1. Zeng, M. and Qin, J. *Metadata* Neal-Schuman: New York. 2008 ISBN: 978-1555706357
2. Kowalczyk, S. and Shankar, K. Data sharing in the sciences. *Ann. Rev. Info. Sci. Tech.* 2011, 45: 247–294. doi: 10.1002/aris.2011.1440450113
3. Bird, C. Willoughby, C., and Frey, J. Laboratory notebooks in the digital era: Record keeping in chemical and other science laboratories *Chem. Soc. Rev.*, 2013,42, 8157-8175

Acknowledgements

Smart Research Framework, e-Research South, Dial a Molecule, UK National Crystallography Service. Organic Synthesis students at University of Southampton.