ORR12: Interview Process for Functional Requirement Gathering

# Objectives

## Primary objectives

* Gather preferred functional requirements of digital notebooks (DiNos)
* Assess demand for sample inventories
* Get first-hand account of potential benefits

## Secondary objectives

* Comprehend the interviewee’s current research data management (RDM) practice to enlighten on non-functional requirements (NFRs)
* Identify if the interviewee has previously used, or currently uses, a digital notebook or sample inventory. If so, which one?
* Assess the interviewee’s willingness for further engagement

# Method

## Primary objectives

### Functional requirements of DiNos

We want the interviewee to be guided into discussing their research data management and offering their own solutions to the problems they encounter *unprompted*. When they do so, they will largely be describing features available within digital notebooks (and sample inventories). The interviewee should therefore naturally start reeling off improvements that cover the attributes we wish to hear about. The complete list of 41 attributes of digital notebooks can be found [here](https://livemanchesterac.sharepoint.com/:x:/r/sites/UOM-RIT-RLP/Shared%20Documents/New%20RLP/Open,%20Reproducible%20and%20Responsible%20Research/ORR12%20Digital%20Notebooks%20%26%20Sample%20Inventories/02%20Definition/Functional%20Requirements/ELN_Attributes.xlsm?d=we602ba5e81fa4831a31462814116f824&csf=1&web=1&e=4mJRaQ), under the AttributeDefinition tab, whilst the reduced subset of 21 attributes which researchers are assumed to be most interested in are found [here](https://livemanchesterac.sharepoint.com/:x:/r/sites/UOM-RIT-RLP/Shared%20Documents/New%20RLP/Open,%20Reproducible%20and%20Responsible%20Research/ORR12%20Digital%20Notebooks%20%26%20Sample%20Inventories/02%20Definition/Functional%20Requirements/FRsforScoringv0.5.xlsx?d=wc819202d40ad45ea9528ebf388c66084&csf=1&web=1&e=Evz5gW), under the same tab label. Anticipated responses to each attribute are found under the ‘DefaultEntries’ tab in the same document, which you should also be familiar with. The attribute subset is also represented more visually [here](https://miro.com/app/board/uXjVLdds0oo=/) and in the table below, with a description of how the interviewee might speak around the topic.

|  |  |  |
| --- | --- | --- |
| **ID** | **Attribute** | **Description/expected content** |
| G-01 | License | An interviewee may be passionate about **open software** and/or **open-source code**, and so therefore may argue for their notebook to also carry an open software licence. Most are probably unaware. If they are however particularly enthused, they may offer up preferred open licence types. |
| U-01 | Customizable user interface | Individual operability of the DiNo can significantly increase acceptance. It specifies whether and how users can **configure** the **interface** according to their preferences. If the interviewee refers to personalisation of say, column headers in the **GUI**, this criterion is referenced. |
| U-02 | Languages | Conversation on **secondary language** availability in the **GUI** is captured here. Users can of course enter data in any language they wish, although the University would encourage English for sharing purposes, users may wish for menus etc. to be in another language, which could be facilitated. Note the language if mentioned. |
| U-03 | Support | Any reference to preferred **training** support should be captured here. |
| CF-01 | Data input | Data entry into the DiNo can be done through various **input options**. On the one hand, different sources can be supported (e.g. voice through dictation function, handwritten through digital pen, barcode scanner) and on the other hand, several input elements or editors can be made available to the user. In addition to free text notes, which are usually possible in every DiNo, input can be made e.g. via structured forms, tables or links to other objects within the DiNo (internal links). |
| CF-02 | Data import (formats) | A DiNo should be able to import and attach all **filetypes**, so this attribute should be referenced when the interviewee notes a specific format which they would wish to preview directly in the notebook. |
| CF-04 | Data export | The interviewee may note specific **export formats**, which are to be captured here. Otherwise, any generic reference to **interoperability**,exporting in **machine-readable** formats or **direct publication** from the notebook should be captured here. |
| CF-05 | Templates | Templates are used for time saving and standardization. There can be templates for documents, **protocols**, etc. In addition to the ability to create own templates, many products include pre-configured templates for specific tasks. Some DiNos also offer the possibility to import and/or share **templates**. Research supporting colleagues (Technical or otherwise) typically have large interest here. |
| CF-06 | Searchability | One of the most significant advantages of electronic lab books over paper books is the ability to find data quickly and easily. In addition to the simple full text search, most products offer an **advanced search**. Moreover, many tools offer additional possibilities, these are e.g. search for chemical formulas, for biological content, database queries, **filtering** possibilities with marking of filtered records. BLAST (Basic Local Alignment Search Tool) is relevant to DNA sequencing. Reference to **file hierarchies** and **tags** should also be included here. |
| CF-07 | Collaboration | Collaborative functions, such as **sharing** documents, data, and workflows, are a key strength of electronic lab notebooks. However, not all users need (and should have) the same access rights. Rights and role management helps to define the access right in different granularity, sometimes down to the user- and file level. **Rights management** is used when access **permissions** of individual or grouped objects can be adjusted, whilst **role management** refers to defining a user or group’s access rights/data visibility. Mention of **real-time collaboration**, **shared project management tools** including **calendars** and **task assignment** with **notifications** are also captured here, as well as **external guest access.** |
| C-01 | Preservation of evidence | The sustainable and secure storage of research data is both required by law in specific branches of research and generally required by funding agencies and research institutions. Specific regulatory codes are captured in the next criterion. Here, specific features which help preserve evidence are captured, i.e. **audit trails**, **signatures**, **timestamping** and **version control.** |
| C-02 | Compliance | Any reference to specific security or integrity standards (see above) should be captured here. |
| EF-01 | Laboratory management functions | Laboratory management functions can be a helpful extension of the DiNo, but it may also be important to allow existing systems to interoperate with the DiNo if (partial) management, e.g., a chemical database, is already in place. The functions mentioned here are often available as an additional module. '**Materials database**' is the broader definition of chemical inventory. '**LIMS Connectivity**' is used to define when integrated LIMS software is available, or integrations possible. '**Sample tracking**' is used to indicate demand for a **sample inventory**. |
| EF-02 | Integrations and extensions | Integrations and extensions provide additional functionalities to a DiNo, e.g. for transferring data to external services or programs. Any mention of **additional software** integrations should be included here (beyond LIMS). |
| EF-03 | Automation | Any automation capabilities are listed here, including functions which can be accessed using an **API**, including "**Device control**" - the transmission of signals/commands to external devices - "**Data analysis**" - for at least preliminary plotting and data manipulation/ visualisation within the notebook (or via an integration). |
| EF-05 | (Individual) Project management tools | Project management tools referred to here are non-collaborative, as distinct from CF-07, such as the desire for a **personal calendar** or **task-board**. |
| EF-06 | Workflows | The term **workflow** here refers to the ability to capture both workflows in the laboratory and technical sense, i.e. for capturing data/decision flow of a DiNo, both internally and external. The documentation and reuse of workflows, SOPs (self-defined standard operating procedures) and processes can bring about a significant increase in efficiency. Traceability of workflows contributes to transparency and can be a means of transferring expertise. It can be worthwhile to develop a concept for the creation and use of workflow templates to derive the maximum benefit from the use of a DiNo. |
| IT-01 | Application programming interfaces | An API (**application programming interface**) is an interface for application programming. It can be used, for example, to integrate software into another system or to transfer data to/from a device. It is important that an API is sufficiently described. If a specific **programmatic interaction** with the notebook is mentioned, that should be captured here, including demand for **SDK (Software Development Kits**), which accelerate software add-on development, and **ODBC (Open Database Connectivity)** database interfacing. |
| IT-02 | Controlled vocabulary | Controlled vocabularies (**ontologies**)are a set of predefined expressions/words that can be provided as input. They improve the coherence of records by preventing different versions of the same term (e.g. typos, plural vs. singular). Using published **controlled vocabularies** (or ontologies) is an important step to increase interoperability and comparability of records between groups. This criterion should be noted only if strict reference to controlling/standardising input is made. |
| IT-03 | Data access (client) | The vast majority of DiNos are **browser-based** and are therefore accessible through any computer operating system. If the interviewee expresses a desire to access the platform via a browser, **tablet/mobile** compatibility(i.e. Responsive Design), **mobile app** or a **desktop client**, this would need to be specified here. |
| IT-04 | Data storage location | The location of storage must be compliant with the privacy policy of the using organization. Depending on the selection, full control over the data and the security protocols lies with the administration of the organization using the DiNo. In the case of cloud storage, the physical location of the servers may determine whether compliance with certain data protection policies may be expected. Despite ITS holding the final decision, end-users may have a preference on whether their data is held on their **own cloud storage**, **locally** (on the Research Data Service – **RDS**) or on the **provider’s cloud** (**SaaS**) – this is captured here. |

As per the first objective, the goal is therefore to corroborate the interview against these criteria, much in the same way that [the survey](https://www.qualtrics.manchester.ac.uk/jfe/preview/previewId/51afd0fa-3f2c-4112-bd58-0da8ccdb9607/SV_8vTzwbvdC97UZbU?Q_CHL=preview) does, to facilitate scoring the active products available on the market. I think the easiest way to do this is to have an awareness of the criteria and to either note them down as they come up (in an offline interview) or perhaps later on from the interview transcription (if permission to record is given in an online interview). Once the interview is complete, each functionality mentioned can be recorded against the attribute in the [stakeholder response log](https://livemanchesterac.sharepoint.com/:x:/r/sites/UOM-RIT-RLP/Shared%20Documents/New%20RLP/Open,%20Reproducible%20and%20Responsible%20Research/ORR12%20Digital%20Notebooks%20%26%20Sample%20Inventories/02%20Definition/Functional%20Requirements/Stakeholders_FRs_Bens.xlsm?d=w9d0e97af6a054764832f14b4d0730ed5&csf=1&web=1&e=veDSwD): ‘StakeholderELN\_FRs’ tab. I think it would be too difficult to try and simultaneously populate the log whilst actively listening to the interviewee.

### Sample inventory demand

Focus is on the digital notebooks at this stage, as there seems to be little difference in the functionality of sample inventories across platforms. The demand for sample inventories is instead captured in EF-01 > Lab Management Functions > Sample Tracking. If however, the interviewee goes into detail on Sample Inventories, this can be captured in an additional tab of the [stakeholder response log](https://livemanchesterac.sharepoint.com/:x:/r/sites/UOM-RIT-RLP/Shared%20Documents/New%20RLP/Open,%20Reproducible%20and%20Responsible%20Research/ORR12%20Digital%20Notebooks%20%26%20Sample%20Inventories/02%20Definition/Functional%20Requirements/Stakeholders_FRs_Bens.xlsm?d=w9d0e97af6a054764832f14b4d0730ed5&csf=1&web=1&e=veDSwD): ‘StakeholderInventory\_FRs’. The attributes (column headers beyond ‘User’) have been added on an ad-hoc basis, so these can be adjusted if necessary. I’m not anticipating any notable discussion on inventories beyond the (shared) DiNo attributes however, within the limited timeframe of the interview.

### Benefits logging

The final tab in the [stakeholder response log](https://livemanchesterac.sharepoint.com/:x:/r/sites/UOM-RIT-RLP/Shared%20Documents/New%20RLP/Open,%20Reproducible%20and%20Responsible%20Research/ORR12%20Digital%20Notebooks%20%26%20Sample%20Inventories/02%20Definition/Functional%20Requirements/Stakeholders_FRs_Bens.xlsm?d=w9d0e97af6a054764832f14b4d0730ed5&csf=1&web=1&e=veDSwD), ‘StakeholderBens’, is designed to capture where any benefits of DiNo + Sample Inventory uptake can bring benefits, which are categorised by each benefit category. Entries in cells are verbose free text, as I couldn’t think of a better way to capture this. Therefore, during the interview, if someone mentions either a benefit or disbenefit, it can be jotted down and placed under whichever header (benefit type) best suits it.

## Secondary objectives

### RDM for NFRs

In previous discussions, discussing a researcher’s current workflow has brought out additional considerations around their research data management. It’s helpful to hear these points to consider the wider picture and any project dependencies. Discussing their current RDM practice can also highlight issues with data security, retention etc., which bolsters the case for DiNos and may inform NFR considerations also. As NFRs will be largely determined by a business analyst (BA), this isn’t of particular concern, hence the secondary objective. Although, it is helpful to build an awareness of NFRs to ensure the BA captures everything. Therefore, feeding this information back to the PM is valuable.

### Previous or current DiNO experience

Simply, in both FR-capturing tabs, there is a column where any current or previously used DiNo or sample inventory can be noted. Doing so will help us identify ‘power-users’.

### Ongoing engagement

Finally, if the interviewee appears either:

* Keen to engage further
* Particularly helpful on the topic
* To have a valuable sphere of influence

See if they will be willing to share their contact details so we can reach out to them specifically in the future.