

Imtiaz A.Khan^{1,2}, Adam Fraser¹, Mark-Anthony Bray¹, Victoria Griesdoorn², Paul J. Smith², Anne Carpenter¹, Rachel J. Errington²

¹Imaging Platform, Broad Institute of Harvard and MIT, Cambridge, USA
²School of Medicine, Cardiff University, Cardiff, UK

We don't OR can't share experimental data!

Problem: Despite the exponential growth of data management systems, it has been shown that more than 75% of experimentalist rarely or half the time use publicly available data.

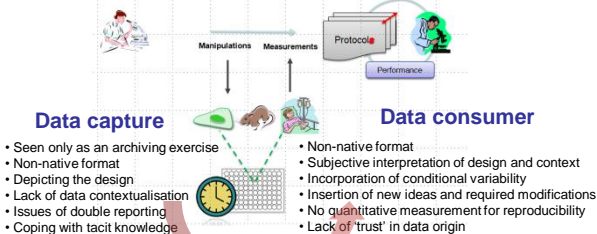
Opportunity: The intricate nature of experimental methods as well as the demand for meaningful communication and effective collaboration about an experimental methodology and data presents a major challenge.

Going beyond just metadata – contextualization via design

Problem: Current form filling or workflow building approaches has improved the *human-computer* interactions, but for effective data sharing, *human-human* interaction is a fundamental requirement. In this regard, metadata and data needs to be presented in context of the experimental design.

Opportunity: Depicting and specifying the design of an experiment is not only technological challenge, but a social one. Since, it is the practice and perception of the research process as well as barriers presented by researchers that imposes the bottleneck.

Defining the 'human factor' problem - social barriers that can lead to poor community compliance and very little opportunity for standardisation



Loss of data provenance = Loss of data 'scientific value'

ProtocolNavigator - a virtual laboratory environment

Here we present **ProtocolNavigator**, a software package that enables experimentalists to emulate their real-laboratory practice in a virtual environment. Focusing on cell biology research and the requirements for systematic cell-based measurement, ProtocolNavigator has implemented an action-based protocol creation approach consisting of three interactive panels with interlinked functionality and display.

The Catalogue panel, an inventory of elements for use in the experiment is presented, including biological/chemical materials, analytical instruments, and sub-processes (e.g. staining process); users can create multiple instances of an element when required. This inventory can be reused for multiple experiments with or without additional modifications, as would be the case in actual practice.

The Bench panel, represents a virtual workbench underpinned by an informatics matrix that requires three fundamental inputs, simply described as "what", "where" and "when". It is here that rather than filling a form the experimentalist undertakes various "actions" by applying previously created instances ("what") to different experimental samples ("where") at certain time points ("when").

The Visualization panel, a time-integrated workflow or map is automatically generated in response to the actions in the Bench that represents the spatiotemporal complexity of the methodological design. This interactive map also serves as the access point for retrieving both action details and experiment-derived data.

Each track through the map is the sequence of actions applied to a set of samples, and thus provides a trajectory or provenance for the data. By encapsulating metadata, provenance and experiment derived data within a visualization framework, we have established a non-textual-based contextualization of experiment-derived data.

Time-stamping provides the structure for the visualisation map, for classical textual based reporting – we have adopted the Google Map ethos, where time-stamped sequential description of steps along with a localised visual representation of the location are provided. We are further aiming to develop the capacity to extend this output format to any journal-specific functionality (like EndNote), where users can generate different format of material and method section from the same protocol.

Outcomes – the highlights

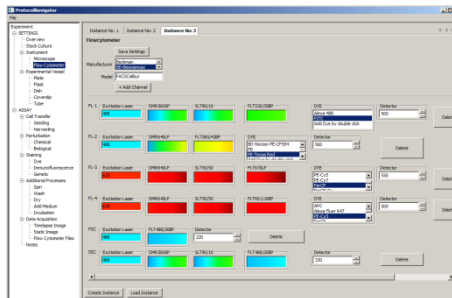
- ❖ **Experimental design** - (at start) is the basis for documentation not archive (at end).
- ❖ **Reuse of inventory** - defines the 'home' laboratory and maximizes reproducibility.
- ❖ **Emulation** of real life practice to increase accuracy and encourage user compliance.
- ❖ **Data embedded** directly in the context of experimental design with traceability
 - enables unambiguous interpretation,
 - transfer and translation of data with provenance,
 - better knowledge sharing, specially in multidisciplinary environment.
- ❖ **The map** as a reflection of the practice can identify "best practice".
- ❖ **Minimize or even eliminate "double reporting"** - reporting of protocols in standardize fashion

Some key references

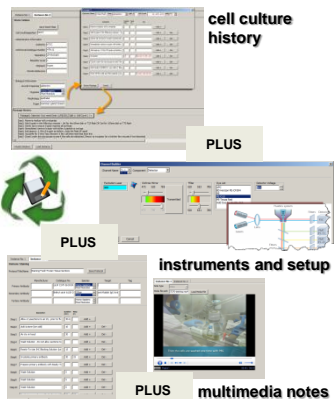
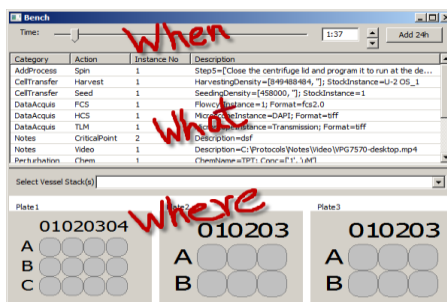
Millard et al. *Nature Methods* 8(6): 2011
 Editorial Science 331, 692 (2011)
 Tenopir et al. *PLoS ONE* 6(6): e21101 (2011)
 King et al. *J R Soc Interface*. 2011



The catalogue panel – defining the laboratory



The Bench panel – defining the user actions



supplementary protocols

at the START - The experimentalist creates/loads an element or instance in the Catalogue panel

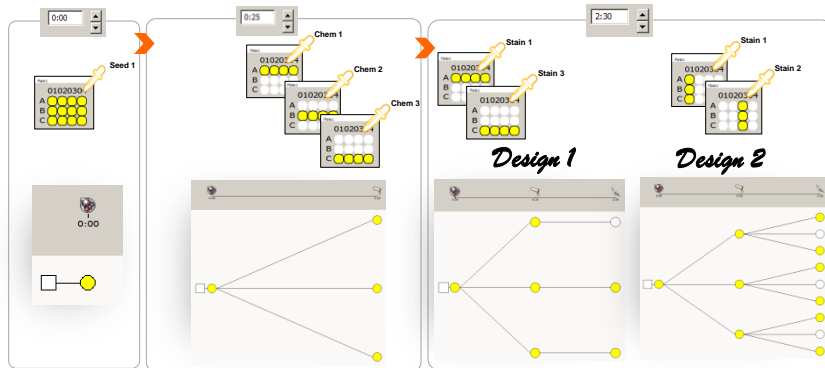
WHAT - It is now available in the Bench panel (under "What" list) for use. This resembles the "shelf" of a laboratory packed with different materials to be used in an experiment.

WHERE - The experimentalist undertake actions according to a timeline.

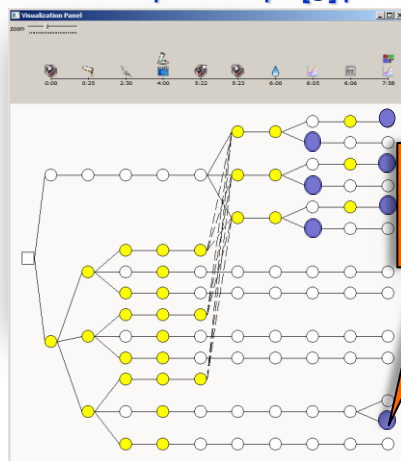
WHEN - These actions take place in different vessels

A map emerges of all these actions

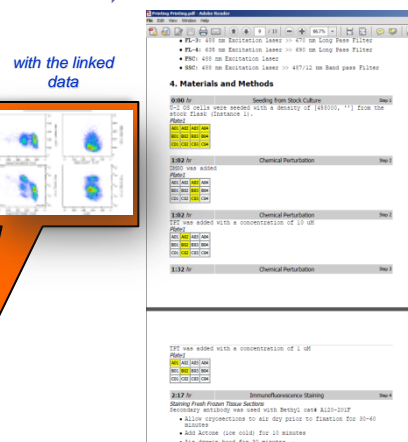
Emulation of the 'real-laboratory' practice leads to the practice map



Practice map with sample [●] provenance



Print this page



Working version of ProtocolNavigator can be downloaded from:
<https://sites.google.com/a/broadinstitute.org/protocolnavigator/>