ISSN 2398-6182
Received on 26th September 2017
Accepted on 30th October 2017
doi: 10.1049/enb.2017.0019
www.ietdl.org

Comprehensive web-based broker for bio-technology design and manufacturing

Stefan Claesen[1] ✉, *Anna Stone*[1], *Mark van Rossum*[1], *Richard I. Kitney*[2]

[1]*Visbion Ltd, Visbion House, Gogmore Lane, Chertsey, Surrey KT16 9AP, UK*
[2]*Department of Bioengineering, Imperial College, London SW7 2AZ, UK*
✉ *E-mail: info@visbion.com*

Abstract: Synthetic biology, particularly in relation to characterisation experiments relating to the description of bio-parts frequently involves the use of a wide range of equipment, including, for example, plate reader's, flow cytometers, and mass spectrometers. This equipment is often from multiple manufacturers. The study describes broker technology that has been developed which has the ability to connect multiple types of equipment into a common information environment; the connectivity from the databases and equipment is achieved using Visbion's 'cube' technology that involves military specification encryption for data security. The broker technology uses a new, developing standard, Digital Imaging and Communication in Medicine (DICOM)-SB, that is based on the highly successful international standard for biomedicine, DICOM. The broker uses a version of the DICOM data model that has been specifically designed for synthetic biology and, in particular, characterisation data.

1 Introduction

At its core, synthetic biology (or engineering biology) applies engineering principles to the rational design of biological parts, devices and systems. This has led to expansion of the synthetic biology toolkit and the development of repositories and computer-aided design (CAD) tools that assist in the design, building and testing of these systems [1]. An increase in the number and variety of software resources available for a broad range of synthetic biology activities marks the transition of the field from one of research to that of an engineering discipline. However, the problem with the large number of data sources comes when users wish to access multiple repositories, devices and tools in order to undertake research and development because different equipment and systems do not always have the necessary level of interoperability. There is therefore a need to develop solutions that facilitate more efficient data communication and sharing. For example, by increasing the efficiency of synthetic biology design and experimentation, the speed with which the industry in the field can develop commercially on a global scale will be increased. Visbion has developed such a system, building on and adapting solutions used for information sharing in the healthcare industry.

2 Digital imaging and communication in medicine (DICOM) standard

In the field of medical imaging, the age of digital radiography brought with it a host of systems for the acquisition of both the image and associated data. It was quickly established that there was a requirement for these different systems to be more interoperable and an important way in which this was achieved was through the introduction of the DICOM® standard.

The 1990s saw the development of two compatible international standards for biomedicine – DICOM [2] and HL7 (Health Level 7) [3] – which are now used universally in healthcare itself and the vast majority of international healthcare technology companies. A key goal of DICOM was to achieve compatibility between different medical imaging systems and information systems in the healthcare environment. Originally, the DICOM standard defined the data formats and communication protocols for medical images. Over the last 15 years DICOM has developed into a much more comprehensive standard, covering broader areas of biomedicine. DICOM is based on ISO guidelines [4]. Currently, there are no established standards in terms of engineered biological function that would be recognised by the International Standards Organisation (ISO) [4]. The development of standards for synthetic biology is now underway. There are currently two compatible standards under development – DICOM-SB and synthetic biology open language (SBOL). DICOM-SB is an important aspect of this paper and will be covered in more detail. SBOL [5] is, essentially, a standard for the exchange of information relating to DNA and, in addition, describes DNA components used in synthetic biology.

DICOM has now formed the basis of a new standard for synthetic biology data – termed DICOM-SB [6]. This new standard allows for the storage and display of raw experimental data alongside protocol and other information. One important application is in bio-part characterisation experiments [6]. The DICOM-SB standard is designed to be compatible with other file formats and in synthetic biology, it can be seen as being complimentary to the SBOL standard that is used for the description of DNA components. DICOM-SB is currently undergoing the formal approval processes with the International DICOM Committee. Systematic design in synthetic biology is based on the engineering principles of modularity, characterisation and standardisation. In synthetic biology for parts, devices and systems to be accurately reproduced they must be standardised. This is particularly important in relation to the translation of research to industrial products. As the field develops, synthetic biology designs and processes will need to interface with existing, complex multistep manufacturing processes.

One of the main advantages of the DICOM approach is it allows for the storage of raw data alongside the meta-data that accompanies it. (Traditionally this might be the raw output from an imaging modality, alongside the patient demographics.) In synthetic biology, the process of characterisation involves the need to run multiple experiments, with various outputs (this is equivalent to the imaging data in healthcare). In synthetic biology it is the raw experimental data (e.g. the output from a flow cytometer). For example, in order to document a characterisation experiment effectively, the parameters and layers of analysis should also be included – these become the meta-data included in the DICOM-SB file.

3 Broker features and capabilities

Having initially been involved with the DICOM-SB proposal, Visbion has taken that work forward and applied the new standard to the development of a software broker for use in the synthetic biology field. The broker addresses the lack of interoperability between the varying data sets that are available to the synthetic biology user. The broker software is designed to sit between the multitude of bio-CAD tools and repositories of bio-part data currently available. Where, currently, the search and retrieval of information from these applications would involve the user moving between differing platforms and file types, the aim of the broker was to allow this search process to be performed on a single platform, keeping everything in one place and improving workflow efficiency (Fig. 1).

Visbion supplies technology to the healthcare industry that is based on the DICOM standard. The technology is designed to assist in the storage and communication of medical images and associated data. The technology that the broker for synthetic biology has been built upon has drawn upon this established Visbion medical imaging platform. DICOM-SB forms one of the core components of the broker, which also utilises a set of

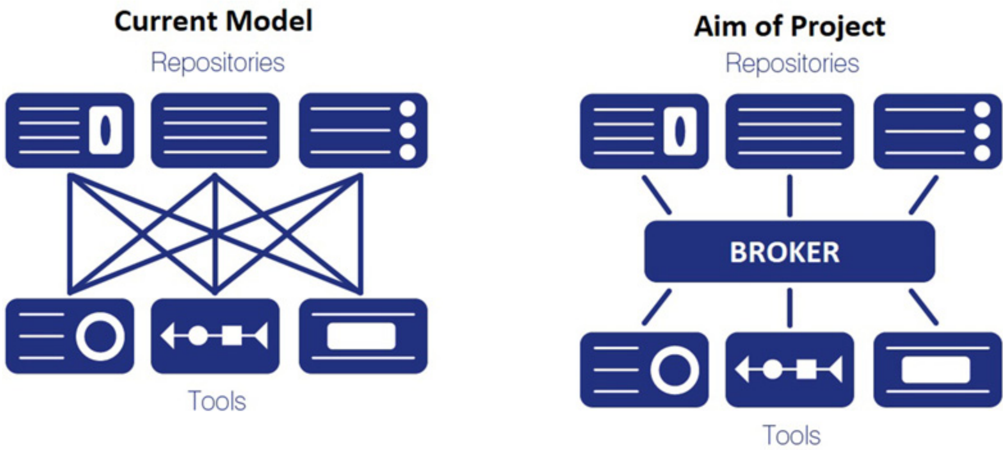

Fig. 1 *Broker is designed to ease the communication of data between differing bio-CAD tools and repositories*

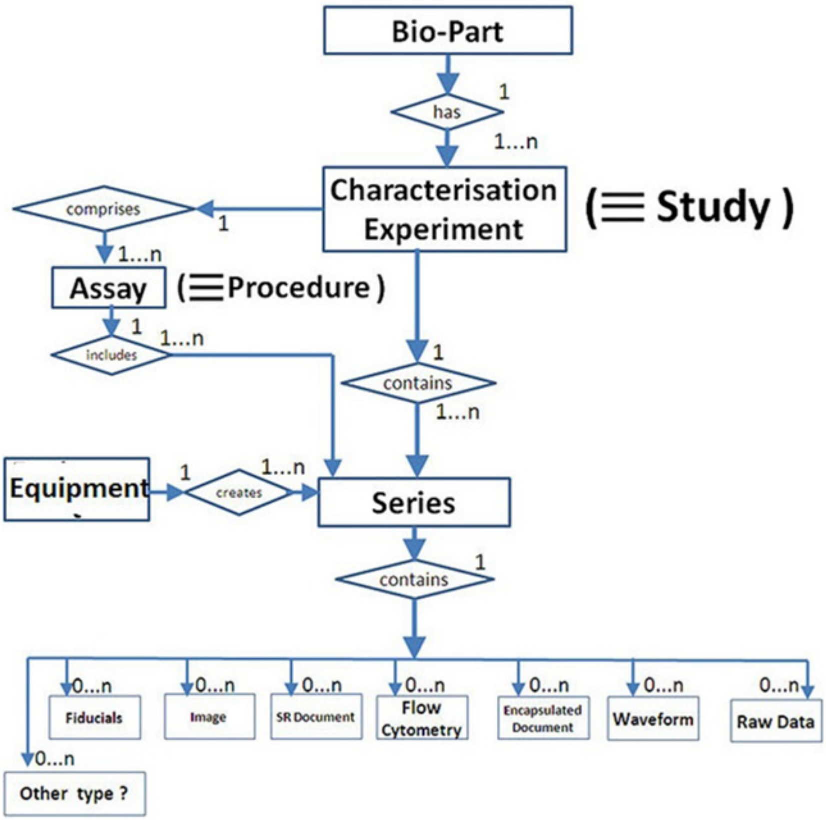

Fig. 2 *Basic DICOM-SB data model*

conversion tools to communicate data in different formats. In addition, the broker technology that has been developed has the ability to connect multiple types of equipment (e.g. plate reader's, flow cytometers, and mass spectrometers) from multiple manufacturers to be connected into a common information environment; the connectivity from the databases and equipment is achieved using Visbion's 'cube' technology that involves military specification encryption for data security.

4 DICOM-SB data model

Fig. 2 illustrates the basic DICOM-SB data model, which has been modified from the original DICOM data model, for synthetic biology. Referring to the figure, the basic tree structure of the data model is translated to Family Type (e.g. constitutive promoters), Parts, Characteristics and Experiments. It is important to note that this is a simple illustrative example and that the data model can be extended, as necessary, to incorporate additional information, such as metadata. For example, BioPart source, design notes, vector resistance, protocols, institution name, date and authors etc. The broker is designed to incorporate the DICOM-SB data model, particularly in relation to characterisation experiment data, metadata and anecdotal notes.

5 Development work

The broker software was developed over a number of stages. Once the DICOM-SB standard had been outlined and documented, there was a requirement to build a number of converters. These were mostly focussed on the conversion of XML to DICOM-SB and back again. These converters were then built into a framework that would form the basis of the broker and provide an overall tool for the performing of query/retrieve actions and integration into live tools and registries.

6 Future work

Alongside the development of the prototype, Visbion have identified a requirement in synthetic biology to store the large amounts of data deriving from synthetic biology experiments. The next stage of work therefore proposes that the broker infrastructure is connected to Visbion's large-scale vendor neutral archive technology that can handle very large amounts of data for storage and subsequent analysis. By combining these different systems and creating more efficient workflows and data management, there is great potential for establishments such as DNA Foundries to greatly improve their internal infrastructure and move towards more scalable methodologies.

7 Acknowledgement

The work described in this project was partly funded by Innovate UK under their Smart Grant initiative.

8 References

1 Kitney, R.I., Freemont, P.: 'Synthetic biology – the state of play', *FEBS Lett.*, 2012, **6**, pp. 2029–2036
2 NEMA PS3/ISO 12052. Digital Imaging and Communications in Medicine (DICOM) Standard, National Electrical Manufacturers Association, Rosslyn, VA, USA. Available at http://medical.nema.org/
3 Dolin, R.H., Alschuler, L., Beebe, C., *et al.*: 'The HL7 clinical document architecture', *J. Am. Med. Inform. Assoc.*, 2001, **8**, (6), pp. 552–569
4 ISO. International Organization for Standardization, nd. Available at http://www.iso.org/iso/home.html, accessed September 2017
5 Galdzicki, M., Wilson, M.L., Rodriguez, C.A., *et al.*: 'SBOL: a community standard for communicating designs in synthetic biology'. DSpace@MIT: 2013 2013/08/02. Report no
6 Sainz de Murieta, I., Butelle, M., Kitney, R.I.: 'Towards the first data acquisition standard in synthetic biology', *ACS Synth. Biol.*, 2016, **5**, pp. 817–826