

Plans and Performances: Parallels in the Production of Science and Music

David De Roure, Graham Klyne,
Kevin R. Page, John Pybus, David M. Weigl
Oxford e-Research Centre
University of Oxford
Oxford, UK
david.deroure@oerc.ox.ac.uk

Matthew Wilcoxson, Pip Willcox
Bodleian Libraries and
Oxford e-Research Centre
University of Oxford
Oxford, UK

Abstract— Whether in the science lab or the music studio, we go in with a plan, we perform, and we make a record of that performance for distribution, consumption, and reuse. Both domains are increasingly data-intensive, with the adoption of new technology, and also socially intensive with democratised and growing citizen engagement. The music industry has embraced digital technology throughout the lifecycle from composition to consumption; scientific practice, and scholarly communication, are also undergoing transformation. Is the music industry more digital than science? We suggest that comparing and contrasting these two systems will provide insights of mutual benefit. Our investigation explores the notion of the Digital Music Object, analogous to the Research Object, for rich capture, sharing and reuse of both process and content.

Keywords—*Digital Music Object; Reproducible Research; Research Object; provenance; scholarly communication; workflow*

I. INTRODUCTION

The music industry has ‘gone digital’ almost end to end, from composition and recording, through production and distribution, to consumption via downloads and devices, and in some cases to reuse by professional and amateur alike. This is not just the course of the audio recording itself, but also of the associated information, from musical scores to front-of-house materials, and social media in promotion and in response to the performance. This transformation has not been without challenge, notably how established intermediaries endeavour to sustain business when digital copying and transmission can occur at near-zero cost, and the digital technology enjoys widespread and pervasive adoption.

Research, and research communication, are also undergoing a digital shift, as we conduct research at new

computational and social scales, and our academic papers are born and consumed digitally. The academic publishing industry debates its future while acting to sustain business. The peer-reviewed scientific article, a format now over 350 years old, mirrors the audio recording in moving from physical artefact to digital download. Libraries, which traditionally created and curated physical collections and facilitated their use, have additionally become distributors and curators of born-digital content.

What can science learn from the music industry, and *vice versa*? In this paper we provide a report on this investigation, as we develop a series of tools which bridge these two endeavours. We open in Section II by looking at the common backdrop, as expert and citizen alike adopt new technologies. Section III focuses on description of process, as this underlies the increasingly automated production methods in both domains. The focus switches from process to object in Section IV, as we compare the lifecycles of article and music. Based on this we discuss Digital Music Objects (DMOs) in Section V. Some of our tools and scenarios are described in Section VI, highlighting the intersections, as these are developed to inform and ultimately to realise the DMOs. We draw conclusions and identify future directions of work in Section VII.

II. AFFORDANCES OF THE DIGITAL

The ‘turn’ of data-intensive science is characterised by new computational capability and by working with the capture, analysis, and curation of large datasets [1]. This is realised as we adopt the increasing capacities of our computational, storage, and network hardware in the pursuit of accelerated research, by innovating in data-intensive methods, and establishing new research practice.

The increasing utility and affordability of the hardware has also put it in the hands of citizens. The result is massive growth of citizen engagement in the digital world, thanks to the Web and the smartphone, and with it access to digital cultural goods. Naughton [2] observes the convergence and blurring of boundaries that is occurring through digital technology because artefacts are reduced to the lowest common denominator—the bitstream.

This work is supported by: *Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption* funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/L019981/1, a collaboration between Queen Mary University of London, University of Nottingham and University of Oxford; *Transforming Musicology*, funded by the UK Arts and Humanities Research Council under grant AH/L006820/1 in the Digital Transformations programme, in collaboration with Goldsmiths University of London.

III. PROCESS DESCRIPTIONS

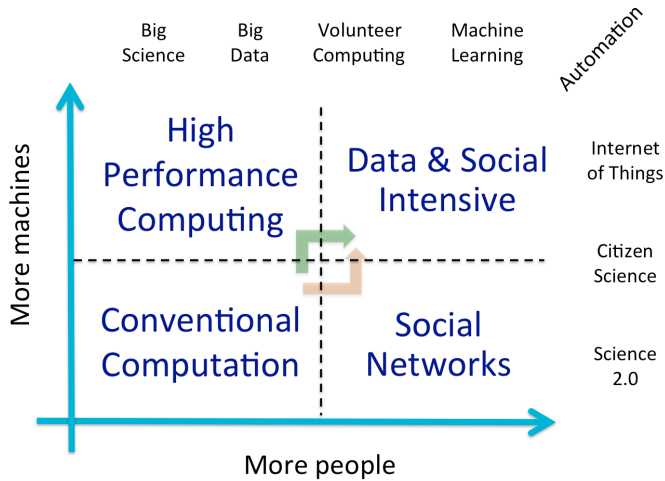


Fig. 1. Computational capability and democratization come together in the fourth quadrant, with increasing automation ahead.

This democratisation is a key digital affordance in our discussion: it means science is not just data-intensive but “social-intensive” too. Content is routinely born digital in lab or home, by professional and amateur. Significantly reuse, repurposing, and republishing is also democratized—empowering the consumer as producer. Distributed teams of empowered individuals form fluidly around content, and new social processes are easily created.

e-Science (and e-Research) today is innovating at the intersection of these trends; for example, massive datasets are generated by online interactions, while other datasets are analysed by citizen science. e-Science has followed a trajectory into high performance and high throughput computing and towards the social, while Science 2.0 and citizen science build on mass citizen engagement and become more computational. This is depicted in Figure 1.

With this comes another key trend and affordance of digital: the rise of automation. We handle the scale, and/or the velocity, of the massive datasets by automating the analytic process. This batch processing is becoming smarter: today citizen science systems use smart algorithms for task allocation, and next we might anticipate machine learning capturing some of the human analytic practices in order to scale to the data volumes of tomorrow’s science.

In this study we aim to concentrate consideration on the *process* as well as data, especially as this underpins the established scientific method—it is how we ‘stand on the shoulders of giants’. We note a significant growth in attention to research data, as stakeholders come to terms with data management plans and data sharing policies and as citizen concern grows for the responsible use of data, but with this comes at least an equal rise in the methods and process of data handling, processing, and analytics. We suggest that issues of sharing and responsible use of process are at least as important as those of data, evidenced through prior work in sharing scientific workflows [3].

A. Descriptions in Laboratory and Studio

A program describes the process that the computer is to execute; we can view the execution as the *performance* of that process. Similarly a computational workflow describes a process and is ‘enacted’. Many decisions are made at runtime as the process is accomplished: which processor core, where to allocate memory and disk, which remote service to use. Exception handling may occur due to events that arise dynamically during execution. While the exceptional decisions may be logged, the routine execution is not usually recorded unless we are debugging. More recently the importance of such provenance traces is recognised in reproducible research [4].

Compare this with a play, musical work, or scientific experiment being performed, from script, score, or lab protocol. In this case the execution is enacted by humans, and unexpected events may be handled by agreed procedures, or creative problem-solving solutions.

The scientist in the lab is aided today by digital capture of data from lab equipment, be it a balance or a mass spectrometer. Logging work via laboratory notebooks is mandatory practice, and increasingly these are electronic lab notebooks (ELNs). Hence the scientists go into a lab with an experiment plan, perform the experiment, and the output of that experiment consists of a diverse array of data, some direct and some contextual, some automated and some reported by human. Crucially, essential problem solving may occur during the experiment. The rich information capture is needed to interpret results and ultimately to support reproducible research. The notion of capture in the lab to facilitate interpretation has previously been articulated in the chemistry domain as *publication at source* [5], even including contextual information such as the status of air conditioning and door interlock, and semantic tooling has been developed [6].

Compare this with the music studio or live recording, which provide a rich digital record of a performance, with high quality audio capture but also the records of recording setup, mixer and virtual instrument parameters and perhaps the self-reported notes of the studio engineer. Musicians are also using digital equipment, their playing supported by scores, notes, and annotations, perhaps using samples and sequences. The compositional phase before the studio, like experiment planning, might also be digitally assisted.

B. Data-Intensive Science

Having established a correspondence between the studio and the laboratory, we might reasonably ask: what about data-intensive science—is there data-intensive music? Digital music datasets are large and created all the time, for example in every performance recorded ‘off the desk’ and uploaded to large curated corpuses such as the Live Music Archive (etree) in the Internet Archive. Multichannel high quality audio capture rates have kept pace so that recording continues to challenge local storage, network, and processing capacity, demanding gigabit per second live data rates. The music studio is comparable to high end laboratory apparatus, with real-time processing.

There are computational challenges in processing the digital audio, especially in real time, and also in analysis and discovery. Techniques established by the vibrant Music Information Retrieval (MIR) community [7] have utility throughout the recording, production, distribution, reuse, archiving, and curation phases of the digital music lifecycle. Feature extraction tools are increasingly set to feature in the toolkit for audio capture and reuse. As in science, the MIR community has developed computational workflow solutions, and a workflow paradigm is proving useful [8].

C. Provenance

In all these examples we can identify (a) a *prospective* plan or script, and (b) a *retrospective* record of what actually happens. These descriptions vary in completeness and specificity. At the comprehensive end of the spectrum, a computational workflow provides a detailed plan, and its provenance trace reports the runtime decisions (e.g. which service was selected?) At the underspecified end, we might have found a few enduring parts of a historical record and piecing it together is itself a process of scholarship.

This is illustrated in Figure 2, where the vertical axis represents decreasing determinism or increasing creativity. We note that the introduction of digital methods brings a potential improvement in completeness of the process record, be it in digital archaeology or digital studios.

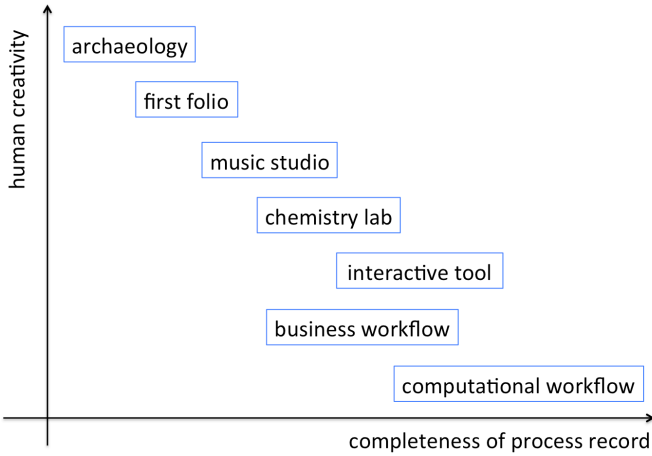


Fig. 2. Illustrating a variety of process descriptions, with more detailed descriptions currently available when machine-generated.

In many cases today, comprehensive descriptions of provenance occur when we are working largely by machine: we see provenance information being ‘born-digital’ alongside content. However it is possible to capture provenance information when humans are involved, such as in a business workflow or a citizen science experiment. The standards and tooling for this exist, notably the W3C PROV recommendations which are motivated by examples of digital systems with human agents [9], as long as the provenance information is substantially complete. We can anticipate further application of PROV in this space, partly motivated by the push to facilitate reproducible research.

IV. LIFECYCLES IN THE ECOSYSTEM

The outcomes of the processes we have described are pieces of recorded music and research results. At first glance, these may seem quite distinct products, that we obtain and use in quite different places for contrasting purposes, with differing scale and demographic of audience. But they are both products of a publishing process. Indeed, before audio distribution at scale was possible, music was distributed as scores—they share the early history of print, and distribution of printed materials.

With product in mind, tangible or digital, one perspective we might usefully take is that of the consumer. Figure 3 is based on the customer activity model of Vandermerwe [10], which we have extended beyond its adaptation presented in [11]. This captures the phases of discovery, use, and post-use from the perspective of the consumer. It enables us to see many similarities, and to tease out differences. For example, discovery, acquisition, and storing in a personal library (stages 1 through 6) have clear commonality. One observed difference might be that citation (5) does not have a direct analogy in music listening from a user viewpoint—however we would argue that unique references to recordings are as much a means to reputation and income in music as they are in scholarship.

Here we are particularly concerned with the way in which these sociotechnical systems have evolved, or indeed reconstructed, due to the affordances of the digital technology. Perhaps the most obvious affordance, and driver of change, is the reduced cost of replication and distribution, where vinyl LP or paper article are replaced by digital download. At the same time this ease of copying facilitates redistribution, a benefit for open content where reuse is encouraged, but a challenge for copyright material such that the industry has sought digital solutions for rights management. Hence stage 8 of the cycle in Figure 3 now demands particular attention due to the ease of reuse, repurposing, and republishing of the digital product.

This product perspective reminds us that the output, be it musical or scholarly content, is also a social object—by which we mean it is the subject of sharing and discourse, around which social networks form and reputation is built. While the CD is a physical object for digital content, today’s social objects can be entirely digital, easily shared and annotated, and they interlink social networks.

We turn then to the lifecycle of the product, with its digital origins: studio practice has changed profoundly due to digital technology [12], and so too has science, for example in chemistry research [13]. We can consider the research cycle diagram in Figure 4, due to Neylon [14], in both contexts.

What are the differences? We might highlight reproducibility, fundamental in the established scientific method, whereby another scientist in another lab should be able to reproduce an experiment *independently* based on the record of this one. At first glance, the recording industry does not rely on such a philosophy. However, we note (a) performers typically wish to replay performances, especially in live settings, and (b) reuse and reinterpretation of audio content underpins the productivity of the industry. Hence the incentive for better capture could be to maximise the reuse of the audio content. We address this in the next section.

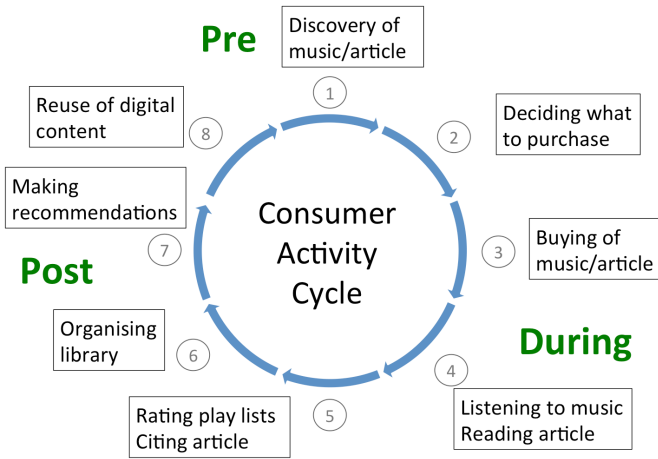


Fig. 3. The Customer Activity Cycle (due to [10]) adapted for music and scholarly articles.

V. DIGITAL MUSIC OBJECTS

The object flowing through the production process in music, from recording through production, distribution, and consumption, was originally a physical object containing an encoding of the audio signal—be it shellac, tape, or vinyl. Now it is the digital audio file. As with tape, the studio generates a multitrack recording (known as *stems*, which are subgroups of mixer channels) and this can be used as the basis for multiple mixes. The consumer receives an audio file, in WAV, FLAC or MP3 format, probably stereo or binaural, possibly in surround sound, for various modes of playing.

Similarly, research data is recorded in the laboratory (or in the field) and leads to publication. Once a printed article, this has been largely replaced by the PDF, very much a digital version of the physical media; i.e. electronic paper. The PDF is produced and distributed for download, like an MP3. Where the research is publicly funded there may be an obligation to publish the research data that supports the results in the paper.

In both cases, then, the content recorded at source is rich and reusable, and the more the richness can be preserved then the better the object can be interpreted, reused, and repurposed. This line of thinking in scientific publishing has led to the notion of the *Research Object*, essentially an aggregation of experimental description and data into a single identifiable and shareable bundle [15].

The goal of Research Objects is "...to create a class of artefacts that can encapsulate our digital knowledge and provide a mechanism for sharing and discovering assets of reusable research and scientific knowledge". Significantly a Research Object is also a social object in that it can be shared and is a subject of discourse, like the physical thing [16].

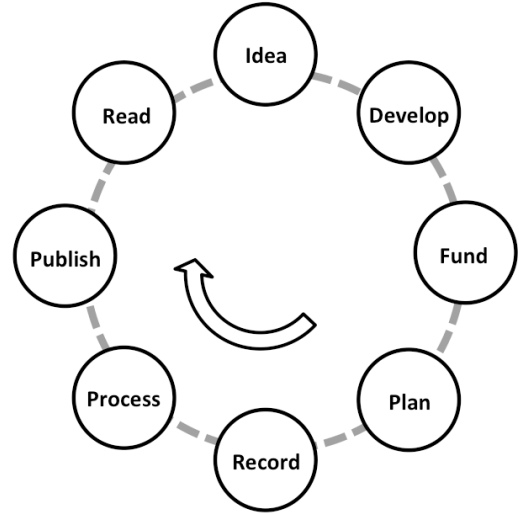


Fig. 4. The Research Cycle (due to [14]).

By analogy we can define the *Digital Music Object* (DMO), which bundles the rich recorded content of music performance for reuse. In this case, the content is of significant size and may be stored in various places and formats for production use. Furthermore, it is temporal in nature. DMOs could be the interchange between studios and production steps, or a DMO could be seen as a mutable object that evolves through the workflow. While the notion of Research Object has been intended to embrace all kinds of content, this has not been the emphasis in today's realizations which more often capture data supporting a published article in a repository, perhaps delivering zip archive or tarball. The DMO must provide a means of gathering distributed content (e.g. a descriptive manifest) and work live as well as in the archive.

The notion of DMO presented here is very much about the object as container, but equally we could be looking at the semantic representation or augmentation of its content. Recent work with the HathiTrust Research Center in the EIEPHaT project has demonstrated the need to construct and interact with complex research collections or worksets. These require extensions to existing bibliographic ontologies in order to represent the items and their relationships comprehensively [17]. Further consideration of DMO as workset is an important line of enquiry which is set to bring insights from digital scholarship back into the music domain. Semantic audio techniques will also inform content representation.

To conceptualize DMOs we are conducting a series of experiments in performance capture and reuse. We introduce five of these in the next section. Together they will deepen our understanding of process descriptions, and inform the design and practice of DMOs.

VI. UNDERSTANDING THE REQUIREMENTS FOR DIGITAL MUSIC OBJECTS

A. Capture “in the wild”

We have conducted an experiment to understand requirements for DMOs based on current democratised recording practice in the pro-am arena. Using portable digital recording equipment (Focusrite Scarlett 18i20) we have captured live audio in three venues and made the multitrack recordings and “off the desk” mixes available for reuse. Here we distributed the stems in multitrack audio format and different people have developed alternative mixes. We have made some of the audio available for annotation via Music Circle [18].

This in-the-wild exercise has demonstrated the multitrack recording as a social object (physical too, as it was also passed around by USB stick). Outputs have been multiple mixes, including a surround mix, and annotations. Additionally we are experimenting with research-object tooling as applied to this music content, in order to ‘see what breaks’ and to influence the design use cases of a new Research Object system which can be used for recordings of musical performance.

The DMOs produced by this project are stems in standard multitrack audio formats. A “production demonstrator” is currently being designed which focuses on the labelling and grouping of the channels in studio recordings, with the computer suggesting labels based upon inferences from features extracted by signal processing and by other people, perhaps from other stages of the workflow. It aims to show how labels can be derived automatically, shared, and renegotiated between people.

Making the stems available for mixing, and sharing the mixes, is a way of building on human experience and creativity; it also informs future automated mixing by machine.

B. Annalist

Annalist is a software system for individuals and small groups to reap the benefits of using RDF linked data, supporting them in easily creating data that participates in a wider web of linked data [19]. It presents a flexible web interface for creating, editing and browsing evolvable data, without requiring the user to be familiar with minutiae of the RDF model or syntax, or to perform any programming, HTML coding or prior configuration. Requirements centre particularly on achieving low activation energy for simple tasks, flexibility to add structural details as data is collected, access-controlled sharing, and ability to connect private data with public data on the web.

Annalist is designed as a web server application, presenting an interface for defining data structure and managing data. Data is stored as text files that are amenable to access by existing software, with the intent that a range of applications may be used in concert to gather, manage and publish data. During its development, Annalist has been used in a range of applications, which have informed decisions about its design and proven its flexibility and robustness in use.

One of these applications has been the capture of data during a live music performance, including “Carolan Guitar” which reports its own life story [20]. Thus Annalist is serving as an on-ramp to capturing many aspects of the DMO beyond the digital audio.

C. Transforming Musicology

Performance of a musical work potentially provides a rich source of multimedia material for future investigation, both for musicologists’ study of reception and perception, and in improvement of computational methods applied to its analysis. This is particularly true of music theatre, where a traditional recording cannot sufficiently capture the ephemeral phenomena unique to each staging.

In the Transforming Musicology project (see <http://www.transforming-musicology.org/>) we introduced a toolkit developed with, and used by, a musicologist throughout a complete multi-day production of Richard Wagner’s *Der Ring des Nibelungen* (“Ring Cycle”) in Birmingham, UK, in November 2014 [21]. The toolkit is centred on a tablet-based score interface through which the scholar makes notes on the scenic setting of the performance as it unfolds, supplemented by a variety of digital data gathered to structure and index the annotations. Subsequently we structured the data for reuse and further investigation using semantic web technologies, and considered the utility of our tooling from both a user perspective and through an initial quantitative investigation of the data gathered. Additionally we captured physiological data (galvanic skin response, heart rate, micro-movements in three dimensions) from multiple participants—in some ways like additional recording channels.

This illustrates the increasing richness of performance capture. In this regard the requirement of DMOs for musicology are similar to those of the science lab and the studio: multiple and complex data sources, possibly stored in multiple locations, temporally synchronized, richly described and linkable. Today music is seen as a continuous cultural practice, perceived in its context, as opposed to the traditional abstraction that is the score [22].

D. Numbers into Notes

December 2015 saw the 200th anniversary of the birth of Ada Lovelace. A major symposium was held to mark the occasion [23], including the discussion of a thought experiment: had Ada Lovelace lived longer, and had Charles Babbage successfully built the Analytical Engine, what might have happened to pursue this observation by Lovelace:

“Supposing, for instance, that the fundamental relations of pitched sounds in the science of harmony and of musical composition were susceptible of such expression and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent.” (note A in [24]).

As part of this we developed an interactive tool for people to generate music from integer sequences. The workflow of the tool mirrors our hypothesised workflow involving the Analytical Engine: the machine runs a parameterised program

to generate a number sequence, and parts of this sequence are then given to different instruments. Inspired by the use of punched cards in the Jacquard loom and the proposed analytical engine, we generate virtual ‘piano rolls’. Hence this is an exercise in algorithmic music generation but with moments for creative human intervention.

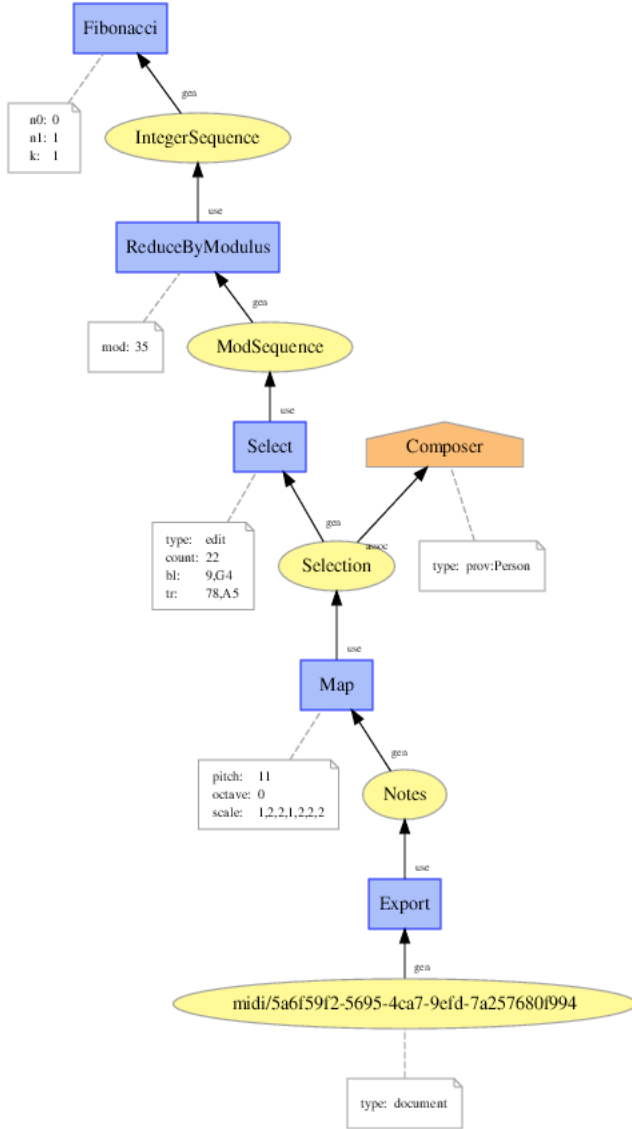


Fig. 5. The Provenance graph describing the output from the “Numbers into Notes” tool, using a Fibonacci sequence. A detailed rendering of this image may be obtained by using the tool.

The programmer and operator (or ‘attendant’) were not allowed to change the numbers generated by the machine, but had full control of the mapping from numbers to notes and then from notes to instruments. The interactive tool, a single page web app (<http://demeter.oerc.ox.ac.uk/NumbersIntoNotes/>), provides several algorithms which illustrate the mathematics of the early 19th century. The primary example involves generalized Fibonacci sequences, reduced by modular arithmetic to produce periodic sequences.

The final stage of the workflow is to export the musical fragment in various formats, one of which is metadata with an automatically generated natural language description of the algorithm parameters, mapping, and selection. We did this to enable someone at a later stage to be able to understand how the fragment was generated or indeed to regenerate the fragment using different tooling; i.e. to reproduce the results of the experiment. For this same reason, one of the output formats is W3C PROV-N, from which an SVG visualization is generated as shown in Figure 5.

E. Lifecycle

We have found it helpful to consider the object as container of digital content but also having a social life—a social object with metadata and annotation, subject of social sharing and of social networks, for consumption by human or machine.

Figure 6 illustrates the lifecycle of the multimedia content associated with a novel performance of the operatic work *Ada sketches* by composer Emily Howard. This was first performed in this format at the Science Museum in London, then based on the London performance we held an event in Oxford, UK in December 2015. The audience completed questionnaires as part of our analysis of the reception of the piece. This is an on-going process, with the next performance in Manchester in July 2016.

The performance was recorded and fed into the Ada Lovelace Symposium, alongside the Fibonacci piece (generated by simulation, but now recorded by a band) and also *Mesmerism*, a second work in Howard’s Lovelace Trilogy. The symposium was also recorded. The resulting materials were used for a presentation at the *Digital Music Research Network* event, and then the entire story was presented reflectively at the *Centre for Digital Scholarship* in the Bodleian Libraries. This is not the end, and nor is there a single archival end point to this process.

F. Discussion

It is clear from these exercises that a DMO must hold together diverse information that might not be routinely captured in individual file formats, including metadata for content and process. However, we also conclude that Digital Music Objects should be realizable through many existing standards; i.e. we do not propose to invent a universal DMO file format that would require pervasive and unmaintainable change.

Our aim then is to define the minimum set of attributes and behaviours that make a DMO, in the spirit of the “Minimum Information Model” for Research Objects [25] which in turn is based on the idea of minimum information checklists. We then need lightweight tooling to support compliant DMOs in some commonly adopted tools, such as Digital Audio Workstations.

Each of these examples has also given us insights into the lifecycle of DMOs. The complexity of reuse in Figure 6 is not just typical of music: in practice, the flow of research data in science looks similar. We also note that our experiments so far have not addressed discovery and curation, and these are also key points for future work.

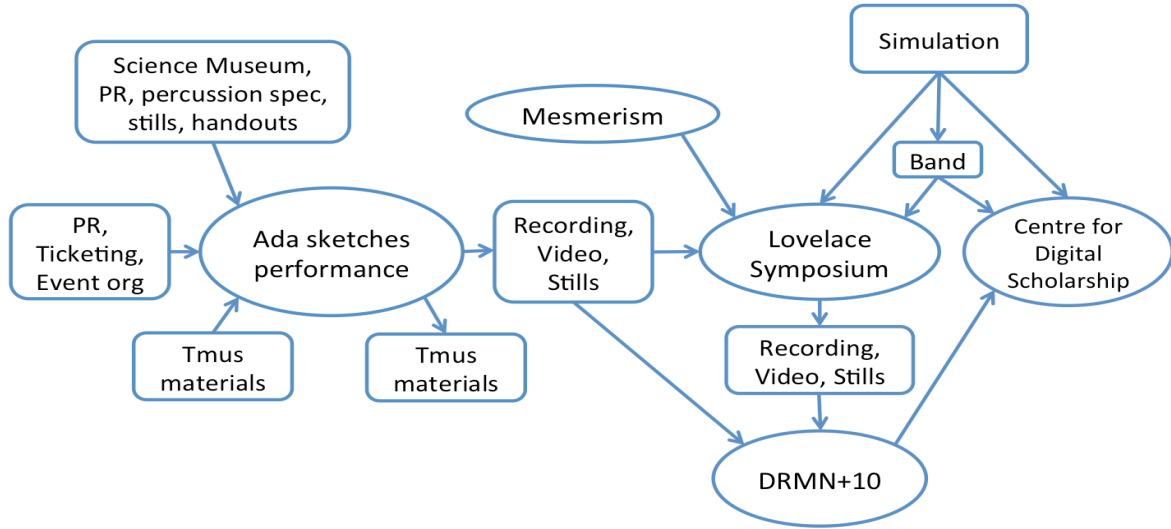


Fig. 6. The flow of physical and digital content in the performance and multiple replays of *Ada sketches* by Emily Howard

VII. CONCLUSION

In this paper we have explained the motivation for a comparison between the creation, production, distribution, consumption, and reuse pipelines of music and science. We have exercised a preliminary comparison framework, looking at the ‘megatrends’ afforded by digital content: duplication and distribution, democratization and disintermediation, and automation, as well as the customer activity cycle and the content lifecycle. On-going “in the wild” developments and capture scenarios have further informed the comparison.

The similarities between laboratory and studio are striking, and they appear to have common trajectories. Our music tooling may directly support or inform tomorrow’s laboratory practice, while reproducible research practice may lead to richer and more re-usable music content—especially in the face of automation.

The next phase of our work will define the minimum DMO concept and deploy tooling for further experimentation. Our key observations so far are:

a) Our work has emphasized the importance of describing process, both prospectively and retrospectively. Our early experiments with W3C PROV confirm its utility and potential. While it demonstrably works well when process descriptions are complete and precise, it does not, of itself, allow for multiple interpretations in multiple contexts. To explore this, our future work will also address provenance of provenance.

b) The digital-social intersections are key, but so too are the digital-physical intersections, as we interact with our (musical and scientific) instruments and materials in the studio, performance space, laboratory, and lecture room. Capturing a full range of contextual information is an essential part of the richness of the DMO.

c) The most compelling insight from our comparison so far is the science/music as performance perspective, with a common theme around the role of creativity. We have gained an appreciation of embodiment—what it means to be the human participant in creating and interpreting performance. This takes us to theories of cognition, and we are currently looking at the notion of *procedural blending* [26], utilising the Annalist tool.

We see that music is an embodied aspect of cognition. Drawing our parallels, how does this apply to science? The situated knowledge approach of Haraway has comment on this, striving for a faithful account of the real world while making explicit our perspective and positioning within the world [27]. All these examples remind us that the audience member (or observer, annotator, scientist) is engaged in a performance, and that annotations captured live are being made by that interpreter in that context and should be interpreted in that knowledge. This argues for more comprehensive consideration of our ‘interpreters’ in the creation and use of Research Objects and DMOs. It also raises questions about the reproducible record of research.

ACKNOWLEDGMENT

The authors are grateful to their colleagues for many discussions, including Steve Benford and Alan Chamberlain (University of Nottingham), Terhi Nurmikko-Fuller, Carolin Rindfleisch and Ségolène Tarte (University of Oxford), Mark Sandler and Geraint Wiggins (Queen Mary University London), Emily Howard (Royal Northern College of Music), Iris Garrelfs (University of the Arts London), and Jeremy Frey (University of Southampton).

REFERENCES

- [1] A.J.G. Hey, S. Tansley, and K. Tolle (eds.) The Fourth Paradigm: Data-intensive scientific discovery. Microsoft Research, 2009.
- [2] J. Naughton. Lecture: Getting from here to there. *MedieKultur. Journal of media and communication research*, v. 30, n. 57, p. 14 p., nov. 2014. Available at: <http://ojs.statsbiblioteket.dk/index.php/mediekultur/article/view/18609>. Date accessed: 30 May 2016.
- [3] D. De Roure, C. Goble, S. Aleksejevs, S. Bechhofer, J. Bhagat, Cruickshank, D. et al. The evolution of myExperiment. *Proceedings - 2010 6th IEEE International Conference on e-Science, eScience 2010*, 153-160. doi: 10.1109/eScience.2010.59
- [4] J. Freire, N. Fuhr and A. Rauber. Reproducibility of Data-Oriented Experiments in e-Science (Dagstuhl Seminar 16041). *Dagstuhl Reports*, 6(1), 2016. doi: 10.4230/DagRep.6.1.108
- [5] J.G. Frey, D. De Roure and L. Carr. Publication at Source: Scientific Communication from a Publication Web to a Data Grid. In *Proc. EuroWeb*, (Oxford, UK, Dec 2002), British Computer Society.
- [6] G. Hughes, H. Mills, H., D. De Roure, D., J.G. Frey, L. Moreau, m.c. schraefel, G. Smith, and E. Zaluska. The semantic smart laboratory: a system for supporting the chemical eScientist. *Organic & Biomolecular Chemistry*, 2, (22), 3284-3293. 2004. doi: 10.1039/B410075A
- [7] J.S. Downie, A.F. Ehmann, M. Bay and M.C. Jones. The Music Information Retrieval Evaluation eXchange: Some Observations and Insights. *Advances in Music Information Retrieval* Vol. 274, pp. 93-115. 2010.
- [8] K.R. Page, B. Fields, B., D. De Roure, T. Crawford, J.S. Downie. Capturing the workflows of music information retrieval for repeatability and reuse. *J. Intell. Inf. Syst.* 41(3): 435-459. 2013.
- [9] Y.Gil and S. Miles (eds.) PROV Model Primer. W3C Working Group Note 30 April 2013. Available as <https://www.w3.org/TR/prov-primer/>
- [10] S. Vandermerwe. 2000. How increasing value to consumers improves business results. *Sloan Management Review*, 42, 3, 2000, 27-37.
- [11] T. Regner, J.A Barria, J.V. Pitt, and B. Nevill. An artist life cycle model for digital media content: Strategies for the Light Web and the Dark Web, *Electronic Commerce Research and Applications*, Volume 8, Issue 6, Nov-Dec, 2009. Pages 334-342, doi: 10.1016/j.elerap.2009.05.002.
- [12] A. Pras, C. Guastavino and M. Lavoie. The impact of technological advances on recording studio practices. *J. Am. Soc. Inf. Sci.*, 64: 612-626. 2013. doi:10.1002/asi.22840
- [13] C. Bird and J.G. Frey. Chemical information matters: an e-Research perspective on information and data sharing in the chemical sciences. *Chemical Society Reviews*, 42, (16), 6754-6775. 2013. doi: 10.1039/C3CS60050E.
- [14] C. Neylon. Science in Society. NESTA Crucible Workshop, Lancaster, 28 June, 2009. Available as http://commons.wikimedia.org/wiki/File:Research_cycle.png
- [15] S. Bechhofer, I. Buchan, D. De Roure, P. Missier, P. et al. 2013. Why Linked Data is Not Enough for Scientists, *Future Generation Computer Systems*, Vol. 29, No. 2. pp. 599-611.
- [16] D. De Roure, S. Bechhofer, C. Goble and D. Newman. Scientific Social Objects: The Social Objects and Multidimensional Network of the myExperiment Website, In *Proc. IEEE Third International Conference on Social Computing* (Boston, MA, USA, 9-11 Oct), SocialCom. pp.1398-1402, 2011. doi: 10.1109/PASSAT/SocialCom.2011.245
- [17] T. Nurmikko-Fuller, K.R. Page, P. Willcox, J. Jett, C. Maden, T. Cole, C. Fallaw, M. Senseney and J.S. Downie. Building Complex Research Collections in Digital Libraries: A Survey of Ontology Implications. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* (Knoxville, USA, June 2015).
- [18] M. Yee-King, M. Krivenski, H. Brenton, A. Grimalt-Reynés, and M. d'Inverno. Designing educational social machines for effective feedback. In *Proc. 8th International Conference on e-learning*. (Lisbon, Portugal, 15-18 July). 2014.
- [19] G. Klyne, C. Willoughby, K.R. Page. Annalist: A practical tool for creating, managing and sharing evolving linked data. *Linked Data on the Web Workshop 2016*. To appear in CEUR Workshop Proceedings (CEUR-WS.org) Vol-1593.
- [20] S. Benford, A. Hazzard, and L. Xu. The Carolan guitar: a thing that tells its own life story. *ACM Interactions*, Volume 22 Issue 3, May - June 2015, Pages 64-66. doi: 10.1145/2745960
- [21] K.R. Page, T. Nurmikko-Fuller, C. Rindfleisch, R. Lewis, L. Dreyfus and D. De Roure. A toolkit for live annotation of opera performance: Experiences capturing Wagner's Ring Cycle. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR 2015)*. Málaga, Spain.
- [22] N. Cook. *Beyond the Score: Music as Performance*. Oxford University Press, Oxford. 2014.
- [23] Ada Lovelace Symposium 2015 - Celebrating 200 Years of a Computer Visionary, Oxford, UK, December 2015. ACM. 2015. ISBN 978-1-4503-4150-9.
- [24] A.A. Lovelace. Sketch of the analytical engine invented by Charles Babbage, with notes by the translator. In *Scientific Memoirs, Selected from the Transactions of Foreign Academies of Science and Learned Societies*, Vol. 3, 1843, pp. 666-731, volume 3. Richard and John E. Taylor, Red Lion Street, Fleet Street, London. Translation of *Notions sur la machine analytique de M. Charles Babbage* by Luigi Federico Menabrea, in *Bibliothèque universelle de Genève*. Nouvelle série 41, 352-76 (1842).
- [25] M. Gamble, C. Goble, G. Klyne and J. Zhao, MIM: A Minimum Information Model vocabulary and framework for Scientific Linked Data, *IEEE 8th International Conference on e-Science*, Chicago, IL, 2012, pp. 1-8. doi: 10.1109/eScience.2012.6404489
- [26] I. Garrelfs. From Conceptual Blending to Procedural Blending: Applying a Model of Cognition to Process in Sound Art Practice. In Denham, S. and Punt, M. eds., *Plymouth: TT OA Papers* pp. 71-88. 2016.
- [27] D. Haraway. Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599.1988.