Moving Object-Based Media Production from One-Off Examples to Scalable Workflows

By Jasmine Cox, Matthew Brooks, Ian Forrester, and Mike Armstrong

Abstract

In this paper, we outline our progress toward creating tools and workflows for object-based media (OBM) production, taking us from one-off demonstrators to scalable production, through the creation of production tools based on shared data models. We feature a recent example of OBM created from the ground up and discuss the lessons learned from this production. We then discuss the progress we are making toward creating a kit of OBM software tools and workflows. Finally, we look at our progress toward building a community of practice for OBM.

Keywords

Data models, interactive, responsive media, storytelling

Introduction

At BBC Research and Development, we have been developing the concept of object-based media (OBM), exploring the role this approach can play in enhancing the experience for our audiences and its impact on storytelling. In each case, the creation of these object-based experiences has been a manual process of analysis, tagging, assembly, and software development. Most of these experiences involved reversioning existing linear media. This paper follows the creation of our most recent example of OBM, the Cook-Along Kitchen Experience (CAKE), which was conceived and produced as an object-based experience from the outset. This paper looks at how we are applying the lessons learned from our previous work to the development of Object Based Media data models and software tools. The paper also discusses how we intend to involve content creators from both inside and outside the BBC and build a community of practice around the development of new forms of media.

This paper looks at how we are applying the lessons learned from our previous work to the development of Object Based Media data models and software tools.

Digital Object Identifier 10.5594/JMI.2018.2806499
Date of publication: 1 May 2018

Background

It is now 20 years since Janet Murray's *Hamlet on the Holodeck* explored the possibilities for personalized narrative,[1] and while computer games have developed greatly over this time, there has been a slower development of more conventional media. The BBC first explored the possibilities of interactive radio drama in 2001 with *The Wheel of Fortune* and *The Dark House* in 2003, both authored as three parallel conventional radio plays each following one of the three characters. The interaction enabled the listener to switch between the different character perspectives,[2] but the production of the three synchronized versions proved highly challenging and very time consuming.[3] Since then, other projects outside of the BBC have explored similar media challenges and developed tools. The *NM2: New Millennium, New Media* project created several interactive TV experiences, most with unusual narrative forms. *Accidental Lovers* applied a circular structure to a drama, whilst *Gormenghast Explore* added spatial navigation to a TV fantasy series and *A Golden Age* provided a sequence of video clips prioritized by topic.[4] The NM2 project also developed a player and production tools for describing the interaction and the video editing process. These ran with an underlying logic engine and were described using their own Narrative Structure Language.[5]

Commercial companies have also explored the potential of these kinds of experiences. TouchCast have created tools to produce what they call Smart Video, where a TV program can be enhanced with links to documents or additional media content. Eko (formerly Interlude) has explored various forms of interactive and branching media, producing narrative examples like *Possibilia* as well as quiz formats, and music videos allowing us to switch between different video tracks.[6] These interactive music videos have a structural resemblance to early interactive BBC radio dramas.

1545-0279/18©2018SMPTE

The first example of OBM from BBC Research and Development was a short audio drama, *Breaking Out*, where the dialog varied according to the location of the listener.[7] We then created a variable length radio documentary reusing the material from a pre-existing program. The narrative themes and structure of the program were broken down into separate objects which could be assembled on the fly to create complete narratives of differing durations.[8] More recent work has explored the use of object-based ideas for production tools: *Discourse* was developed to enable audio editing of speech via a text-based interface, whilst *Squeezebox* provided automated video re-editing via semantic mark-up. *Storyarc* provided a new structured continuity database for the long-running Archers radio program, and *Visual Perceptive Media* was the first experiment conceived and built as an object-based experience from the start.[9]

Cook-Along Kitchen Experience

The Cook-Along Kitchen Experience was built as a showcase for the object-based concepts developed in previous demonstrators. The aim was to create a compelling audience-facing proposition that demonstrated the value of the responsive capabilities enabled by OBM in teaching skills, and further our understanding of appropriate object models. Through this project, we set out to understand the workflow for building a visual object-based production: From conception and planning, through shooting and recording, to post-production and playback.

CAKE is a responsive learning experience implemented in the web browser. It enables someone cooking at home to follow recipes at their own pace. It is a tailored cookery show that can tailor its ingredients according to the number of people being catered for and the availability of appliances in the home. The viewer selects one or more recipes, then adds the number of people they are cooking for, and the number of hob rings available. The cooking steps are dynamically scheduled by the application, which resolves a number of constraints to ensure that the elements of the meal are ready in time for dinner. CAKE offers the audience video clips describing each step, alongside an alternative of recipe cards and waits for the viewer to complete each step (**Fig. 1**). The viewer then triggers the start of the following step.

We ran a prototype object-based shoot to tryout various methods of tracking the production, and create test material. From this, we learned how to organize the data into a structured form that minimized duplication of video sequences. Care was required to plan the structure of the cooking steps to enable them to support multiple recipe combinations. Sequences that required all steps to be completed in a series were distinct from those not tied to any stage in the process. We also ensured that the language used by the presenter contained no mention of quantities or references between different recipes, so that they could be used in any combination. This was achieved by manually mapping recipe "stories" with cards on a board that was used to create a data structure in a spreadsheet. We employed a domain expert, Bella Wright, to create the recipes. We asked her to give us five recipes that both worked well on their own and in combination with each other.

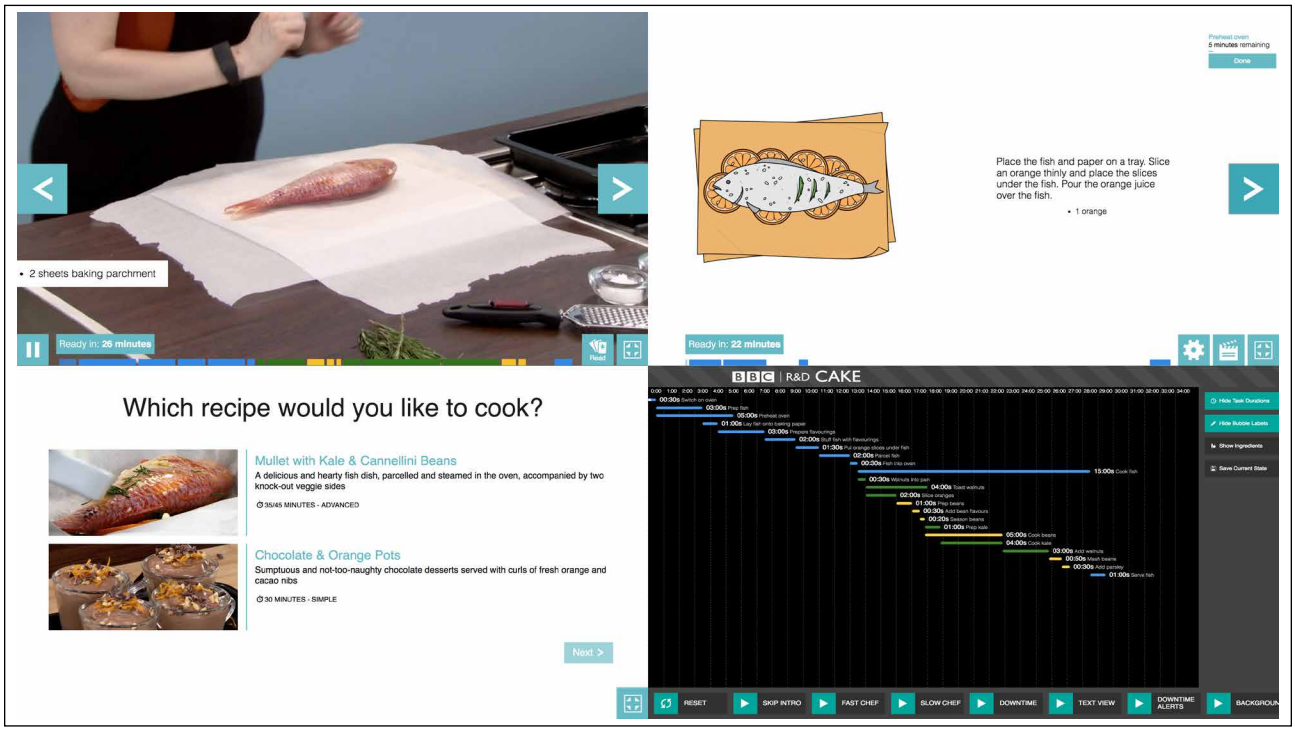

FIGURE 1. Screenshots of the CAKE application audio visual mode, text mode, recipe selector, and backend.

These recipes were then broken down into individual steps which defined the individual objects to be created.

The shoot was a conventional multicamera production with two television cameras and operators and four fixed, wide-angle, 4k cameras. Shooting time was minimized by avoiding the duplication of steps in the captured video and placing the variations into the final application. The shooting script was derived from the pre-production spreadsheet, and items were enumerated by objects and shots, while logging was entirely manual. Post-production involved developing a custom software application, the careful mapping of shots to objects, and conventional video editing. We created a linear video for each dish and associated it with metadata representing the objects within it. The application then plays out objects as defined by the in and out points in the metadata.

A scheduling algorithm is used to organize and play-out media objects and recipe instructions in an optimal order. It takes into account the recipe data, recipe choices, and current progress through the tasks. It uses critical-path methods to identify the shortest possible completion time through a set of related activities and resource-leveling to distribute the cooking steps around the fridge, hob, and oven. This orders steps logically for a particular kitchen and set of recipes. It also allows tasks to be carried out simultaneously where possible.

We ran a proto-pilot with 14 people using a low-fidelity version of the final experience. We found that simple features, such as stopping after each step, were seen as extremely valuable, and this informed our layout design for the controls. We also refined the data model to better fit the requirements of the experience, acknowledging that this data model was tied directly to the domain of cooking, rather than being an extensible or generalizable model for other object-based experiences.

The final design was built using the BBC's GEL frameworks[10] along with some custom elements to provide a familiar web video experience with built-in natural interaction. A locally hosted version with a dynamic back-end visualization of the schedule was created as a demonstrator for IBC2016 as shown in **Fig. 1**. This succeeded in conveying the concepts behind OBM and prompted a lot of interest. The discussions highlighted how OBM is widely applicable to teaching skills—e.g., people could easily conceive interleaved learning experiences beyond public service broadcasting. An optimized version was launched on BBC Taster in January 2017 allowing anyone with a tablet or laptop to try it out for themselves.[i] A controlled home trial will give us more in-depth understanding of the interaction data, and where CAKE may be able to add value in terms of support for learning experiences. The key aspect of CAKE is that it places audience members at the heart of an internet protocol (IP) delivered media experience in an entirely new way. In creating it, we have highlighted the need for new tools and workflows to make the process sustainable. Our work has moved on to deliver on four areas we see as essential to the sustainability of OBM, namely, new exemplars, data models, tools, and a community of practice.

[i] *http://www.bbc.co.uk/taster/projects/cook-along-kitchen-experience*

Data Models for OBM

Looking back over our audience-facing OBM projects, we can identify clear problems with our tooling. The majority of projects were authored using short–term, unsustainable approaches such as spreadsheets and python scripts (with the exception of projects like Squeezebox, which were themselves prototype tools). This leads to a situation where the construction of the media was reliant on the "engineer in the middle" who was often required in order to make changes to experiences. We also identified the problems caused by having individual (sometimes unexpressed) data models for each experience.

Crucially, we saw our lack of sustainable tooling as a barrier to the adoption of OBM by any potential community of practice. We want to transfer the creation of OBM experiences out of the realm of Research and Development engineers and into the world of craftspeople—it is the community of practice that will explore the potential of OBM, provide us with real-world use cases, and give us feedback to help develop our tools and workflows.

A Generalized Model for OBM Storytelling

We wanted to create a core generalized data model to describe all the object-based experiences we can imagine in the future, informed by our earlier work. The data model expresses OBM narrative structure, presentation structure, and production workflows. It allows us to have tangible conversations about OBM, enables interoperability, and allows the wider community to create their own tools.

We have chosen to model OBM narratives as hierarchical finite-state machines—a simple method for expressing interactivity more familiar to the world of gaming than broadcasting. At the top level, we have a *Story*, which holds a collection of *Narrative Objects*. Narrative flow possibilities are expressed using one or more *Links* to other *Narrative Objects*. The narrative concerns are held entirely within these two levels of the data model (**Fig. 2**).

Each *Link* has an associated *Condition*—a Boolean expression composed of any state available to the OBM experience—e.g., variables holding information such as time of day, locale, and the viewer's choices, preferences, or profile. These links are evaluated in sequence until a *Condition* returns true—establishing which *Narrative Object* will be visited next. Additional *Link* types handle *Story* beginnings and endings.

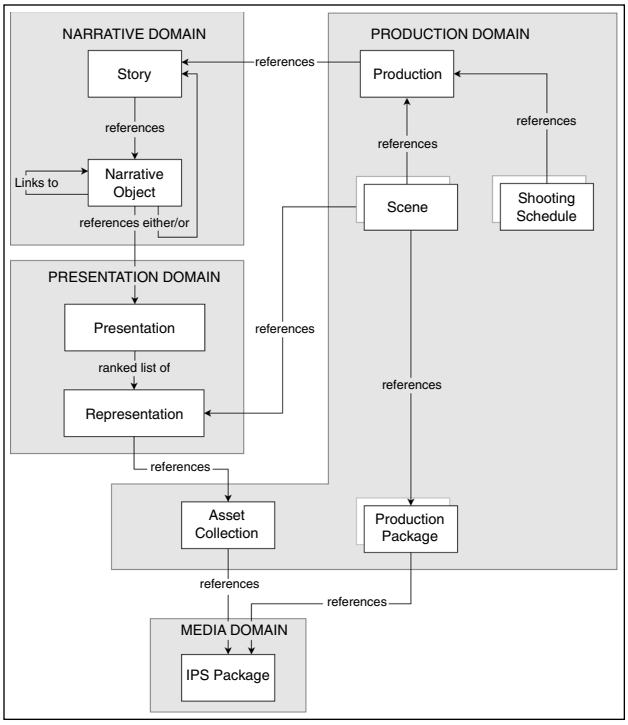

FIGURE 2. Data structure outline.

A *Narrative Object* needs a media representation to convey meaning to the viewer. For this, *Narrative Objects* reference a *Presentation Object* (or, to handle grouping and hierarchy, another *Story*). The *Presentation Object* deals with how an individual part of the story is presented to the viewer. It consists of one or more *Representations* which enable identical narrative content to be conveyed by different media. For example, in CAKE, each recipe step could be viewed as a video, or as a recipe card. In the *Presentation Object*, each *Representation* has an associated *Condition* of a similar format to those in *Narrative Objects*—in this way, presentation can be editorially controlled by using an experience state.

Each *Representation* points to an *Asset Collection*, which references an IP Studio[11] package holding the media necessary to present the *Narrative Object*. *Representations* themselves are not concerned with bitrates or codecs—these concerns will ultimately be handled by an OBM player that consumes the *Story* data model, evaluating network conditions and device capability to request media at an appropriate fidelity.

A Generalized Data Model for OBM Production

We also model associated production workflows. Each *Story* references a single *Production Object*—a container referencing all media and metadata pertaining to the production of a *Story*. *Scene Objects* are associated with *Story Representations*, acting as a container for all production content relating to the representation. The *Scene Object* holds links to captured media under a rushes list—each rush holding production metadata such as *Take* number and *Shot* number. The *Scene Object* also holds links to edited content created from the rushes, in the form of a *Production Package*—a wrapper around an IP Studio package associated with production metadata. During playback, the viewer will experience the most recent version of a *Production Package* referenced by a *Representation*. Additionally, we can use the *Scenes* within a production to create a digital *Shooting Schedule* for the production.

Tooling and Workflow for OBM

We are creating a number of tools as part of an OBM toolkit, which will describe flexible OBM workflows (**Fig. 3**). The tools are built around our data model and on the top of our Object-Based Production Tools in the Cloud platform.[12]

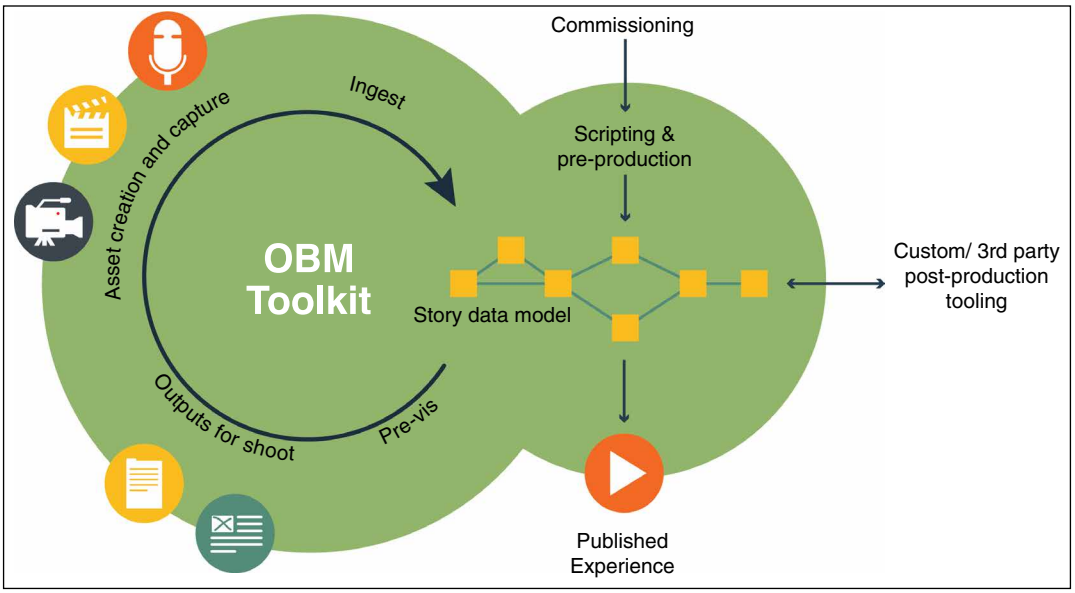

FIGURE 3. OBM production workflow.

The **StoryFormer** tool allows rapid wireframing of OBM experiences, authoring the *Story* data model. The tool enables *Narrative Elements* to be created and linked, with *Conditions* on *Links* used to express narrative flow. Before any assets have been filmed or recorded, proxy assets can be assigned to *Representations*—e.g., a representative image or text snippet. In this way, an interactive experience can be previsualized, iterated, refined, and tested before any media is recorded.

The **StoryShooter** tool is designed to help manage the flow of media into a *Story's Production Object* and is designed for use during a shoot or recording session. As every *Representation* is associated with a *Scene* in the production domain, we can construct and monitor the progress of a *Shooting Schedule*, which ultimately populates *Production Packages* used by *Representations* in a *Story*. The linkage between *Scenes* and *Representations* allows us to digitally describe the media assets that require creation, eliminating manual, error-prone artifacts such as the preproduction spreadsheets used by CAKE. This linkage also eliminates manual logging and any postshoot manual ingest process—media created to represent a *Narrative Element*, and any logging information, are associated automatically with the *Narrative Element*'s *Representation*. Relevant footage for each *Narrative Element* can then be surfaced in **StoryFormer**. As each *Scene* may contain multiple takes and logging information, a craft edit will be required to produce the final asset for *Representations*. This will be achieved with a bridge to existing nonlinear editing tools, or a simpler web-based editing interface. We envisage that a *Story* will complete multiple iterations through **StoryFormer** and **StoryShooter** before it is completed. The final tool of the initial suite is the OBM player, which consumes the *Story* data model, handles viewer interaction, and produces final output for the viewer.

It could be considered that traditional workflows start with the first ingest—other digital artifacts such as scripts and pre-production data exist, but are not tied to the production as a whole. The OBM Toolkit sees value in connecting all production artifacts with a *Story* and associated *Narrative Elements*, as these production artifacts contain rich sources for metadata that can be used to describe an experience at a highly granular level. If content is described at a highly granular level as part of a regular production process, numerous higher order applications become possible by default. For example, semantic navigation and search are possible, and meta-OBM experiences (like trailers, series summaries, and catch-up services) can be automated. Previously, this has been achieved in post-production, either manually or through automated metadata recovery.

Next Steps

There are still many specific details to resolve, both with our data model and our tooling. Key opportunities include the methods that enable authors to express *Link* conditions in editorial terms, without the necessity to write Boolean logic. Specifying user interfaces that populate the experience state (e.g., explicitly choosing a recipe). Handling the composition of additional media into narrative (e.g., background music or graphical overlays). Enabling simultaneous editing of a *Story* by multiple users.

Building a Community of Practice

We need to support a culture of experimentation around the craft and quality of OBM. In order to expand our perspective on what this technology can provide and how it could be developed, we are in the process of encouraging the growth of a community of practice in this area. This has three key aims: **Awareness**, seeking out people and organizations already interested in or working on adaptive narratives though talks, workshops, and conferences. **Advocacy**, presenting our work and demonstrating best practice in our work and methods as we explore OBM and connecting people through networks like the Storytellers United slack channel and helping to share perspectives and knowledge. **Access**, providing early access to our emerging software tools to a wide community for trial use and to generate feedback on their development. We are also partnering with the Department of Theatre, Film, and Television, University of York, including part of the original MN2 project bringing invaluable experience and knowledge from the NM2 work. We have been talking to other sectors including museums, healthcare transport, and smart cities, and this has generated example scenarios where OBM could be applicable. For example, in film distribution where a single release could be responsive to the need for either matinée, primetime, or late-night shows, dependent upon the audience. Finally, we are looking to work with content creators and interactive developers to inform how we take this work toward maturity, developing interoperability between software tools and data models to mature OBM toward business as usual, with sustainable toolsets and workflows.

Conclusion

The creation of one-off OBM has demonstrated the potential of this approach to personalized content. The most recent experience, CAKE, was created entirely from scratch and aimed at interactive skills teaching. CAKE has enabled us to develop a deeper understanding of the challenges of producing flexible and responsive media. We are now using this insight to drive our ongoing work on the development of data models and production tools for the creation of powerful new experiences with OBM. The effectiveness and sustainability

of the models and tools will emerge over time as we develop our understanding of the needs with developers and content creators, and engage with a wider range of audiences and use cases.

Acknowledgment

We would like to thank Newcastle University Open Lab for their collaboration during the production of CAKE. Thanks also to BBC Research and Development team members Ben Robinson, Chris Northwood, Jon Tutcher, Thomas Preece, Max Leonard, and Rhianne Jones.

References

1. J. H. Murray, *Hamlet on the Holodeck*, MIT Press, Cambridge, MA and London, UK, 1997.
2. R. J. Hand, *Listen in Terror: British Horror Radio From the Advent of Broadcasting to the Digital Age*, Manchester University Press, Manchester, UK, 2014.
3. D. Greco and N. Fisher, "God, Random Access, or Roulette?," *HyperText Now,* 2001. [Online]. Available: http://www.eastgate.com/HypertextNow/archives/Fisher.html
4. M. F. Ursu et al., "Interactive TV Narratives: Opportunities, Progress, and Challenges," *ACM TOMM*, 4(4):25, Oct. 2008.
5. M. F. Ursu et al., "ShapeShifting TV: Interactive Screen Media Narratives," *Multimedia Syst.*, 14(2):115–132, 2008.
6. R. Khatchadourian, "The Movie With a Thousand Plotlines," *The New Yorker*, Jan. 30, 2017. [Online]. Available: http://www.newyorker.com/magazine/2017/01/30/alternate-endings
7. A. Churnside and I. Forrester, "The Creation of a Perceptive Audio Drama," *Proc. 2012 NEM Summit*, Istanbul, Turkey, pp. 16–18, 2012.
8. M. Armstrong et al., "Object-Based Broadcasting—Curation, Responsiveness and User Experience," Presented at *IBC* 2014, Amsterdam, The Netherlands, Sep. 2014.
9. M. Evans et al., "Creating Object-Based Experiences in the Real World," Presented at *IBC* 2016, Amsterdam, The Netherlands, Sep. 2016.
10. British Broadcasting Corp. (BBC), "2016 Global Experience Language (GEL) Guidelines." [Online]. Available: http://www.bbc.co.uk/gel
11. P. J. Brightwell et al., "The IP Studio," Presented at *IBC* 2013, Amsterdam, The Netherlands, Sep. 2013.
12. C. Northwood and R. Wadge, "An Architecture for Cloud-Based IP Video Production Tools," Presented at *IBC* 2017, Amsterdam, the Netherlands, Sep. 2017.

About the Authors

Jasmine Cox is a designer and producer with BBC Research and Development's Future Experience Technology group. Championing human-centered design in the field of human computer interaction Cox is helping to take object-based media to the next level by supporting content creators with tools and workflows that enable real-world production for responsive object-based experiences. With a background in product design, she has expertise across interactive, electronic, industrial, and mechanical designs. She crafts playful experiences, devices, and control systems—embedded into the web.

Matthew Brooks is a lead engineer in the BBC Research and Development's Future Experience Technology group. Brooks leads the team's object-based media workstream, which explores how object-based approaches can be used to create responsive, personalized, and flexible audience experiences, and intelligent, assistive media production tools. Brooks is particularly interested in the creative application of technology to storytelling, and with 15 years of experience in videogames development, brings new perspectives to the traditionally linear world of broadcasting. Brooks has written publications for the Audio Engineering Society (AES), IBC, and TVX, covering object-based media, haptics, and dynamically positioned subtitles. He has exhibited work at IBC and Sheffield Doc/Fest, co-created interactive installations at Tate Modern and Manchester MOSI, launched the world's first variable length radio program, and worked on several major videogame franchises.

Ian Forrester is a well-known and likable character on the digital scene in the U.K. He works in the BBC Research and Development's Future Experiences Technology group. His current research is in the area of future narrative and storytelling; using a new approach to broadcasting, which pairs the best of broadcast with the best of internet technology; and creating experiences like sitting around a camp fire telling stories.

Mike Armstrong has led the BBC's research into access services over the past few years, leading to significant improvements in the BBC's subtitle services and in the understanding of audience needs. He has a broad expertise, from access services, speech audibility, video quality, and visual perception, and is now focusing on the challenges of creating responsive narratives with object-based broadcasting and on how metadata can be harnessed to automate production processes.

Presented at IBC 2017, Amsterdam, The Netherlands, September 14–18, 2017. This paper is published here by kind permission of the IBC and BBC.