City University London
School of Informatics
MSc in Human-Centred Systems
Project Report
2011

Playing with difficult objects: game designs for crowdsourcing museum metadata

Mia Ridge

Supervised by: Dr Sara Jones

March 17, 2011

A note on this Open Access version, January 2013

This version is as submitted (barring minor typo corrections) but does not contain any of the Appendices except Appendix A. This is mostly because the Appendices contained interview transcripts that would need redaction to remove identifying information.

If you prefer to cite published papers, I have also published some of my research in:

Ridge, Mia. 2011a, Playing with Difficult Objects – Game Designs to Improve Museum Collections. In J. Trant and D. Bearman (eds). *Museums and the Web 2011: Proceedings*. Toronto: Archives & Museum Informatics. Online at: http://conference.archimuse.com/mw2011/papers/playing_with_difficult_objects_game_designs_to

Ridge, Mia. 2011b. Crowdsourcing Games: Playing With Museums. In Beale, Katy (ed) *Museums at play: games, interaction and learning*. Edinburgh: MuseumsEtc.

Otherwise, please cite this dissertation as below, in whichever style you require:

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

or

Ridge, M. 2011. *Playing with difficult objects: game designs for crowdsourcing museum metadata*. Dissertation. (MSc). City University.

This dissertation received a Distinction. (Yay!). My work in this area is continuing in my PhD research on historians and scholarly crowdsourcing.

By submitting this work, I declare that this work is entirely my own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work I acknowledge that I have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. I also acknowledge that this work will be subject to a variety of checks for academic misconduct.

Signed:

Abstract

This project explores the potential for casual browser-based games to help improve the quality of museum catalogue records. The project goal was to design and build casual yet compelling games that would have a positive impact on a practical level, helping improve the mass of 'difficult' - technical, near-duplicate, poorly catalogued or scantily digitised - records that make up the majority of many history museum collections. The project was successful in designing games that created improved metadata for 'difficult' objects from two science and history museum collections: *Dora*, a tagging game, and *Donald*, an experimental 'trivia' game that explored emergent game-play around longer forms of content that required some form of research or personal reference.

Keywords: museums, collections, games, crowdsourcing, websites.

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

i

Table of contents

1   Introduction and objectives

This project investigated the design of casual browser-based games to help improve the mass of 'difficult' - technical, near-duplicate, poorly catalogued or scantily digitised - records that make up the majority of history museum collections.

1.1   The problem

Museum collections websites, whether object catalogues or thematic sites with interpretative content, sometimes fail to achieve levels of public usage commensurate with the resources taken to create them. Many collections websites lack the types of metadata that would aid discoverability, or fail to offer enough information and context to engage the casual or non-specialist visitor who finds themselves on a collection page. As Trant (2009) found, the 'information presented is structured according to museum goals and objectives' and the language used is 'highly specialized and technical, rendering resources inaccessible or incomprehensible'.

This is not a cheap problem to solve. As Karvonen (2010) says:

> Digitising museum objects is expensive. The physical characteristics of museum materials make them unsuited for mass digitising, and because of their uniqueness, creating descriptive metadata for museum objects is a painstakingly slow process.

The words people use to describe an object during a game - that is, the metadata they create about the object - could be useful for other people trying to find the same object. Interesting connections, facts or stories contributed by game players may help other people understand the objects they've discovered or help a museum discover new information about its objects. They could help make collections more engaging, more intellectually accessible, and more browseable. It may even offer 'surprise, coincidence, provenance, and the discovery of new relationships' (Tiltfactor, undated).

1.2   Objective

This project investigated the question:

- can metadata games be designed to encourage people to play with 'difficult' museum objects while producing content that will improve museum websites?

The project also explored a supplementary question that emerged after reviewing related cultural heritage metadata projects:

- can games be designed to take crowdsourcing beyond tagging? That is, games to encourage the creation of metadata that requires more effort, time or skill than tags (but without creating additional institutional resource requirements for advance data cleaning, object selection or manual game content validation)?

1.3   'Difficult' objects

Tagging applications and games have been successfully built in the past. Games with a Purpose (GWAP) proved that games could bring mass audiences to computation problems. The *steve.museum* project showed that the public were interested in tagging art objects, and that the resulting content was beneficial and met a real need (Trant, 2009).

However, art museums and galleries tend to have smaller collections compared to natural history or social history museums, and as representations, artworks can be easily tagged in terms of styles, colours, material, period, content (things, people and events depicted), and can even evoke emotional and visceral responses. Art objects are also more likely to be unique and visually distinct. Social history collections, however, can contain tens or hundreds of similar objects, including technical items, reference collections, objects whose purpose may not be immediately evident from their appearance and objects whose meaning may be obscure to the general visitor.

The makers of the original metadata crowdsourcing games, Games with a Purpose, said:

> ...the choice of images used by the ESP game makes a difference in the player's experience. The game would be less entertaining if all the images were chosen from a single site and were all extremely similar (Von Ahn and Dabbish, 2004).

And yet that exactly describes the 'difficult' technical and social history objects held by many museums, as evidenced by three objects labelled 'toy' in the Powerhouse Museum collection, below. Does that mean crowdsourcing games will not work on difficult objects?

(Accession num: B794) (Accession num: B1121) (Accession num: B1120)

Figure 1 Sample objects from a search for 'toy' in the Powerhouse Museum collection - Models of vertical steam engines

1.3.1 Not all objects are created equal

Both museum objects and the records about them vary in quality. Just as the physical characteristics of one object - its condition, rarity, etc - differ from another, the strength of its associations with important people, events or concepts will also vary. To complicate things further, as the Collections Council of Australia (2009) states, this 'significance' is 'relative, contingent and dynamic'.

When faced with hundreds of thousands of objects, a museum will digitise and describe objects prioritised by 'technical criteria (physical condition of the original material), content criteria (representativeness, uniqueness), and use criteria (demand)' (Karvonen, 2010).

In theory, all objects are registered by the collecting institution, so a basic record exists for each. Hopefully, each has been catalogued and the information transcribed or digitised to some extent, but this is often not the case. Records are often missing descriptions, and most lack the contextual histories that would help the general visitor understand its significance. Some objects may only have an accession number and a one word label, while those on display in a museum generally have well-researched metadata, detailed descriptions and related narratives or contextualised histories. Variable image quality (or lack of images) is an issue in collections in general. This project excludes object records without images but does include many poor-quality images as a result of importing records from a bulk catalogue.

This project posits that objects can be placed on a scale of 'distinctiveness' based on their visual attributes and the amount and quality of information about them. Within this project, bulk collections with minimal metadata and distinctiveness have been labelled 'sprockets', the smaller set of catalogued objects with some distinctiveness have been labelled 'lockets', and the unique, iconic objects with a full contextual history have been labelled 'rockets'. This concept also references the English Heritage 'building grades' model (DCMS, 2010). During the project, the labels 'heroic', 'semi-heroic' and 'bulk' objects were also used.

These labels are not concerned with actual 'significance' or other valuation or priority placed on the object, but relate only to the potential mental models around them and data related to them - the potential for players to discover something interesting about them as objects, or whether they can just tag them on visual characteristics.

In theory there is a correlation between the significance of an object and the amount of information available about it; there may be particular opportunities for games where this is not the case.

Project label	Information type	Amount of information	Proportion of collection
Rockets	Subjective	Contextual history ('background, events, processes and influences')	Tiny minority
Lockets	Mostly objective, may be contextual to collection purpose	Catalogued (some description)	Minority
Sprockets	Objective	Registered (minimal)	Majority

Table 1 Objects grouped by distinctiveness

This can also be represented visually as a power law or pyramid model.

Rockets

Lockets

Sprockets
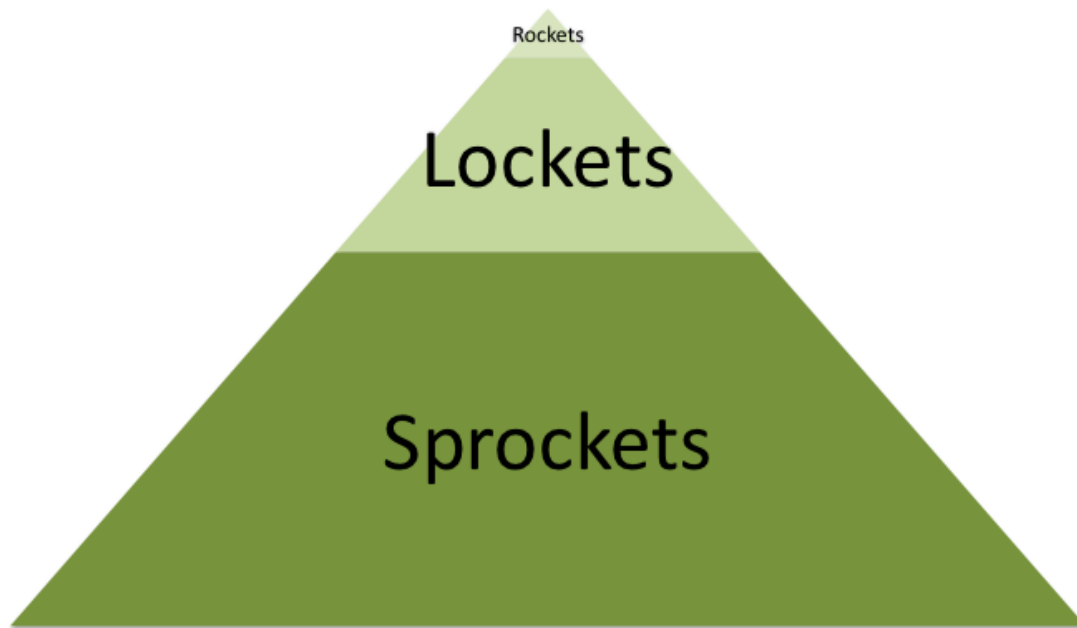
Figure 2 A figurative illustration of the relative numbers of different levels of objects in a typical history museum.

1.4   Definitions

'Metadata games' describes games that involve data about things. Most word guessing games, such as *20 Questions*, *Pictionary*, *Charades* or *Taboo* are metadata games. Clues about the word might be provided by saying related words, or by asking questions to discover the possible categories in which the mystery thing falls, or in the case of *Pictionary* and *Charades*, by drawing pictures of or acting out a thing.

For Von Ahn and Dabbish (2008), most games can be specified through the non-trivial goal the players are trying to achieve, and the rules that define the allowable actions they can take. Gameplay is experienced through the 'challenges and actions that entertain', with challenges representing the obstacles to a player achieving the goal of the game, and actions being the permitted activities that address the challenges (Adams and Rollings, 2007).

Von Ahn and Dabbish (2008) defined a 'game with a purpose' as 'a game in which the players perform a useful computation as a side effect of enjoyable game play'. People playing GWAPs perform basic tasks that cannot be

automated. Under their definition, GWAPs always contribute to solving a computational problem and include a form of 'input-output behavior'.

For the purposes of this project, social history objects are defined as objects that may have appeared in ordinary contexts such as homes and non-specialist workplaces, such as a wall clock. Technical objects are those which have a particular purpose in a specialist field that an ordinary person would not usually have encountered, such as an astrolabe.

2 Research context and literature review

Howe (2006b) was the first to document crowdsourcing, describing companies that were 'using the Internet to exploit the spare processing power of millions of human brains', noting that the distributed nature of the crowds played an important role: '[t]he most efficient networks are those that link to the broadest range of information, knowledge, and experience'. He observes that the crowd has a short attention span, is full of specialists and is good at finding the best content online (Howe, 2006a).

The Games with a Purpose project has produced games including the *ESP Game*, also available as the *Google Image Labeler*[1]. This is a tagging game that produces 'meaningful, accurate labels for images on the Web as a side effect of playing the game' and had gathered more than 50 million labels for images from 200,000 players as of July 2008 (Von Ahn and Dabbish, 2008).

2.1 Tagging and cultural heritage collections

Tagging games have subsequently been applied to art collections (e.g. Brooklyn Museum's *Tag! You're it!*[2]), and manually-curated collections of contemporary audio-visual material (*Waisda?* in Oomen et al, 2010a, 2010b) and archival images (Tiltfactor, undated).

According to Oomen et al (2010b), cultural heritage organisations are motivated to crowdsource the tagging of collections because tags can bridge 'the semantic gap' that exists between controlled vocabularies used by professionals and the search terms of end users, enrich collections 'with factual and contextualized information' and increase 'connectedness with the archive'.

By providing a 'wider and more inclusive array of keywords', tags can increase the findability of collections (Oomen et al, 2010b). The *steve.museum* research

---

[1] http://images.google.com/imagelabeler/
[2] http://www.brooklynmuseum.org/opencollection/tag_game/

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

8

project analysed crowdsourced tags, finding that tags 'provide a significantly different vocabulary than museum documentation: 86% of tags were not found in museum documentation' and 88% of tags were thought by museum staff to be useful for searching (Trant, 2009).

2.2 Issues specific to museum audiences and collections

Internal museum audience research[3] suggests that participatory projects in museums often have lower rates of participation than those in other contexts due to audience concerns around the authority and knowledge required to speak in the same space as the prestigious museum voice. This is exacerbated by the use of technical language and assumption of specialised knowledge evident in some museum interfaces.

Nina Simon, who has written widely on participatory projects in museums, states that participants in museum projects 'need clear roles and information about how to participate', noting that audiences do not need to participate 'in a uniform way or at the same level of commitment'. Simon acknowledges the sometimes-awkward shift required by museums from providing high-quality content to becoming a 'platform' and creating opportunities for audiences to share their own content in 'meaningful and appealing ways' (Simon, 2010).

One potential model for engagement with museum content comes from the UK government's Department for Culture Media and Sport 'Culture and Sport Evidence' programme which defines four types of engagement, each of which builds on the previous type: 1) 'attending'- paying conscious, intentional attention to content; 2) 'participating' - interaction that contributes to the creation of content; 3) 'deciding' - making decisions about the delivery of resources for content creation and 4) 'producing' - creating content 'which has a public economic impact' (CASE, 2011). It could be argued that museum metadata games can include all four types of engagement.

---

[3] Unpublished research, Science Museum Audience Research team

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

2.2.1 Opportunities beyond tagging?

Collections Council of Australia (2009) summarises the impact of institutional history on the reception of an object as it was 'constructed into particular histories, taxonomies and assemblages'. Participatory projects, whether within museums or based on machine-readable access (e.g. through APIs or linked data) to collections have the potential to help constitute new meanings for objects, but this is not yet generally possible as the majority of records are so poor. Crowdsourcing the improvement of collection metadata would therefore make other participatory projects possible.

Crowdsourcing games could also be designed to elicit content beyond tags. As Kaitavuori (2010) states, audiences can share 'a certain kind of hidden knowledge' about objects they've used, particularly for 'industrial design and object design'. This includes historical information otherwise inaccessible to 'museums and academically oriented researchers' about the history and use of objects and 'autobiographical memories' or experiential accounts of the objects in use that could support the interpretation of the objects and provide new perspectives and broader context.

2.3 Casual games and design for participation

The growing genres of 'casual' games - 'games with a low barrier to entry that can be enjoyed in short increments' (IGDA, 2009) are ideal for most crowdsourcing games. Casual game genres include puzzles, word games, board games, card games or trivia games. Features of casual games include easy-to-learn game-play, simple controls, addictive game-play, 'forgiving' game-play with low risk of failure, carefully managed complexity levels with a shallow learning curve and guidance through early levels, and inclusive, accessible themes (IGDA, 2009). Instant game-play can also help turn attention into engagement (McGonigal, 2008). Adams and Rollings (2007) recommend giving casual players 'a sense of rapid progress and achievement' and providing 'emotional rewards for success' and disincentives for failure through text messages, animations and sound. Casual games have potentially huge audiences - the Casual Games

Association (undated) states that over '200 million people worldwide play casual games via the Internet'. However, given the competition for participants, crowdsourcing projects must also be designed to scale up or down (McGonigal, 2008).

The IGDA (2009) recommends casual games should be designed so they '[l]ook like a minimal time investment' as players who feel they can leave at any time will play for longer. Short rounds and frequent 'closure points' can encourage players to keep playing, especially if they feel their progress will be saved. Casual games should need minimal exposition and should offer immediate gratification. Directions and instructions should be very short (i.e. one sentence), and ideally the game should build any instructions or game skill learning into the flow of the game.

2.3.1 Casual game design for women

A 2006 survey from Information Solutions Group commissioned by casual game company PopCap found '76% of casual game players are female, with an average age of 48'. IGDA (2009) reported 'the largest demographic in the casual games audience is comprised of women 30-45 years old', generally playing free online games for 'stress relief' (90% of respondents) while 73% identified 'cognitive exercise (mental workouts)' as a reason for playing.

A survey of 983 gamers (Ghent University, 2010) found that women were motivated to play when able 'to choose their own assignments and to explore freely'. Other factors with strong appeal included 'control, challenge, competition and immersion'. Female gamers prefer clear game rules and feasible challenges, and are 'less motivated to sacrifice time to get the knack of the game', so a game must be playable without preamble.

Malone (1982) found that gender differences should be considered when designing game themes or 'fantasies' ('a system that evokes mental images of physical objects or social situations that are not actually present'). Game themes

'made more difference in the appeal of the game' than simple feedback; this impact can be both positive and negative: 'unless the fantasies are carefully chosen to appeal to the target audience, they may actually make the environment less interesting rather than more'. If themes cannot be designed for all target audiences then game designers should 'provide several fantasies for the same system so that different people can select different fantasies'. Lazzaro (2005) says game mechanics that are about 'helping instead of hurting people' and 'creating joy in life' are appealing to the casual game demographic.

Casual games can benefit from the inclusion of 'emotional in-game characters' who can exhibit a 'variety of emotional states for the player to experience and react to' (Lazzaro, 2005). However, while a 'well-known scenario increases the game's usability and makes it easy to get started' (Lazzaro, 2005), this may offer some challenges for museums, as the nature of individual roles and tasks within a museum may not be widely understood.

2.3.2 Game design process

Adams and Rollings (2007, pp53-83) divide the game design process into three stages. While the build and test processes are iterative, they believe the choice of concept, audience and genre should be decided at the beginning and not changed once design has commenced. They recommend making sure player actions are clear at an early staging, pointing out that if it is difficult to describe the role, it could be difficult for players to grasp it.

In the concept stage, designers consider how gameplay will entertain someone and create a compelling experience. Most of the design details, including look and feel, 'physical, temporal, environmental, emotional and ethical dimensions' that serve and support the gameplay but 'also entertain in their own right' are added in the elaboration stage. Design decisions and core mechanics (the game challenges and actions) are refined through prototyping and play-testing. No new features are added during the tuning stage, but the process of small

adjustments, 'removing imperfections and making the game shine' makes the difference between good and great games.

2.4 Designing games with a purpose

Von Ahn and Dabbish (2008) define requirements specific to GWAP, stating that the rules should:

> 'encourage players to correctly perform the necessary steps to solve the computational problem and, if possible, involve a probabilistic guarantee that the game's output is correct, even if the players do not want it to be correct'.

There must also be 'tight interplay between the game interaction and the work to be accomplished' and iteratively polished task design is important: 'goals that are both well-specified and challenging lead to higher levels of effort and task performance than goals that are too easy or vague' (Von Ahn and Dabbish, 2008).

Usefully for museums with limited design budgets and large collections to cover, the IDGA (2009) report recommends favouring 'a variety of content over a variety of mechanics in a single game'. Citing Sudoku, they report that adding similar content to the same game structure leads the player to 'greater feelings of mastery', as they can apply their existing knowledge of the game mechanics.

Oomen et al (2010b) suggest that the use of specialist interfaces that reinforce the altruistic nature of play increases participation. Discussing the impact of messages about the value of participation in *steve.museum* project, which suggested that 'helping out' a museum may motivate some taggers, they say: 'four times more' works were tagged compared to a general interface. Demonstrating the use of data so that players can see the impact of play is also recommended (Trant, 2009).

McGonigal (2008) states that participants vary according to what they and 'how

much the participant is willing to contribute.' It is important for museum metadata games to match the abilities of the participant - their skills, knowledge and experience - to the activities offered. Importantly for museums looking for content about difficult objects, the IGDA (2009) also recommend games are clear about the knowledge and skills the player will need, establishing any mental requirements immediately by making them integral to the game.

Von Ahn and Dabbish (2008) suggest evaluating GWAP in terms of 'expected contribution' for time of play per player or as 'a combination of throughput and enjoyability'. McGonigal (2008) suggests thinking of participation requirements in terms of 'thought hours' ('a measure of the mental effort required to create something of value') and 'community scope' ('a measure of the diversity of community members the project requires'). Additional considerations such as learning outcomes and engagement with objects may also be an important consideration (CASE, 2011) for museums.

2.4.1 Issues around crowdsourced data validation

Validating the data generated through game-play is a vital part of their design process. Data validation methods include repetition or other player actions within the same game or in the suite of games (Von Ahn and Dabbish, 2008). Discussing the validation issues for 'cultural human activities', Cook (2008) notes the need to 'use other human beings as measurement instruments', such as 'the rating techniques of sites like Hot or Not or Amazon.com'. Validation activities provide tasks to suit players at different levels in the 'pyramid of participation' (McGonigal, 2008) but a comment on Cook's post raises the issue of 'inherent lag' built into 'social games where you use other people as your measurement system' that makes timely feedback difficult: '[y]ou draw a picture and then days later enough people offer ratings'.

2.4.2 Risk, rewards and randomness?

There are various arguments for and against designing for the possibility of failure. Given the issues around fears of contributing to museum projects, it may be best not to make explicit failure a feature of museum crowdsourcing games;

but some level of risk and uncertainty is required in a game (Adams and Rollings, 2007). As Malone (1982) says '[f]or an activity to be challenging, it needs to have a goal whose outcome is uncertain'. Variable levels of difficulty not only help keep the player in flow (the zone between boredom and anxiety), by varying the challenge in relation to skills (Csikszentmihalyi, 1990) but it introduces interesting uncertainty about the outcome. Variable difficulty levels may be evident as 'successive layers of complexity', score keeping, timed responses, high score lists or randomness (Malone, 1982; Von Ahn and Dabbish, 2008, Snow and Barnes, 2010).

Museum metadata games have an advantage here, as the variety of objects and record quality introduces an inherent level of randomness, creating varying levels of uncertainty about successful task completion within the game rules. As Von Ahn and Dabbish (2008) say, '[b]ecause inputs are randomly selected, their difficulty varies, thus keeping the game interesting and engaging for expert and novice players alike'.

Careful reward design helps maximise the amount of data generated during play. For example, Von Ahn and Dabbish (2008) state that points increase motivation 'by providing a clear connection [between] effort in the game, performance (achieving the winning condition), and outcomes'. Evaluation of their GWAPs 'suggests that many players continue playing just to reach a new rank'. However, goals should also be 'multi-level' to motivate 'extended game play—and related data generation'. Brooklyn Museum's game, *Tag! You're It* provides a good model for presenting both a 'local', immediately achievable goal (beat the player ranked just higher than you) and a 'global', long-term goal (beat the highest score). High score lists could be presented geographically (by city, region, country, continent) as well as temporally (hourly, daily, weekly, monthly, etc). Players respond differently to different types of goals (Lazzaro, 2004, Yee, 2006), so displaying a range of potential goals helps motivate more players overall.

The metrics that are rewarded are important (Snow and Barnes, 2010). Is data judged on quality (bearing in mind the issues around data validation) or quantity? If quantity, are rewards allocated for the number of items played or the amount of content generated?

2.5   Designing for participation

As crowdsourcing becomes more popular, there is a growing body of literature on the challenges of engaging mass audiences in the face of competition for 'participation bandwidth' (McGonigal, 2008) and the distribution of participation in crowdsourcing projects. Clay Shirky's claim that there is a 'common power-law distribution across all emerging participatory systems' (Shirky, 2008, in McGonigal, 2008) is borne out by cultural heritage crowdsourcing projects that seem to demonstrate similar 'power laws' in participation rates. Both *Waisda?* and *steve.museum* found that a small number of users - 'super taggers' - contributed the majority of content (Oomen et al 2010a, Trant 2009).

McGonigal (2008) suggests that 'micro-tasks' ('one-off tasks requiring minimal effort') also provide useful actions for the mass of 'individuals at the bottom of the distribution curve'. Designing a seductive initial task 'that can be accomplished quickly and easily' is important: '[i]t is less important at the onset to make something interesting or challenging than it is to make something easy' (McGonigal, 2008). This could include simple data validation tasks such as voting for or 'liking' player-contributed facts or stories, or clicking to report tags for review.

Snoek (2010) describes how *Waisda?* kept barriers to participation very low while designing interactions that lead the user to increase their level of activity: users 'can provide their feedback just by clicking buttons', providing information on the validity of a tag with a 'thumbs up'. They also design interactions that entice the user to increase their level of participation: '[i]f they press the thumbs-down button, the user is asked to correct the label' (Snoek, 2010).

3   Methods

This section summarises the steps undertaken in the project, including genre choice, persona research, platform selection, initial backend build, tagging and fact game designs and evaluation. The game build sections discuss the backend application that delivers the user experience to the screen. The game design sections discuss the screen design, including the user interface and game mechanics as visible to players.

3.1   Game design workshop

The game design workshop aimed to produce workable game ideas, or parts of workable games ('game atoms') that would be suitable for a range of players and produce data suitable for use on a museum collection site.

Specific goals included:

- to design reward systems for various player actions and weight them to encourage data enhancement and satisfying game play
- to consider whether metadata games can become more compelling if social elements are added
- to explore how metadata games can provide enough challenges and variety to keep the player in a state of flow

3.1.1   Participants

Attendees for the workshop were recruited through internal contacts in the Science Museum/NMSI, via the Museums Computer Group (sector specialist) mailing list, and word of mouth on Twitter and Facebook. Participants included curatorial, web and new media staff from the Science Museum, National Railway Museum, Wellcome Collection, a publishing agency and a PhD student from Southampton University. It was important to involve curators and content teams to ensure the games generated useful data that museums would be happy to display alongside traditional collections data.

3.1.2 Structure and activities

The workshop format was an experimental structure devised for the project. It mapped the stages in the game design process (Adams and Rollings, 2007) to a RESCUE-style creativity workshop, particularly the divergence and convergence structure related to idea generation, refinement and validation stages, with reference to creativity theories (Jones et al, 2008).

Preparation and an amount of incubation took place before the workshop, referencing the Poincare model. The participants were sent documents that included a research summary, persona (Appendix C Game Design Persona), a link to the simple tagging activity4 that used sample collection objects (Appendix E Game Design Workshop Sample Objects) and a list of similar projects, as well as the questions they would be asked in the introduction session (see Appendix D Game Design Workshop Briefing document). During the workshop itself the introductory questions were designed to create a shared understanding of the range of play motivations and styles, but their role in the preparation phase was to get participants reflecting on their own definitions of fun, successful, compelling gameplay. The convergence stages were presented through a game-like 'survival of the fittest' meme for ideas. After each activity, the ideas generated were reviewed by the group and the best ideas were taken forward into the next stage.

The activities in the workshop are outlined below. Following the advice of the project supervisor, participant groups for activities were planned in advance and symbols were used to link groups to activities (brainstorming pairs by shape; elaboration by colour; play testing by symbol fill). The workshop was recorded as audio recording, photographs, drawings and short videos. The workshop slides are included in Appendix F.

Preamble (1 hour)

---

[4] http://museumgam.es/tagging-activity-page/
Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

As discussed, the ice breaker question, 'what's your favourite game' was designed to create shared understanding of the range of play motivations and styles and aid preparation (and hopefully future illumination).

This stage also included discussion of the requirements of museum audiences, the characteristics of the objects and the types of crowdsourced data useful for a museum. This helped create a shared understanding of the design constraints for the project.

Concepts brainstorm (30 minutes)

Participants worked in pre-determined pairs to generate as many ideas as they could. Each was given a different design constraint in the form of an assigned class of objects grouped by distinctiveness (e.g. rockets, lockets and sprockets). This formed the concept (game design) phase and was part of the divergence (creativity) phase. Ideas could be expressed as short, high-level descriptions, marketing taglines, etc. The metric for this phase was quantity of ideas, and the pair with the most ideas was rewarded at the end of the activity.
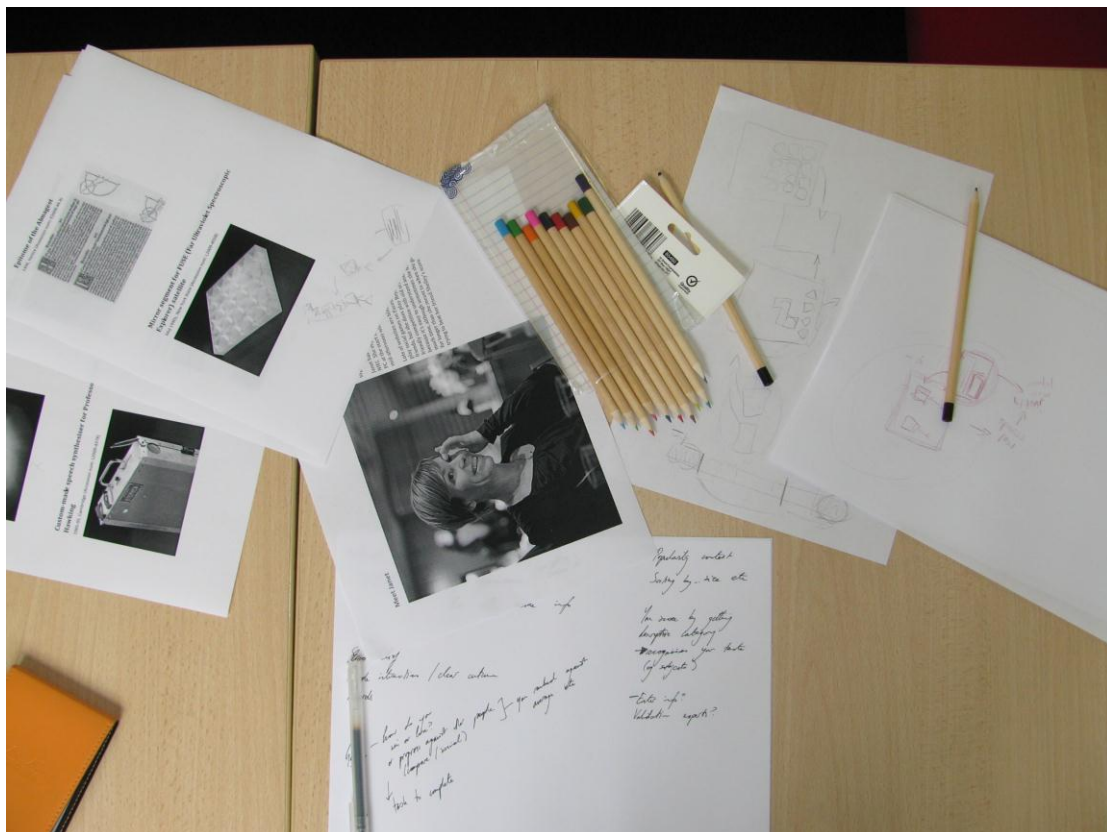
Figure 3 Sample objects, the persona and early participant concepts

Elaborating concepts (two sessions of 45 minutes)

After presenting the results of the concepts brainstorm, participants were put in groups of three and asked to pick a few of their best ideas and explore the data inputs required, the potential data outputs, game play, rewards and challenges. This marked the start of convergence (creativity) and elaboration (game design).

Further convergence occurred during this activity as concepts were presented back to the group for discussion. The lunch break was timed to fall between these sessions to allow space for illumination. In the final session, participants created paper prototypes for use in final stage.

Play testing and further tuning (two sessions of 45 minutes)

Groups of participants ran play-tests with a member from the other group (using the paper prototypes created and refined in the previous sessions. Further

tuning following the tests. This activity supported the verification (creativity) and tuning (game design) phases.

Reflection

After final game design stages, the last session was a reflection on the day and sharing thoughts on workshop structure and activities.
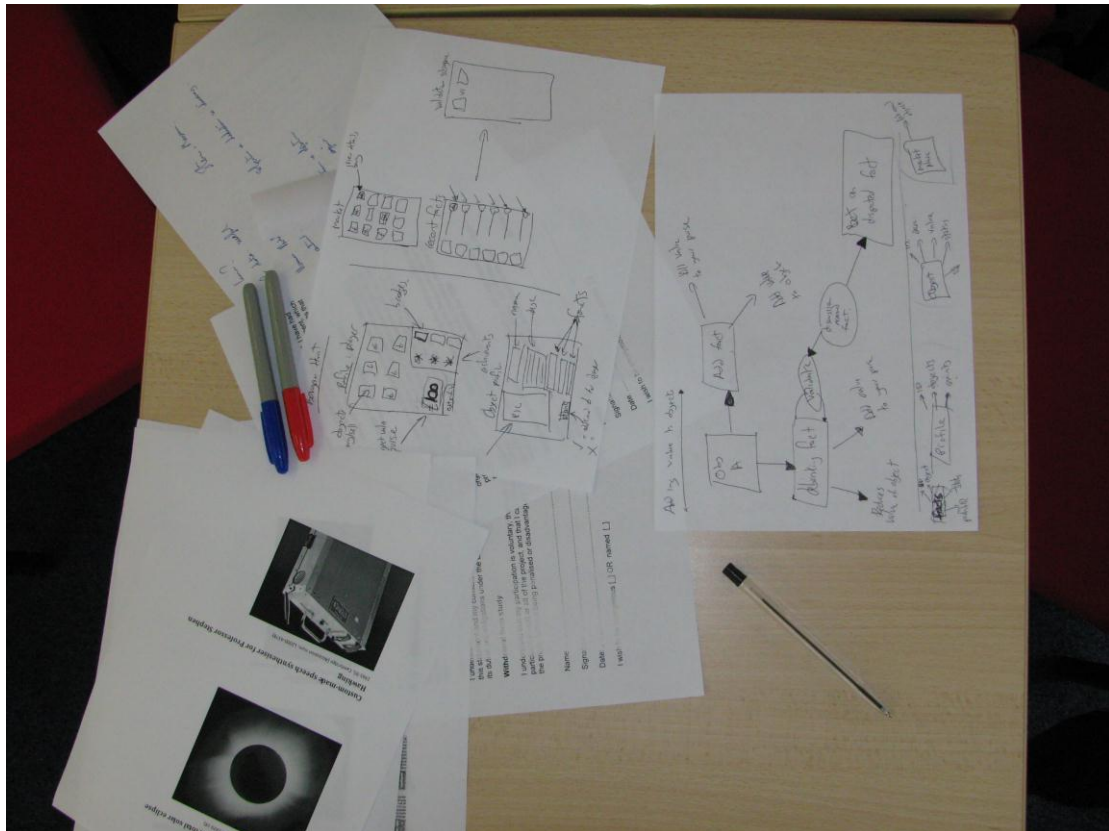
Figure 4 Sample objects and the 'Bargain Hunt' game in development

3.2   Game design

The literature review and analysis of existing sites provided some parameters for the design of fun, casual, browser-based games. The emergent research question around the use of difficult objects programmatically selected from bulk collections provided additional constraints. Further constraints are discussed below.

Other important factors for museum metadata games are the abilities of the

participant - the sum of their skills, knowledge and experience. Audiences may include subject specialists and researchers; people who have helped create or had direct experience using the objects; collectors and hobbyists; or may include those without specific abilities in these areas. When defining projects it is important to consider how much skill is required to recognise, classify and describe objects with more than general physical descriptors (e.g. colour, material, size).

Several early potential games were considered and rejected for failing to satisfy sufficient design parameters. The design process applied insights from evaluating the simple tagging activity, ideas generated through research and review, and ideas from the workshop. Once the game interfaces were built they were iteratively evaluated with play-tests and the designs reviewed.

The overall iterative process took sketched concept designs (see Appendix B), turned them into screen designs, produced backend code to support player interactions, and deployed and evaluated the designs. A review process then lead to the next iteration.

The review of existing crowdsourcing sites showed that tagging works well, even without a game around it and at the most basic level, tagging does not require an investment of time or expertise. It also has immediate benefits for museums in making content more discoverable by providing extra keywords for objects. Early evaluation discovered that potential players are now familiar with tagging from sites like Facebook. During the review of other crowdsourcing projects in the cultural heritage sector, a requirement for registration before play (particularly when the game play could not be previewed before registration), was identified as a particular barrier to participation.

The final design goal, based on the persona and research context became: to create a fun, compelling tagging game that the persona Janet enjoys but justifies playing by telling herself that it is helping a museum.

3.2.1 Genre choice

As discussed in Section 2, research was undertaken into the appropriate game genre for crowdsourcing metadata, with casual games emerging as most suitable genre. Time was also spent time investigating social and casual games (Frontierville, Angry Birds, and Bejeweled) to develop an understanding of the game experiences enjoyed by the target audience and to analyse the implementation of various game mechanics. Points of analysis included the use of narrative characters, methods for providing instructions and hints, and methods for directing player attention and action towards defined goals (such as in-game purchases for these commercial games).

3.2.2 Persona

At the suggestion of the project supervisor, a player persona was created. The research discussed in Section 2 was undertaken into the audiences for casual games to determine the demographic characteristics and motivations of the persona.

The persona, Janet, is presented in Appendix C. Her age, lifestyle, family situation, access to computers and the internet, etc, were chosen to reflect a realistic audience for a casual game. The persona also became a key input into the game design workshop preparation and was designed to help participants relate to Janet and understand her motivations (and excuses) for playing.

3.2.3 Design parameters

In addition to the general constraints for museum metadata games for difficult objects discussed above, there were several constraints on the design process based on the project scope and resources, including: expected low participation rates outside the user testing sessions; lack of resources to market the games to specialists, creating constraints on the skills and experience required of players; the lack of illustrator resources for cartoon characters; the limited graphic design skills of the project investigator; and the potential importance of moderation and data validation if registration was not required.

The following design challenges emerged:

- pitching activities at appropriate level of skills, personal experience and knowledge of general audiences
- audience expectations around 'correct', respectful behaviour in museum environments
- providing varying levels of difficulty from a mass-imported dataset where the challenge afforded by any one object is unknown and the audience's cultural heritage skills remain at a constant level
- allowing players to learn and master new skills while increasing challenge within a game at the right rate to keep them in 'flow' between boredom and anxiety

The necessity for low participation single-player games in particular provided several challenges, including automatically moderating data when a critical mass of players for mutual validation (e.g. tag agreement) or duplicate validation (e.g. tags repeated over time) or flagging offensive terms could not be assumed, and building reward systems that did not rely on extrinsic motivators such as social competition (e.g. dynamic leaderboards) or player feedback (e.g. 'likes') to operate successfully. However, single-player game design also allows the game to be played by specialists and hobbyists (who can bring more expertise and provide greater content coverage) who may not exist in numbers sufficient for more general crowdsourcing games.

3.2.4 Interface design

The aim was to keep key design elements, such as player actions and feedback, visible on the first screen that loaded (on desktop computers), and to provide a simple, compact form design to aid game play.

A simple 960 pixel grid-based framework was devised to provide consistency throughout the project, with the aim that variations in graphic design would not affect the reception of interaction design (see Figure 4, Appendix B).

Wireframes were drawn showing blocks or modules of content, grouping interface elements for related user actions and in-game feedback (see Appendix B, Figure 3). WordPress templates were chosen to support those layout blocks. This ensured consistent page layouts, with defined areas for player actions, progress markers and supporting information. The design also included white space and large fonts with generous line heights, to encourage cognitive accessibility.

The Thematic Theme Framework[5] was chosen because its structure provided the ability to change background colours easily, it supported a 'wash of colour' per game, and it supported WordPress widgets. The website Color Scheme Designer[6] was used to generate complementary colours for different games based on a common 'look and feel' and colour palette.

Once the potential role of narrative characters had emerged, the cartoon characters were created with avatar creation tools designed for social networking sites. Cartoon-style rather than realistic characters were chosen for their ability to reference the current generation of social games on game portals and social networks like Facebook. A range of facial expressions could also be generated quickly.

3.2.5 Initial prototype design

The production of the proof-of-concept 'tagging activity' was timed so it was available online before the workshop so attendees would have a chance to view the types of content the game would work with, and understand the basic dynamics of the activity. The timing of the prototype build also reduced the risk of uncovering serious issues later in the project timeline.

---

[5] http://themeshaper.com/thematic/
[6] http://colorschemedesigner.com/

The 'tagging activity' interfaces were designed to support the core activity with a simple data entry form while excluding game-like design features and interaction dynamics. Minimal feedback was provided for user actions, and progress markers and rewards were excluded. Instructions were written in straightforward but cool language.

Testing was undertaken on this interface with an in-person think-aloud evaluation session to provide a baseline understanding of the potential user satisfaction with tagging tasks.

3.3   Game build

The games were built in PHP, SQL, HTML and CSS as WordPress plugins and themes. WordPress is a PHP-based application originally designed as a blogging platform. The plugins delivered the game functionality while themes managed the presentation layer. The object display code was written as functions shared between the games and 'simple activity' pages. The page logic presented different fields depending on the availability of data such as object title or date and place made, or the length of the object description.

The build process used a local development server, a staging server (http://www.museumgames.org.uk) and a production server (http://museumgam.es).   This tiered structured was designed to support the project's multiple design and testing iterations, by providing continued access to a stable live environment while changes could be developed locally, then tested on the staging server before deployment to the production server. Changes were deployed as updates to the plugin or theme.

A hosted solution was used for code version control, branch and merge management. The project code is available at https://github.com/mialondon/metadata-games. The code written by the project investigator is also supplied in Appendix H, which also includes an

overview of the application structures and functions, alongside Appendix G which documents the underlying database structures.

3.3.1 Data import

Two museum collection APIs were used: the 'Cosmic Collections' API from the Science Museum[7], containing records about objects prepared for display in the 'Cosmos & Culture' exhibition and released for use in a website mash-up competition (Ridge, 2010); and the Powerhouse Museum API[8]. The collections of both museums cover social history and the history of science and technology.

The format of each API was reviewed to devise a shared structure that could support the desired player experience (see Appendix G). Each API was then mapped to the common structure and import scripts were written in PHP and tailored to match each endpoint. As the Cosmic Collections API only contained objects related to astronomy no further selection process was necessary. Objects from the Powerhouse API were selected programmatically (i.e. without manual intervention) on the basis of selected subject searches. The import scripts rejected Powerhouse Museum objects without images.

3.4 Game play-testing and evaluation

The project was evaluated using both qualitative and quantitative techniques. The main evaluation method was play-testing sessions with semi-structured interviews, designed to test the potential of the game designs to support player motivations and the impact of the design on game playability.

Play-tests with think-aloud protocols were conducted on the simple tagging and fact-adding activities, two versions of the tagging game and two versions of the fact game. Some play-test sessions included participants who had tested earlier versions of the games. Semi-structured interviews were conducted once per participant for repeat participants. Participants were recruited through personal networks and were largely a sample of convenience, though an effort was made

---

[7] http://www.sciencemuseum.org.uk/objectapi/cosmosculturepublic.svc/MuseumObjects
[8] http://api.powerhousemuseum.com/

to match play-test participant demographics to the design persona. The semi-structured interview questionnaire was developed iteratively in the early evaluation period.

The questionnaire was also sent to selected players who had registered on the site. Additional feedback was gathered through comments sent by players to the developer and posted on social networking sites such as Twitter and Facebook. Further analysis was supported by data gathered through general game sessions on the http://museumgam.es site.

Table 2 Participants and form of feedback, by interface

Key	Play outside tests?	Demographics	Simple tagging	Simple fact	Dora v1	Dora v2	Donald v1	Donald v2
A	Extensive	female, 30s			Facebook			
F	None	male, 20s			Play-test		Play-test	
H	Extensive	female, 30s	Play-test	Play-test	Play-test	Play-test	Play-test	Play-test (v1.5-2)
J	Extensive	female, 30s				Email questionnaire		Email questionnaire
K	None	female, 30s			Play-test	Play-test (v1.5-2)	Play-test	
M1	Some	female, 30s			Email		Email	
M2	Extensive	female, 30s				Facebook		
N	Some	male, 30s			Email x 2			
Y	None	female, 30s			Play-test			

3.4.1   Online promotion

The games were promoted on personal Twitter and Facebook accounts as these social media channels were a good match for potential online casual game

players. It should be noted that while the analysis period covered two months, owing to extenuating circumstances[9] the games were only actively promoted at the start and end of the period.

---

[9] The project investigator broke an arm

4 Results

The project successfully produced and evaluated two games designed to answer the research questions. Overall the reception of the games was positive, with the games receiving surprising levels of play outside the evaluation session; and the game statistics and qualitative responses show that for some players, the gameplay was compelling. However, many players found that the 'difficult' objects were indeed difficult, particularly in early versions of the games with less clear task prompts.

Encouragingly, for some players, the games were so effective at providing fun supported by altruism that they could validate procrastination. One play tester commented, "I'd play [the game] when I'm supposed to be working. ...[the design] makes me feel like I'm contributing" (Appendix I Play-test and interview transcripts, other responses, p17).

The game versions of the simple activities demonstrate the power of appropriate game themes (Malone, 1982; Lazzaro, 2005). Interestingly, people seemed to relate to Dora's story of losing her data so well that some[10] actually blamed her for the loss or for not having a backup.

The games also demonstrate the usefulness of characters in addressing another design challenge: the issues about authority and fear of 'saying the wrong thing' seemed not to reoccur in evaluation of the game versions, though removing the objects from a typical museum setting probably also had an impact.

4.1 Game design workshop

The game design workshop produced useful results, some of which were

---

[10] For example, http://sosialtmuseum.blogspot.com/2011/01/tester-museumspill.html A loose translation provided by Google Translate reads, 'When I first read the text, I feel immediate sympathy. Heaven, she deleted all the information on his first day at work? Poor Dora! So I am slightly annoyed, why has not taken back up?' or in the original Swedish, ' Når jeg først leser teksten føler jeg umiddelbar sympati. Himmel, har hun slettet all informasjonen på sin første dag på jobben? Stakkars Dora! Så blir jeg litt irritert, hvorfor har ingen tatt back up?'.

incorporated into the final prototypes. The concepts generated helped validate the investigator's earlier game concepts and the discussion during the day provided useful insights. The workshop also provided a useful milestone by creating a focus for delivery of the initial 'tagging activity' and backend build required to support it. The shared understanding generated through constraint brainstorm (Jones et al, 2008) was useful as it provided the basis for suggesting the removal of a constraint for one group.

4.1.1   Game ideas

Concept brainstorm results: the pair of participants thinking of game concepts for the 'rockets' had 4 ideas; those working with 'lockets' had 14 ideas; the 'sprockets' ideas pair had 11 ideas.

4.1.1.1 Game design: *Object Wars*

In *Object Wars*, players are given two objects and asked to pick one. Once they have voted on n pairs of objects, they're given a badge that reflects their choices. While it provides a 'personality quiz through objects' for the player, it generates preference data that could be used to select objects for other games in the ecosystem or to display online retailer-style recommendations for related objects. This game might use tags from earlier games to generate the badge labels or could ask beta-testers to label their collected 'favourites'. As Participant A said, part of the fun is also "in the way it's written, in the museum being irreverent about its objects".
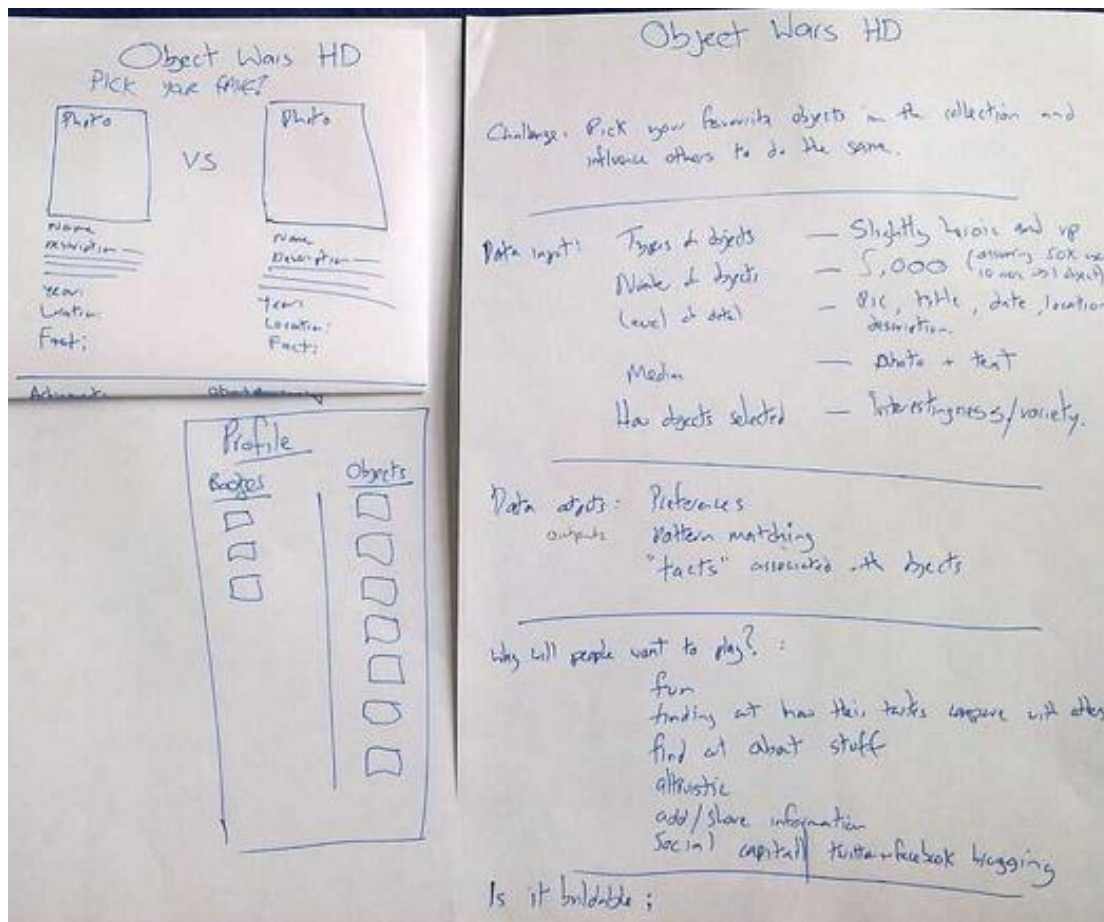
Figure 5 Object Wars overview and design sketch from the design workshop

4.1.1.2 Game design: *Mind Market*

Mind Market is 'a cross between Wikipedia and eBay'. The game has two core gameplay mechanics around which various challenges could be built: play as a 'dealer', getting objects, adding value to them then selling them back into the marketplace; or play as a 'debunker', reviewing recently added facts looking for made-up or misunderstood facts from other players. Facts nominated for debunking are marked as being in 'discussion' mode until the fact is either verified or deleted. The game also supports simple 'likes' on contributed facts. The selection of input objects for this game was complicated by the requirement to assign an initial in-game value for use in trading. Participants felt the ability to choose a category of objects (e.g. related to a subject) might help engage more players. The game would output interesting and validated facts.
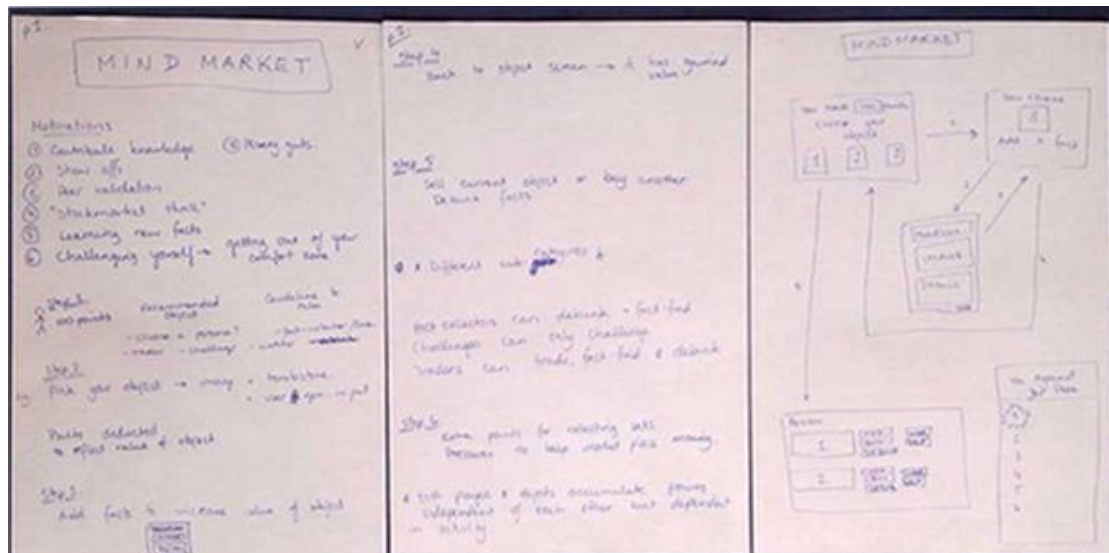
Figure 6 Mind Market overview and design sketch from the design workshop

The *Object Wars* game was more complete as it was less complex. Building this game would have been a realistic output for the project except for the reliance on previously-created tags to generate badge labels, though it would need to be evaluated for compatibility with casual game genres preferences (Lazzaro, 2005). *Mind Market* could be very compelling for some players, and includes interesting activities for data validation ('debunking mode') that could support new types of game-play; but was too complex to build as an entire system within the scope of the project. Both games supported the concept of an 'ecosystem' of games operating on the same objects and user content in the data they required and generated.

4.1.2 Insights from the workshop

Each pair of participants also produced some interesting insights into potential issues for the object 'distinctiveness' type they'd worked on. For 'rockets' the issue was that a lot is known about them already. As one participant said, "basic things like names and places have been researched to hell... a lot of our [ideas are] about qualitative and personal experiences of those rather than raw, hard data about them".

Participants struggled with definition of and boundaries around 'slightly heroic' objects (later 'lockets'), partly because significance is contingent. It may be that the distinction between different levels of objects is only useful as a design tool, and that while the concept is useful for understanding which objects from a particular collection will work in different games, it should not form part of the public interface.

A discussion of the Lewes Chessman raised the issue that 'notability' cannot be assumed if audience is international. This could be particularly important for online games played by international audiences. As Participant T said, "your international identity makes you experience material culture in different ways".

4.1.3 Observations on the process

Participants who had worked on one idea may have found it difficult to switch to tuning and polishing a different game as groups switched members following play-testing.

Some participants who had a lot of experience with video games found it difficult to relate to simpler casual games. It appears that people will angle game-play towards their own preferences and while they could appreciate the persona, they may have found it difficult to relate to her. This diversity in preferences for 'fun' has advantages in that a diverse range of views are represented, but when a persona has already been chosen it may be useful to work with participants who are close to the persona or already familiar with their needs.

A game design workshop is an excellent opportunity to discover points where your audiences might have difficulties, and helps gain an understanding of where assumptions need explaining or even re-visiting.

4.2 Game design

The games are hosted at http://museumgam.es/. During the evaluation period, players were directed to a front page which asked 'Are you ready to play some games with museum objects?' and offered the choice of two games labelled as

'Donald's detective puzzle' (http://museumgam.es/donald/) and 'Dora's lost data' (http://museumgam.es/dora/).

The site included separate pages for a contact form, a combined 'about' and 'Frequently Asked Questions' page and Terms and Conditions. The site also featured examples of previously added content, with a special list of facts used to support *Donald*.

The games were designed to support a 'lazy registration' model, and players could register or login at any time with a form on the right-hand side. The list of pages also appeared in the right-hand menu with a 'leaderboard' of the all-time top scores and 'Share this game' buttons for popular social media sites.

Both the immediate access to gameplay and the tagline 'play games, make museums better' were designed to instantly convey both the fun and the altruistic nature of the site.

The overall site was also iteratively improved during the project. For example, after initial evaluation, a message highlighting the contributions players were making was added: 'So far players like you have improved 335 records for 2 museums through games on this site' and text was added to make it clear that links to a museum object page opened in a new window (Participant M1, Appendix I, p17).

4.2.1 'Simple tagging activity' prototype

This 'tagging activity' provided the core functions necessary for successful tagging, and provided an option for players to skip the object, but it was carefully designed to exclude any game-like elements such as reward systems, progress markers or tutorial modes.

Figure 7 Screenshot of simple tagging prototype

This prototype was evaluated in one play-test before the game design workshop. This participant was an academic in another cultural heritage field, yet found the task intimidating: "What if I write the wrong information? ... I'm nervous, maybe the words I see aren't worthy of being added". Interestingly, the number of tags per object emerged as a key metric for this participant: "It'd be good to have a count of how many words so I could try to beat myself". As would be expected from the literature, this test established a base level of satisfaction with the tagging activity: "there is something compelling about wanting to know what the next object is", but the participant also encountered issues with technical objects: "I personally have no way in, no idea what to describe" (Appendix I, p2).

4.2.2 'Simple fact adding' prototype

As above, a non-game version of the basic 'fact adding' activity was created to help understand the basic dynamics. The user was prompted 'Find an interesting fact to share about this object'. They also had the option to pass on the object, or save the address to come back to the page after completing their research.

Telescope fittings and eyepieces for a 6" telescope.

View object on the Powerhouse Museum site (opens in new window).

(Accession num: H10343)

Thumbnails under license from Powerhouse Museum.
Add a fact about this object
Headline

Find an interesting fact to share about this object.

Fact summary   Source

Add your fact

The headline should 'sell' your fact. The source should provide evidence for your fact.

Tip: save this URL if you want more time to think or research: http://museumgam.es /?page_id=21?obj_ID=473

Not sure about this object? Get a different object.

Figure 8 Screenshot of simple 'fact adding' prototype

While this interface was only tested in one play test, the verdict was damning: "Complicated and a bit scary". When probed, they were worried about "tone, level of detail, reliability of source" but they did feel they knew what kind of fact to add (Appendix I, p3).

4.2.3   Dora

Dora is a tagging game. The player is presented with a character, Dora, who has to replace lost data about a museum's collections. The page presents a random object and asks the player to 'add words to describe the object'.

As the player tags objects, they are given 5 points per tag, and the objects they have tagged are added to a module on the right-hand side of the page.

4.2.3.1Dora v1

The screenshot of the first version of *Dora* shows the 'game activity' module within the wider page and the results of the design process outlined in Section 3. On submission, the player received a simple 'thank you' message from Dora and another object loaded for tagging. The page provided options for bookmarking an object for playing later, or skipping the object without tagging. It also provided links to the originating museum site where there was a specific page for that object.

Figure 9 The first page of Dora v1

Most participants in play-tests grasped the tagging task quickly - "it was obvious that what you wanted me to do was describe the physical properties of [the object]" - and responded positively to the accumulation of points.

One of the main responses to this design in play tests and wider social media commentary (Appendix I) was the lack of validation of tags entered. Some participants also felt the lack of extrinsic social or timed challenge, combined with the repetition of similar objects to make the game feel more of a chore:

> "It feels like I'm being asked to do the same task again and again. It's not a lot of variation. Especially when I got another telescope." (Participant F, Appendix I, p 19).

Another common response was the apparently endless nature of the task and lack of feedback about progress against the assumed goal:

> "Probably by this point, I'd be thinking, 'how many objects do I need to tag', 'is the game going to conclude', or you know, does it go on continuously." (Participant K, Appendix I, p 13).

4.2.3.2 Dora v2

The main change in version 2 was a redesign of the user interface and underlying functionality to add explicit 'sets' of five turns per round to reward players for tagging five objects. The current or 'active' row was highlighted in a different colour, and Dora congratulated the player and summarised their performance after 'completing a level'.

The text used to introduce Dora was revised so that it explicitly introduced the character by name, explained her role - "I'm a junior curator", and contextualised the game task: "I accidentally deleted all the information we were going to add to our collections online. I need to re-label them, and quickly...".

The text directing players' attention to various features was updated based on play-testing results. For example, 'I've added your object to your collection over on the right' was added after evaluation showed that players were not making a connection between their actions and the growing collection of objects on the right-hand side.

The HTML and CSS implementation of the design was tidied and the prominence of different page elements and interactions was reviewed. For example, the page list was moved down the page, and the individual score was moved up.

Dora's lost data

"Hi, my name is Dora, and I'm a junior curator. It's my first day and I've made a big mistake – I accidentally deleted all the information we were going to add to our collections online. I need to re-label them, and quickly...

Can you help? **Add words about the thing in the picture that would help someone find it on Google** – how it looks, what does, who might have used it – anything you can think of."

Scale model (1:10) of Swift gamma-ray burst satellite

Object from: Science Museum.

Date: c. 2000 Place: University of Leicester (Accession num: L2009-4034)

Add words ('tags') to describe this object

Tags [   ]

Tip: separate each tag with a comma, like this: tag, label, a phrase, name, names.

[Tag!]

Not sure about this object? Get a different object. (It won't affect your points.)

Tip: save this URL if you want more time to think or research: http://museumgam.es/dora/?obj_ID=225

Figure 10 Detail of the improved first page of Dora v2

Dora also summarised the player's actions after each turn, offering advice specifically tailored to the number of tags the player had entered and showing them how their tagging compared to others for the same object. This tailored feedback was designed to encourage players to continue tagging and was the

result of evaluation showing that players would appreciate feedback on the types of tags a museum might find useful.
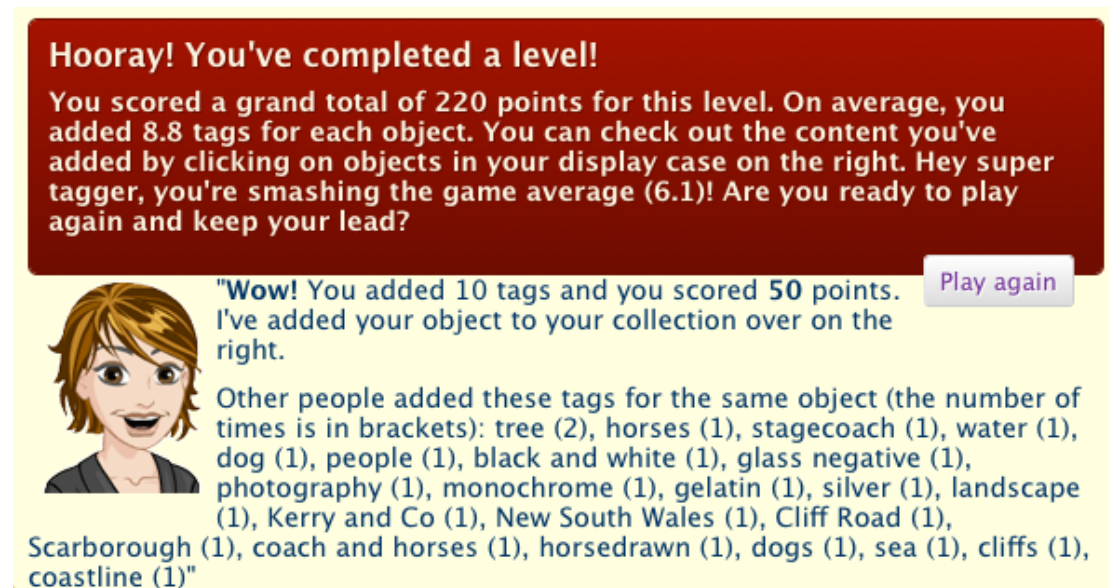
Figure 11 Dora congratulates the player on completing a level and gives them feedback on their performance, both on the set and the individual turn, inviting the player to 'play again'.

The redesign was positively received in play-tests:

> "This is all much clearer and cleaner than before, easier to follow...
> [submits tags]... Wow, Dora got a smiley face!" (Participant H, Appendix I,
> p 32).

4.2.4 Donald

The idea for a game exploring the possibilities for player-created content beyond tagging stemmed from the discovery of other tagging games in the cultural heritage sector such as *Waisda?* and Tiltfactor's suite of games. *Dora* built on existing work to investigate the possibilities for tagging difficult objects, but the project investigator wondered whether more could be done for difficult objects. *Donald* was designed to explore these issues.

As more work was required to research and summarise a fact, the reward structure needed to encourage players to do that work. The two key design

issues that emerged from the design and evaluation process were task (and to an extent, theme) design, and validating content.

4.2.4.1 Donald v1

Part of the design process involved thinking about the types of questions that could be asked as part of the game task. However, while more specific questions provide better tasks, leading to a better game, the more specific a question or game action is the less likely it is to apply to all the objects selected at random. When *Donald* v1 was designed, it was still assumed that audience levels would be very low, and that crowdsourcing the matching of questions to objects would not be realistic.

In the end, the question given to the player was quite general, and the game narrative was written to support a vaguer question. The types of facts gathered through game play and the play tests were analysed for evidence of players' mental models around this type of exercise.   In this way, the game itself was a data-gathering process for future game design iterations.

Fact headlines are limited to 100 characters so they will fit in a 140 character tweet (with a shortened URL). A page bookmark link was created so that players could feel confident about going off to research the object and find the right page again when they were ready to add their fact.

Donald's detective puzzle

"Hello, Holmes! Thank goodness you're here!

Can you help us solve The Case Of The Mystery Objects? The dastardly Moriarty has left behind these objects, but we don't know why. Can you **use the information on this page to find an interesting fact or link about this mysterious object**?

You may need to hunt around for some relevant facts – try searching books or the internet. Then **report back** to Headquarters by filling in the form below. If you succeed, you'll eventually get a promotion for your hard work!

If it's been a while since your last case with us, here's a hint to get you going: if you can't find anything specific about this object, try to find something about the type of object or what it's used for instead.

Pocket terrestrial and celestial globe
Object from: Science Museum.

Date: 1824 Place: London (Accession num: 1936-53)

Tip: copy this URL to come back later if you want more time to think or research:
http://museumgam.es/donald/?obj_ID=288

Mystery object report form
Add the information you've discovered to your report about this object
Headline

The headline should 'sell' your fact to the reader.

Fact summary

Summarise your fact in your own words (short quotes are ok).

Source

Pages
- Welcome
- Dora's lost data
- Donald's detective puzzle
- Tag the object
- Contact
- About this site
- Site terms and conditions

Objects you've tagged in this game

? ? ? ? ?

So far players have added information to **84 objects** through games on this site.

Your Points
- Total : 1275 points
- Donate Points

Your score for this game:
- 1250 points.

Manage your account
- Logout

Top registered players (all games)
- Han (2100 points)
- mia (1275 points)
- yvette (150 points)
- marthasadie (150 points)
- jt (95 points)
- dmje (35 points)
- questions (35 points)

Share this game
- Bookmark on Delicious
- Digg this
- Recomend on Facebook
- Share on netvibes
- Share on Posterous
- Share on Reddit

Figure 12 Screenshot of the opening page of Donald v1

Play-test participants were not overly engaged by the game narrative and the reward structures it offered. This was partly because they assumed that their facts could not easily be validated immediately:

> "text says if you succeed you'll get a promotion, I'm kinda thinking, how are they going to verify what you've said, and I'm guessing that it doesn't bother and it's just 250 points whatever you've said, but I'm kinda intrigued by that ..." (Participant F, Appendix I, p 22).

They found it difficult to find information that related directly to the 'sprocket' object the game randomly assigned them:

> "when it's a dissociated object like memorabilia … it's kind of difficult to know what sort of fact to even start with" (Participant H, Appendix I, p 14).

Examples are particularly important for a task that require a greater investment of player time ("there's a lot of parts to it before you can even do anything"):

> "Can I look at the fun facts that other people have done so I can get an idea of what kind of fun facts I'm meant to be doing?" (Participant H, Appendix I, p 16).

The reward structure - 250 points per fact, compared to 5 points per tag - was a factor in repeated play for some participants:

> "I think I've got over the stage fright of not quite knowing what to do… and I want to make points!"

However, one repeat player (3 facts added) had not signed up to save their points, so perhaps they found the activity rewarding for its own sake.

One participant suggested:
> "there should be someway of selecting the types of objects you get offered … I think then whether you were a student or an expert you'd be more likely to do the research or write what you knew."

4.2.4.2 Donald v2
Evaluation showed that the type and the distinctiveness of the random object 'assigned' to a player had an impact when playing *Donald v1*. *Donald* was redesigned to allow players to choose to play one of 8 random objects from a

summary view with a 'call to action' framing the gameplay: 'Take the fact challenge for this object'. The page design was tidied and the narrative text reviewed so that information was delivered at the point when the player would need it. A link to previous reports was added to support players uncertain about appropriate content.

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

46

Donald's detective puzzle

"Hello, Holmes! Thank goodness you're here!

Can you help us solve The Case Of The Mystery Objects? The dastardly Moriarty has left behind these objects, but we don't know why. Can you **use the information on this page to find an interesting fact or link about one of the things in the images below**?

You may need to hunt around for some relevant facts – try searching books or the internet. Then **report back** to Headquarters by filling in the form below. If you succeed, you'll get **250 merit points** towards a promotion for your hard work!

I've selected some objects at random – take your pick. I know you like to prepare, so this link will open a new window and show you some previous examples of reports submitted.

Toy motor car of a Vanguard Estate

Toy motor car, 'Micro Models', Vanguard Estate, made by Goodwood (Australia) Productions Pty Ltd, Australia, 1952–1961 Toy car model of a vanguard estate car with a diecast zinc two-piece rivetted construction. The body of car has a dark green enam...

Object from: Powerhouse Museum. (Accession num: 91/319)

Image credit: Powerhouse Museum.

[Take the fact challenge for this object]

Toy fish by Mettoy

Toy fish, tin plate, made by Mettoy, Great Britain, c.1955 Toy tin fish, with a yellow and green body and scales printed in black. Fins are coloured in red, green, black and white across the top middle and sides of fish There is a clockwork motor t...

Object from: Powerhouse Museum. (Accession num: 85/2575-64)

Image credit: Powerhouse Museum.

[Take the fact challenge for this object]

Optical instrument for measurement of astronomical photograph plates

Optical instrument, for measurement of astronomical photograph plates, metal / glass / wood, used at Sydney Observatory, designed by H H Turner, made by Troughton and Simms, London, England, 1892–1915 An optical machine for measuring astronomical p...

Object from: Powerhouse Museum. (Accession num: H10140)

Figure 13 Opening screen of Donald v2 showing updated text, design and the ability to select an object for play.

While this version was more positively received, it is clear that for this type of game to become more engaging the task would need to be more closely defined.

One participant pointed out that a lack of information about the real audience (outside the 'magic circle' of the game) would experience the content made the task more difficult:

> "I wasn't sure how the Donald info would be displayed/used in an online museum context, so that made it harder to choose facts and write the headline. I didn't know who the audience was, I suppose. Perhaps the character of Donald could have guided more closely on this." (Participant J, Appendix I, p 46).

4.2.5   Content created through gameplay

Sample content can be seen on the site (http://museumgam.es/content-added-so-far/) or in session reports at the end of play-test transcripts (Appendix I).

The most tagged object was a 'Glass plate negative of view of the Shipard family at Bungowannah near Albury' (Powerhouse Museum) (http://museumgam.es/content-added-so-far/?obj_ID=672) with 76 tags and 1 fact, including tags describing the image content (named people, places, things, items of clothing) subjective descriptions, potential misspellings of given names, and image reproduction technologies.

The fact given on this object also illustrates the way in which some fact entries contain a mixture of personal and 'official' sources. Intriguingly, this entry also seems to include a reference to family history.

Glass plate negative of view of the Shipard family at Bungowannah, near Albury

View object on the Powerhouse Museum site (opens in new window).

(Accession num: 2008/165/1-57)

Thumbnails under license from Powerhouse Museum.
So far people have added content in 10 turns about this object.

Tags

family, dog, photograph, negative, glass plate, Shipard, Albury, verandah, black and white photograph, relaxed group, informal, Albury, family group, glass plate negative, photograph, family, dog, shipard family, Family, Shipard, photograph, adults, male, beards, children, dog, outside, trees, costume, beards, waiscoat, dog, family, shipyard, braces, boots, glass, plate, negative, photography, family, domestic, bungowannah, children, dog, outdoor, portrait, Negative, glass plate, monochrome, Sydney, family portrait, dog, house, children, 1900, Bungowannah, Albury, Joseph Shipard, Arthur Phillips, family, dog, black and white, tree, old, old fashioned, antique, photograph, family, dog, Australia, house, tree, father, children

Headline: Family pictures and social history 1906 Bungowannah

Summary: Photographic glass plates pre-date film and negatives and were common up until the early 1900's and could still be processed upto the 1980's. Photographic prints can then be made from the glass plates. The plate above is held at the powerhouse but members of the family also have copies of the Photgraph. The plate shows Joseph Shipard was born 22/4/1859, died 17/9/1945. Never married. Ephraim Shipard (Teddy) was born 22/10/1866, died April 1938. The children Stanley John was born 17/7/1898, died 1968, and Bessie Maud, born 22/2/1904, died 28/10/1995. All seated outside their home in Bungowannah outside Albany in 1906. The plate negative is part of a larger collection at the museum that shows locations and buildings around Sydney and New South Wales. Importantly it shows the social context and economic history of the time.
Source: Powerhouse museum, Raymond Phillips

Figure 14 Screenshot of the most tagged object, including fact added through Donald.

The object with the most facts was 'Toy model of a tractor made by Matchbox'

(Powerhouse Museum) (http://museumgam.es/content-added-so-far/?obj_ID=587)

The most skipped object was the 'Robilt toy steam locomotive' (Powerhouse Museum) (http://museumgam.es/content-added-so-far/?obj_ID=602), which was passed by 25 players but still tagged in 9 turns.

4.3   Game build

The build process was successful in developing casual, browser-based games on the WordPress platform, in programmatically importing objects from museum APIs for use in metadata games and in supporting the front-end game design. The build produced a WordPress plugin, 'mmg' that contains the core functionality and data structures for the games. It also delivered a related plugin, 'mmg-import' that can be used to populate the games data structures with objects from various collections APIs. The use of museum APIs meant the project was able to create a shared experience with objects from two separate collecting institutions.

The following custom tables were created for the plugin: wp_mmg_turn_facts, wp_mmg_turn_tags, wp_mmg_turns, wp_mmg_game_scores, wp_mmg_objects, wp_mmg_objects_shown. The database structures are documented in Appendix G. The project code (c2700 lines) is in Appendix H.

Installation

Once the plugin is configured, uploaded, installed and activated, the 'shortcode' (a special tag that invokes custom content or functionality when the page is parsed by WordPress) is entered into normal WordPress pages. The related theme must also be installed and activated.

4.3.1   Data import

The following objects were imported to help answer the research questions. The technical and social history objects were leavened with potentially 'empathic' objects or photographs depicting people and/or animals from the social history collections of the Powerhouse Museum.

Table 3 Number of objects by institution and subject

Institution	Type	Number of objects
Powerhouse Museum	social history (112 'toys', 41 'dogs')	153
Powerhouse Museum	technical (astronomy)	100
Science Museum	mostly technical (astronomy)	90
Total		**343**

4.3.2 Disadvantages of WordPress as a platform

One limitation of the games is that players are not logged in automatically after creating an account. The login management plugin used did not support automatic login and the project investigator did not manage to find an appropriate hook in the WordPress codex. This meant the user had to go on to another page after creating an account to have their game score saved against their new username, creating a jarring user experience.

The plugin used to manage player scores also appears to have a bug where the option to 'donate points' cannot be turned off. This interface element confused some players.

4.4 Game play-testing and evaluation

The games were successful in attracting players outside the formal evaluation sessions. Over the evaluation period November 30 - March 1 (the date starts earlier than public release as it includes test data, though not the very earliest prototype data), 196 game sessions were played, with a total of 1079 turns (average 5.51 turns per session); 47 users registered for the site during the analysis period. These 1079 turns created 6,039 tags (average 18 tags per object), 2232 unique tags, and 37 facts for 36 objects.

There were, however, 8 objects that did not gather any user content. There is no immediately obvious difference between these untagged objects and similar objects that did collect content.

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

51

The highest number of turns for a single session was 56, and the average was 1.76 turns per session. In the second version of Dora the average number of turns per session was 2.34. As one early leaderboard winner, a 30-something woman, commented on Facebook, "The tagging is somehow addictive".

The project showed a similar 'power law' rate of participation to that reported by other projects (Oomen et al 2010a, Trant 2009). The majority of sessions comprised fewer turns, but there was a 'long tail' of a smaller percentage of players with a high number of turns per session.
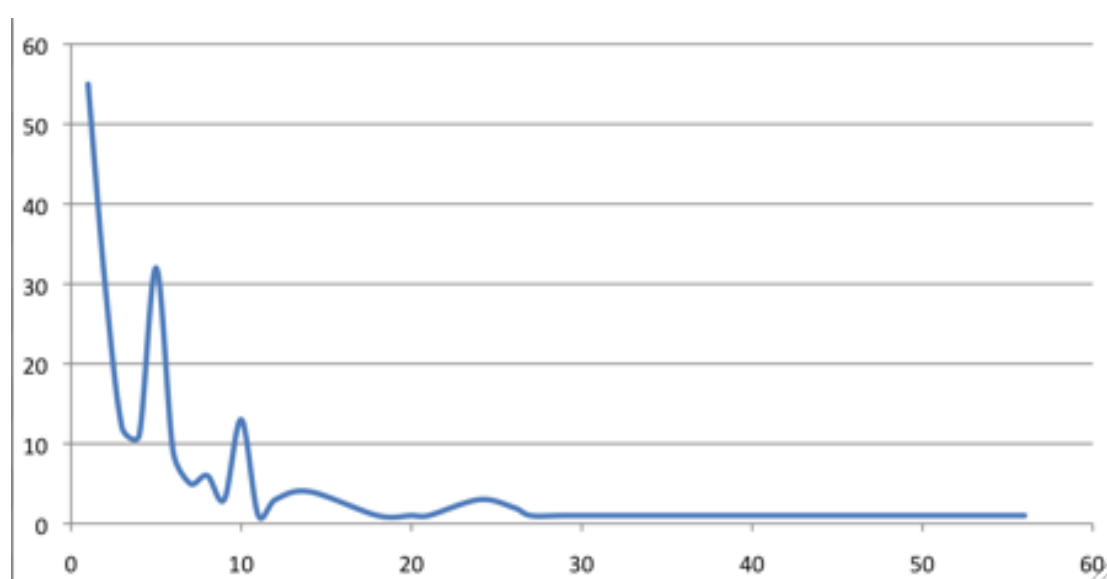
Figure 15 Number of turns (Y axis) by number of sessions (X axis)

The number of facts per player showed the same distribution patterns as the 'super taggers' reported by Oomen et al (2010a) and Trant (2009), suggesting the possible existence of 'super fact finders' if the right games can be designed to engage them.
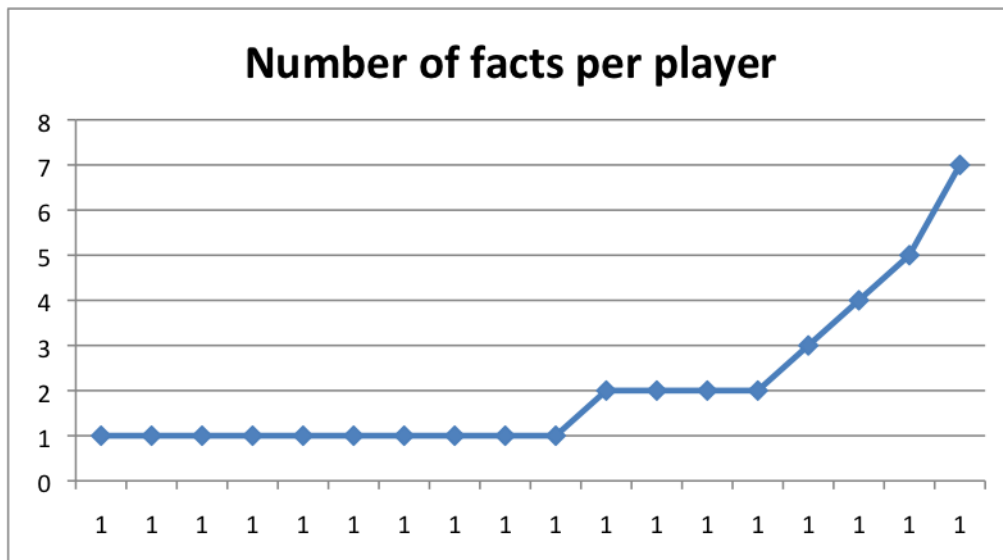
Figure 16 Number of facts per player (by username or session ID)

Table 4 Total turns per interface

Game/interface	Version	Number of turns
simpletagging	NULL	42
funtagging	1	389
funtagging	1.5	70
funtagging	2	541
factseeker	1	16
factseeker	2	21

4.4.1 Website visits

In the analysis period (December 3 - March 1), there were 969 visitors from 46 countries, with 1,438 visits and 5,512 page views on the website.

Figure 17 Number of site visits per day, December 3 - March 1

Average time on site, pages per visit and bounce rate (the percentage of visitors who only visited one page) varied according to visitor intentions (game players or curious peers) over time.

Overall site averages were 3.83 pages per visit, with an average visit length of 3:52 minutes. The bounce rate was 47.64%. When 'bounce' visits were excluded, the average time on site was 07:11 minutes, with 6.45 pages per visit. This suggests a deep level of engagement with the content and activities.

It is worth nothing that during the two-month open evaluation period for the games, only two spam entries were added, and no vandalism was found (though some players tried 'test' entries).

4.4.2 Evaluation results

The results of the play-testing sessions have largely been discussed in the 'design' sections as the design in each iteration responded to the previous version. A few themes were apparent throughout the evaluation:

- players struggled with the repetition of largely-similar objects, but also seemed happy to skip object types they had already tagged. This could be alleviated through a staggered release of similar types of objects, or a greater pool of objects, decreasing the overall probability of encountering similar objects.
- players appreciate clear directions and/or viewing examples before entering content. This may be because they want to help a museum and do not want to add 'bad' data
- some players were happy to tag based on the visual attributes of an object (or by picking out terms in the title, description or web page on the

owning institutions' site); others felt this was inadequate or were not sure they were making a meaningful contribution

- points work for most players, either as a social item (friendly competition with other players) or to mark increased mastery of the task
- a significant number of players would have preferred points to be given on a qualitative rather than a quantitative basis
- related to this, the validation of content was an issue, particularly for *Donald*

4.4.3 Game promotion

The notification messages on Twitter and Facebook were picked up and retweeted widely (mostly by peers - technology or museum people) and shared on Facebook (by random selection of friends).

In an approximately 6 hour period one evening, a call for players on Facebook yielded 180 turns (176 tagging turns, 4 fact turns), 1179 tags and 4 facts on 145 objects (facts weren't requested, but two players started competing for top spot and one used facts as a way of getting points relatively quickly). This may have been in part because the appeal ('Mia has until midnight to get 85 more tagging goes on Dora, can you help'), was bounded by the amount of effort required and had a time-limit, as well as an altruistic appeal.

Dashboard Feb 27, 2011 - Mar 1, 2011

Visits — All Visits — referred_from_facebook Graph by:
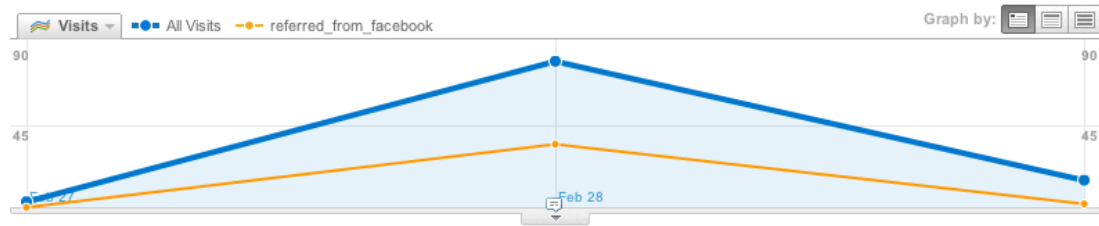
Site Usage

All Visits : **99** Visits
referred_from_facebook : **37**

All Visits : **45.45%** Bounce Rate
referred_from_facebook : **37.84%**

All Visits : **433** Pageviews
referred_from_facebook : **213**

All Visits : **00:05:05** Avg. Time on Site
referred_from_facebook : **00:06:30**

All Visits : **4.37** Pages/Visit
referred_from_facebook : **5.76**

All Visits : **81.82%** % New Visits
referred_from_facebook : **75.68%**

Figure 18 Google Analytics figures showing the impact of a call for players on Facebook. The 'time on site' and 'pages per visit' figures are reasonably high.

5   Discussion of Results

The project investigated the question, 'can metadata games be designed to encourage people to play with 'difficult' museum objects while producing content suitable for re-use?' and found this to be possible. The project demonstrated that metadata games about difficult objects can be fun and compelling experiences. The project investigated the use of content drawn from the entire catalogue and found that this approach was adequate for tagging games when supported by appropriate game design.

The project also explored a supplementary question that emerged after reviewing other cultural heritage metadata projects: can you design games that take crowdsourcing beyond tagging, eliciting content that requires more effort, time or skill than tags without creating additional institutional resource requirements? Initial findings suggest that games such as Donald that ask the player to undertake a more specific task seem to benefit from tasks and activities matched with objects manually. This supports Von Ahn and Dabbish's (2008) statement that 'goals that are both well-specified and challenging lead to higher levels of effort and task performance'.

Donald showed that more advanced activities are possible but would benefit from additional metadata and input into object selection; a possible solution of an 'ecosystem' of games providing the data needed for this is discussed below. The issues around validation for museum metadata games are also discussed below as it appears that cultural heritage collections should not rely on term agreement for content validation when specialist input or long-form content (such as facts or personal experience) is sought unless a critical mass of concurrent players can be achieved.

5.1   Game design workshop

The game design workshop produced useful results. The game designs met the criteria 'fun to play' and 'creating useful metadata'. While one of the final designs

was quite ambitious, and not in itself an easy match for a casual game, many atomic elements of the game ideas produced were eventually used in the final game designs. This project did not have the resources to implement all the reward systems discussed, but the results of evaluation show the value players place on reward systems and progress markers; future work would investigate some of the outputs of the workshop.

The workshop structure and activities were well-received by participants. In particular, some attendees shared positive feedback about the paper play testing. The testing during the design process also provided valuable early 'reality checks' in terms of expected user interactions with typically random objects.

Recommendations for future workshops:
- include more creative stimulus, possibly including board games or props
- include an introductory activity where participants can play a few rounds of the types of games the design persona enjoys. This would help ensure a common understanding of the type of gameplay
- a preparatory exercise on the persona might have been useful, such as asking participants to think of someone they've known who is like the personas
- possibly include participants as close to the target audience or persona as possible to help participants understand their motivations and the types of gameplay that appeal to them
- include one round of play-testing within the design groups so that the group has a shared understanding of their game-play
- reconsider the way the participants were moved between groups - it may be better to let each group continue to work together throughout the afternoon to support constant refinement of the rules towards a shared vision for their game designs.

5.2   Game design

Dora showed that a character and a minimal narrative helped players demographically close to the design persona understand their role in the game immediately. Compared to the control non-game activity, the ability of the character and narrative to invoke the 'magic circle' ('the boundary that divides ideas and activities that are meaningful in the game from those that are meaningful in the real world', entered into when the player decides to play, Adams and Rollings, 2007, pp508) might explain the general lack of concern about lack of expertise or 'saying the wrong thing'.

The game design process and workshop discussion highlighted the tension between the need for mass participation - usually supported by simple, minimal interactions at the lowest level - and the relationship between task complexity and results: the more complex the task, the more immediately usable the resulting data. As workshop participants put it, there is a sliding scale between 'super fun stuff' and generating good data.

The design process also surfaced the conceptual difference between creating content about an individual object, for example, a spark plug, and the concept of spark plugs in general. It can be assumed that tagging is useful for all levels of objects, but for objects at the lowest level, satisfying gameplay for 'fact' games would be better supported by representing them as an object type around which content can be created. The point at which an object gains enough significance or uniqueness-of-data will vary between collections and would be an interesting area for future research.

The review of existing projects showed that registration is a barrier to participation. Compulsory registration before gameplay also violates the requirements for minimal time investment, instant game-play and immediate gratification discussed in Section 2. Game sessions were recorded by IP address and through session ID cookies, allowing the bulk removal of suspicious data if necessary. Where desired, players can be encouraged to register at appropriate

points through prompts such as 'register to save your score' and 'lazy registration' design patterns[11]. The relatively high participation rate (13% of visitors played at least one turn) is unlikely to have been achieved if registration was required to preview or play the games.

Establishing that the games did not expect particular expertise from players was important for these games, but specialist games may want to include a pre-game activity designed to test for the skills or knowledge required.

For museums, curiosity about the next object also contributes to the 'just one more' feeling that increases the number of turns per session. Participants seemed to enjoy spending time engaged with the objects, although repetition of similar objects reduced this effect.

5.2.1   Use of personas in game design
The ideal context of use for the player persona Janet was 'a quick game over a cup of tea and a break', with the games' altruistic purpose acting as her justification for playing. As discussed in Section 4, this was achieved for some players.

Results of evaluation seem to suggest that players closer to the persona in demographic and life circumstances (i.e. women in their 30s and older, with demanding professional and busy personal lives) found the games more compelling, engaging in longer sessions with more turns and higher scores, while players further from the persona (i.e. younger men) were less likely to be engaged by the games. Of the top ten players on the site leaderboard on March 1 (excluding the project investigator), seven were women, two were men and one was of unknown gender; eight are over 30 years of age. The top five players are all women over 30, and generally have domestically and professionally busy lives (source: personal knowledge and play-test questionnaires).

---

[11] http://ui-patterns.com/patterns/LazyRegistration

The results of project suggests that persona-based game design produces results that strongly appeal to players similar to the persona; further, persona-based design could be useful when designing games to appeal to 'super taggers', especially those with specialist skills or experience.

Oomen et al (2010a) reported that while only 2.7% of players contributed more than a thousand tags, 'together they were responsible for adding the largest number of all contributed tags'. The value of the contributions from this kind of 'super tagger' lead then to conclude that cultural heritage tagging projects should find ways to specifically target 'super taggers' while supporting a wide audience. Persona-based design and evaluation is one way of achieving this.

5.3   Game build

The benefits of using WordPress as a platform included: ease of installation and the ability to manage the deployment of updates to plugins or themes; the provision of a page layout and navigation framework; the ability to write extra functionality in PHP as required; built-in user account management and internal data management libraries; and the availability of free plugins and themes that saved some prototype development time. Using an existing platform meant the project began with a robust, documented, security-tested, accessible site. The disadvantage was a reliance on the WordPress codex to provide the 'hooks' needed to attach functionality to user actions in the browser (for example, page loads).

Objects can be imported for use in the game through another plugin, MMG-Import, that queries various museum web services (at the time of writing, the Powerhouse and Culture Grid APIs) and imports the results so the new objects can be played in the games; or through custom SQL imports for collections without APIs. The MMG plugin is available for installation on other WordPress sites.

5.4   Game play-testing and evaluation

The iterative development and testing model proved successful and would be a suitable model for future development.

As discussed in Section 2, crowdsourcing projects require iterative rounds of design and testing with the target audiences to ensure they meet their content goals. Metadata games with difficult objects face particular challenges in helping their audiences feel engaged with the objects and game-play, and able to contribute appropriate content. Play testing was valuable from the first project prototype, and could in general help museums understand the best way to market their game to their target players.

The session data shows that score tables can strongly motivate some players, but the play-test results show the games also provide compelling activities for players who are not motivated by extrinsic factors such as social competition. The implications of this could be explored through further testing, particularly around achievement levels and social play.

The evaluation showed that displaying data already entered helps players understand the types of content other players have provided, and was cited in play-tests as a reason to take more care when entering data. This echoes some of the findings of Oomen et al (2010a, 2010b) and Trant (2009).

Future work on the games would attempt to resolve the issues of specialist data validation (discussed below) to create a more nuanced reward and level structure, and modify the game theme and create more specific tasks for *Donald*. Updating *Donald* after an analysis of the existing data created would be fruitful. For example, making 'source' a repeating field would be useful as currently people enter more than one source per fact, and it allows people to supplement authoritative with personal knowledge. Exploring the implications of this in terms of players' mental models would also be interesting.

5.4.1 Game promotion

The game was 'marketed' sporadically on personal social media channels (Twitter, Facebook) and discussed on museum technologist lists and sites. Interestingly, Facebook referrals show a comparatively high number of page views and longer session times than other social media visits. Excluding bounce visits, Facebook referrers viewed 8.18 pages and spent 09:59 minutes per visit. This may be because Facebook users are comfortable with notifications about new games, or because Twitter referrals from my personal network tended to be curious peers rather than game players.

In future, a 'URL shortener' with parameters embedded to allow traffic from non-web referrers such as twitter and email clients could be used to provide more precise information on the source of website traffic. This could in turn be used to improve site performance and understand the audience.

6 Evaluation, recommendations for further study, conclusions

6.1 Evaluation against original project definition

The original proposed focus of this project (Appendix A) was the design and evaluation of interfaces and interactions applied to online museum collections in order to encourage members of the public to undertake specific tasks that will help improve the quality of the experience for later visitors, and the development and evaluation of interaction models for crowdsourcing metadata enhancement.

The original problem was refined during the project, in part in response to the discovery of a number of tagging projects which seemed to satisfy the original basic question, and in part as a result of iterative design testing and analysis of the emergent gameplay around the sample objects.

The original project objectives were:

- Design game-like interaction models applicable to cultural heritage content and audiences through research, analysis and creativity workshops
- Build an application and interfaces to implement the interaction designs and to create and store user-created content linked to collections content
- Evaluate the effectiveness of game-like interaction models for eliciting useful content

As this report shows, these objectives were met. The project was also able to answer the redefined research questions. Some interesting ideas for further research are discussed below.

The original project plan included interviews, but during the review of other

Mia Ridge, 'Playing with difficult objects: game designs for crowdsourcing museum metadata' (unpublished Masters Dissertation, City University, 2011)

64

museum crowdsourcing sites and games it became apparent that interviews with project staff would probably not provide the insights desired, particularly for public citation.

6.2   Impact

The games have been well-received within the museum sector, and many museums were happy to share the link with their followers on social media: for example, the Getty Museum tweeted the link[12] to their 150,000 twitter followers and closer to home, the Birmingham Museums & Art Gallery tweeted it to their 2,700 followers.

Originally, the built game designs were intended to be rapid, discardable prototypes, and the primary product of the project was to be the research and evaluation. However, there has been on-going interest from UK, European and American institutions in hosting their own version of the games or in contributing records to http://museumgam.es. The project investigator has been invited to a 'hackathon' with access to a Europe-wide collections API (Europeana) in order to continue research into museum metadata games.

The project research outputs have also had international impact. The project investigator was interviewed about the research and use of the Powerhouse Museum API for the museum technology blog, Fresh + New(er)[13], by Seb Chan, Head of Digital, Social & Emerging Technologies at the Powerhouse Museum, who said: "I'm hoping we can use Mia's findings to help us design better minigames in our new collection database, and I'm also hoping others, especially those outside of the museum community, will use her findings to build better games..."

In a conclusive demonstration of the potential usefulness of the project, the Heritage Lottery Fund (HLF) nominated the project as a case study of 'outstanding digital practice in the heritage sector in the UK and internationally'.

6.3   An ecosystem of museum metadata games?

The potential for a model for applying different types of games to different content gathering and validation requirements emerged during the research and development phases of the project. Other specialist metadata projects have uncovered a similar requirement: '[h]aving a suite of games enables database managers to custom link to the most effective and appropriate game front ends for their data' (Tiltfactor, undated).

For example, the process of matching potential tasks to objects (or classes of objects) could also be crowdsourced. The object selection process can also benefit from data gathered in other games, including additional content to support the catalogue entry or by dropping out objects that do not pass a threshold limit (e.g. for the number of times an object was skipped or excluding objects that gathered three or fewer tags per turn).

Von Ahn and Dabbish (2008) describe three GWAP 'templates' (output-agreement games, inversion-problem games, and input-agreement games) that 'can be applied to any computational problem' that could be usefully investigated in this context.

6.3.1   Player-contributed data lifecycle in an ecosystem of games

The points at which player-contributed data can be passed between games as input into game-play, for validation in other games, or into another game are outlined below.

Content is created about objects in the game; the content is validated; a game-dependent value is assigned to the content; and the player is rewarded. The value of a piece of content may also be validated (e.g. for 'interestingness') when

other players show preferences for it. At this point, the object and the new content about it can be used in a new game or presented on a collections page.

For some content types, the content may be validated by players in another game after a default value has been calculated and the player has been rewarded. If the content is found to be inaccurate or otherwise invalid, the reward can be revoked.
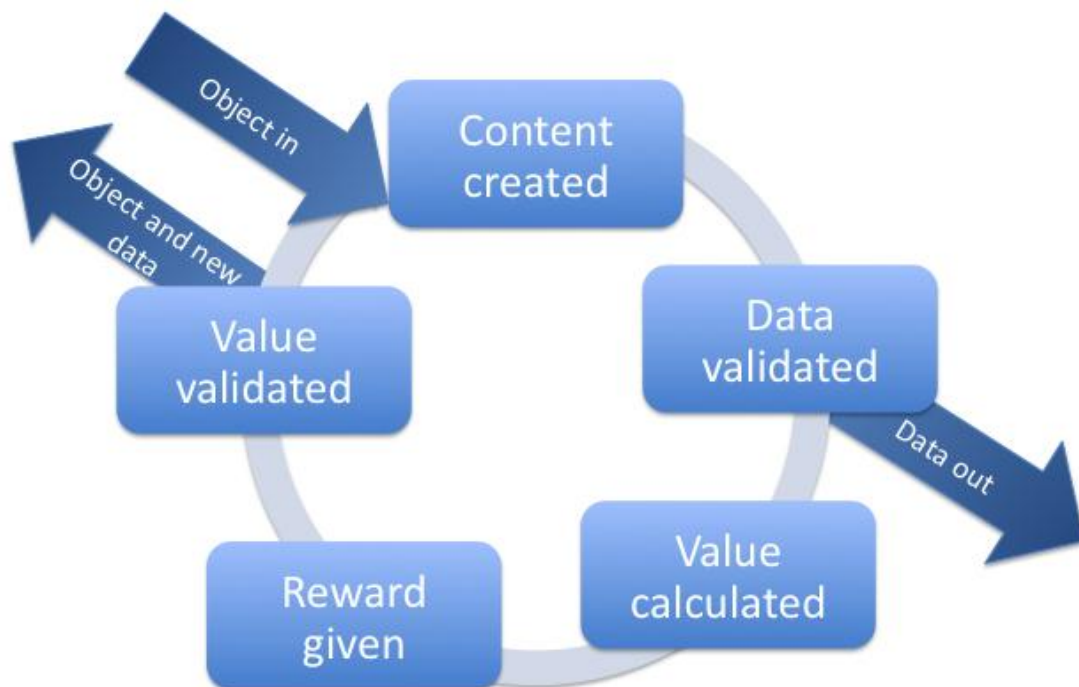
Figure 19 A model for the lifecycle of player-contributed data within a game

6.3.2 Activity types and data generated

The activities listed below can be applied to museum objects, and built into games through the design of rules and concepts or themes. They could also be mapped to typical game challenges (e.g. Adams and Rollings, 2007, p23). The type of data input required will depend on the collection and 'distinctiveness' of the object. Each activity listed assumes the presence of an image or access to the original object.

The folksonomies generated by tagging activities could be used to generate labels for classes of objects for use in higher level games (e.g. the concept of spark plugs rather than 128 individual spark plugs[14]) to address some of the issues found with repetition of similar objects that affect the playability of games like *Donald*. Other activities could also help match objects with appropriate tasks and narratives for 'fact' games, or solicit potential tasks or questions about objects as free-text input; again, this could have a substantial impact on the player experience of more complex games.

Table 5 Activity types for museum metadata games

| Activity	Data generated	Validation role, requirements
Tagging (e.g. Brooklyn Museum *Tag! You're It*, variations include two-player 'tag agreement' games like *Waisda?*, extensions such as guessing games e.g. GWAP *ESP game*, *Verbosity*, Tiltfactor *Guess What?*; structured tagging/categorisation e.g. *Verbosity*, Tiltfactor *Cattegory*)	Tags, folksonomies, multilingual term equivalents. General; specialist. Objective; subjective.	Repeated tags provide validation through agreement. Some automated validation on common terms.
Debunking (e.g. flagging content for review or researching and providing corrections).	Possible provides corrected data to replace erroneous data. General; specialist. Objective only.	Can flag tags, links, facts for review. Should not be used on subjective personal stories.

Recording a personal story	Oral histories; contextualising detail; eyewitness accounts. Generally subjective.	Moderation should be handled sensitively.
Linking (e.g. objects with other objects, objects to subject authorities, objects to related media or websites).	Relationship data; contextualising detail; information on history, workings and use of objects; illustrative examples. General; specialist. Generally objective.	Can be validated through repetition, preference selection, debunking, or through recording links followed if used on collections sites.
Stating preferences (e.g. choosing between two objects; voting; 'liking'; GWAP Matchin, workshop Object Wars).	Preference data, selecting subsets of 'highlight' objects or 'interestingness' values for different audiences. May also provide information on reason for choice. Generally subjective.	Can help validate most forms of data, though the terms on which items are being valued should be considered.
Categorising (e.g. applying structured labels to a group of objects, collecting sets of objects or guessing label or relationship between presented set of objects).	Relationship data; preference data; insight into audience mental models; set labels.	Repeated labels or overlapping sets provide validation through agreement.
Metadata guessing games (e.g. guess which object in a group is being described)	Tags; structured tags (e.g. 'looks like', 'is used for', 'is a type of')	Successful clues provide validation through agreement.
Creative responses (e.g.	Relevance,	Consider acceptable

write an interesting fake history for a known object or description of purpose of a mystery object.)	interestingness, ability to act as social object; common misconceptions.	levels of criticism from other players.

6.3.3   Pyramids of participation in museum metadata games

The model below is based both on the literature review of participatory experience design conducted for Section 2 and on the results of the design, build and evaluation in Sections 3, 4, and 5. This could be related to the activities described above as well as various player motivation models (e.g. Lazzaro, 2004; Yee, 2006).

It appears that the abilities of mass participants maps to the needs of museums in terms of object coverage. For example, every object can benefit from tags, and tagging is the easiest activity with the highest participation rates of the activities discussed for this project. However, it is not a productive use of player time to write a fact or story for largely duplicate sprockets, but it is less overwhelming to write about a class of sprockets (though this may bring new issues around writing about the subject of sprockets rather than sprockets in the context of a museum collection). It would make sense to focus players' attention on the objects with greater significance or distinctiveness, but this requires more specialist skills, reducing the size of the potential audiences. Testing the utility of this model could be a useful area for further research.
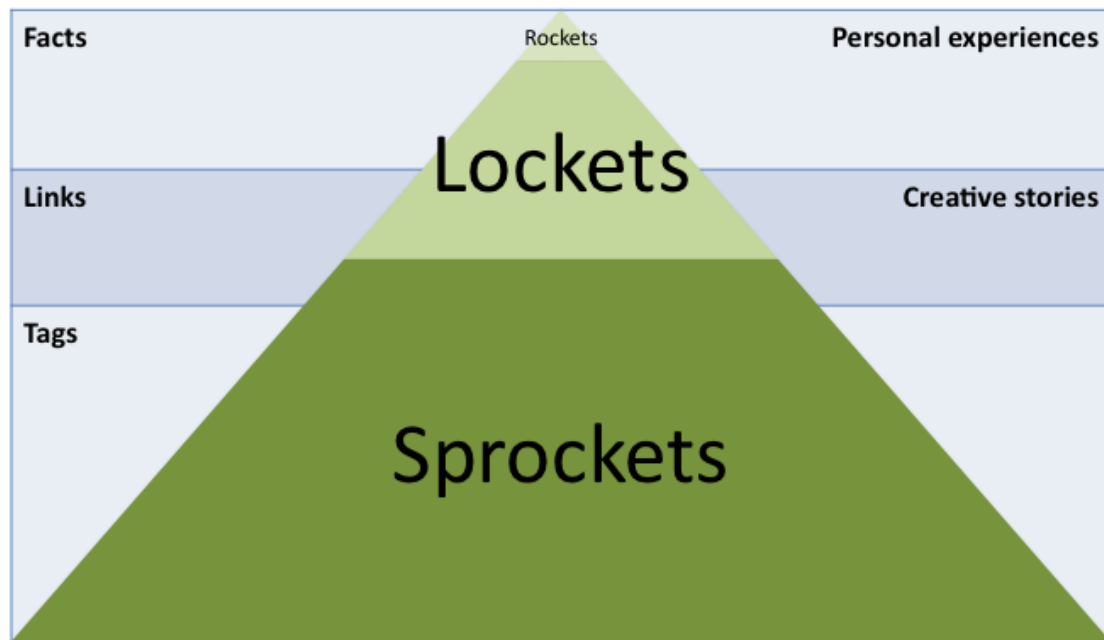
Figure 20 Possible mapping between object 'significance' and highest potential activity

6.3.4 Validating unusual data

The project games have a simple reward structure, with a set number of points assigned to each tag or fact added. Some play-test participants expressed a desire for a metric based on tag or fact quality, but implementing this was beyond the scope of the project.

The ideal model for these games would place a high reward value on valid and unusual tags, a medium value on common tags, a lower value on tags that are visible in the object metadata, and would penalize invalid nonsense words. As players explore the game, they will test the boundaries of the reward system with nonsense and irrelevant words, and it is important to detect these experimental forays and only reward appropriate tags.

While matching tags from independent players is a common 'win condition' that provides validation through agreement on tags about a media item; and while 'taboo' words can be introduced to encourage players to find more granular descriptions and encourage greater coverage in tagging (e.g. Von Ahn and Dabbish, 2008, and Tiltfactor's 'overused' words), this does not provide a

solution for validating specialist tags or a model for validating other types of long-form content such as facts. Museums, like archives, 'require specialized, precise information in order for their contents to remain relevant or useful to the people who wish to use them' (Tiltfactor, 2010a).

Tags that reference specialist knowledge are less likely to have been previously entered in the game and are therefore penalized under models that validate tags against a corpus of previous content. The more precise or specialist a term, the greater the likelihood it will be rejected as a nonsense word, yet previously unknown tags could contain valuable information such as: specialist knowledge related to the collections; corrections of common misidentifications or stereotypes; languages other than English, or be informed by particular domain knowledge or personal experience of the objects.

Asynchronous (i.e. deferred) manual review (where other players validate the tags applied by previous players) requires a critical mass of players or institutional resources, and introduces awkward issues of delayed gratification or penalties for points validated or rejected after the play session. Automated solutions for detecting and rewarding specialist terms without allowing players to 'cheat' the game by entering nonsense or irrelevant words would be a useful area for future research. The lack of a qualitative validation model for long-form or specialist data is a barrier to the wider success of metadata games.

6.4   Further research

6.4.1   The impact of record quality and object distinctiveness on games

The 343 object records imported into the project dataset were analysed to provide a subjective rating of various aspects of record quality, including the type and length of description (technical, descriptive or contextual history; less than x or greater than y words); quality of the image (colour versus black and white; reference- versus publishable-quality); relative age of the object, and whether it could support an 'empathic' viewing (assuming that images that show people have a certain universality regardless of the audience's knowledge of

time and location). Analysing the potential correlations between this subjective rating of 'record quality' and the amount and quality of in-game data created is beyond the scope of this project. However, it would be an interesting area for further research. It may also be worth investigating potential correlations between the 'distinctiveness' model outlined in this paper, in-game data and record quality.

6.4.2   New relationships with specialists

Following the discussion of specialist data, validation models for data beyond the 'lowest common denominator' (Tiltfactor, 2010a) would be extremely useful future research.

A critical mass of available objects creates opportunities for specialist metadata gathering while forging new relationships with audiences close to a museum. Von Ahn and Dabbish (2004) suggest 'theme rooms' where self-selecting players can choose to play with images 'from certain domains or with specific types of content', in turn generating more specific tags. Potential specialist audiences include subject specialists, students and researchers; people who have helped create or had direct experience using the objects; collectors and hobbyists. This is one solution to the issue of specialist content validation discussed elsewhere, and also creates opportunities for different forms of cooperative and competitive game-play.

However, museums need to consider the requirements for adequate collections data as input into the games, validation requirements, specialist marketing and outreach, and the design issues around signalling the competencies required to play the games. For example, how much skill is required to recognize, classify and describe objects with specialist terms? Is the size of the potential audience large enough to justify the resources required, or do they already have access to sites that meet similar needs? If personal experience of the objects is within living memory are there any issues with the age or technological skills of the audience?

6.5 Conclusions

In summary, the project produced two games, *Dora*, a single-player tagging game, and *Donald*, an experimental single-player 'trivia' game. The production and evaluation of the tagging game Dora addressed the primary research question and showed that games can be designed to encourage people to play with 'difficult' museum objects while producing content that will improve museum websites.

Donald, a more experimental game that explored the design issues around longer forms of content that required some form of research or personal reference, showed that game can be designed to take crowdsourcing beyond tagging.

Analysis of the results lead to the proposal for an 'ecosystem' of metadata games in which objects can be incrementally improved and selected for play in more advance games, and in which player-contributed content can be validated for correctness and 'interestingness'. Related to this, future work on validating long-form and specialist content would be useful.

7   References

Adams, E., and Rollings, A. (2007). *Fundamentals of Game Design*. New Jersey: Pearson-Prentice Hall.

Black, G. (2005) *The engaging museum : developing museums for visitor involvement*. Routledge

Casual Games Association (undated) *f.a.q.* [Online] Available from: http://www.casualgamesassociation.org/faq.php#casualgames [Accessed March 6, 2011].

Coldicutt, R. and K. Streten, Democratize And Distribute: Achieving A Many-To-Many Content Model, in J. Trant and D. Bearman (eds.). *Museums and the Web 2005: Proceedings*, Toronto: Archives & Museum Informatics. [Online] Available from:
http://www.archimuse.com/mw2005/papers/coldicutt/coldicutt.html [Accessed February 21, 2011]

Cook, D. (2008) *What actitivies [sic] can be turned into games?* [Online] Available from: http://www.lostgarden.com/2008/06/what-actitivies-that-can-be-turned-into.html [Accessed December 30 2010]

Culture and Sport Evidence (CASE) programme. (2011). *Evidence of what works: evaluated projects to drive up engagement.* [Online] Available from:
http://www.culture.gov.uk/what_we_do/research_and_statistics/7290.aspx
http://www.culture.gov.uk/images/research/evidence_of_what_works.pdf
[Accessed February 20, 2011]

Department of Media, Culture and Sport (DCMS) (2010) *Principles of Selection for Listing Buildings* [Online] Available from: http://www.english-heritage.org.uk/content/imported-docs/p-t/principles-of-selection-for-listing-buildings-2010.pdf [Accessed 4 March 2011]

Deterding, S. (2010). 'Just add points? what ux designers can (and cannot) learn from games' *#UXCamp* Europe Berlin, May 30, 2010 [Online] Available from: http://www.slideshare.net/dings/just-add-points-what-ux-can-and-cannot-learn-from-games [Accessed July 22, 2010]

Howe, J. (2006a). 5 Rules of the New Labor Pool. Wired 14.06. [Online] http://www.wired.com/wired/archive/14.06/labor.html [Last accessed February 23, 2011]

Howe, J. (2006b). The Rise of Crowdsourcing. *Wired* 14.06. [Online] Available from: http://www.wired.com/wired/archive/14.06/crowds.html [Accessed October 8, 2010].

Ghent University (2010) *Girl gaming on the rise*. Press release. Published 5 July 2010 [Online] Available from: http://www.alphagalileo.org/ViewItem.aspx?ItemId=80450&CultureCode=en [Accessed 20 July 2010]

Information Solutions Group. (2006). *Women Choose "Casual" Videogames Over TV; 100 Million+ Women Now Play Regularly, For Different Reasons Than Men*. [Online] Available from: http://www.infosolutionsgroup.com/pdfs/women_choose_videogames.pdf [Accessed August 6, 2010].

International Game Developers Association (IGDA). (2009). *2008-2009 Casual Games White Paper*. [Online] Available from: http://www.igda.org/sites/default/files/IGDA_Casual_Games_White_Paper_200 8.pdf [Accessed December 31, 2010].

Jones, S., Lynch, P., Maiden, N.A.M. and Lindstaedt, S. (2008). "Use and Influence of Creative Ideas and Requirements for a Work-Integrated Learning System", in *Proceedings RE08, 16th International Requirements Engineering Conference*, Barcelona, Catalunya, Spain, September 8-12, IEEE Computer Society Press, pp 289 - 294, 2008.

Karvonen, M. (2010). "Digitising Museum Materials - Towards Visibility and Impact". In Pettersson, S., Hagedorn-Saupe, M., Jyrkkiö, T., Weij, A. (Eds) *Encouraging Collections Mobility In Europe*. Collections Mobility. [Online] Available from: http://www.lending-for-europe.eu/index.php?id=167 [Accessed February 26, 2011].

Koster, R. (2005). *A Theory of Fun*, Paraglyph Press, Scotsdale, Arizona.

Lazzaro, N. (2004). Why We Play Games: Four Keys to More Emotion Without Story In *Game Developers Conference* [Online] Available from: http://xeodesign.com/xeodesign_whyweplaygames.pdf [Accessed October 17, 2010]

Lazzaro, N. (2005). *Diner Dash and The People Factor.* [Online] Available from: http://www.xeodesign.com/xeodesign_dinerdashcasestdy500n031405.pdf [Accessed August 13, 2010]

Malone, T. (1982). Heuristics for designing enjoyable user interfaces: Lessons from computer games. In *Proceedings of the 1982 conference on Human factors in computing systems* (CHI '82). ACM, New York, NY, USA, 63-68. DOI=10.1145/800049.801756 http://doi.acm.org/10.1145/800049.801756

McGonigal, J. (2008). *Engagement Economy: the future of massively scaled collaboration and participation.* [Online] Available from: http://www.iftf.org/system/files/deliverables/Engagement_Economy_sm_0.pdf [Accessed June 10, 2010].

Oomen, J., Verwayen, H., Timmermans, N., and Heijmans, L. (2009) Images for the Future: Unlocking the Value of Audiovisual Heritage. In J. Trant and D. Bearman (eds). *Museums and the Web 2009: Proceedings*. Toronto: Archives & Museum Informatics. Published March 31, 2009. [Online] Available from: http://www.archimuse.com/mw2009/papers/oomen/oomen.html [Accessed February 20, 2011].

Oomen, J., Brinkerink,, M., Heijmans, L., van Exel, T. (2010a) Emerging Institutional Practices: Reflections on Crowdsourcing and Collaborative Storytelling. In J. Trant and D. Bearman (eds). *Museums and the Web 2010: Proceedings*. Toronto: Archives & Museum Informatics. Published March 31, 2010. [Online] Available from: http://www.archimuse.com/mw2010/papers/oomen/oomen.html [Accessed February 20, 2011].

Oomen, J., et al (2010b) *Emerging Practices in the Cultural Heritage Domain - Social Tagging of Audiovisual Heritage.* In: Proceedings of the WebSci10: Extending the Frontiers of Society On-Line, April 26-27th, 2010, Raleigh, NC: US.

Ridge, M. (2010). Cosmic Collections: Creating a Big Bang. In J. Trant and D. Bearman (eds). *Museums and the Web 2010: Proceedings. Toronto: Archives & Museum Informatics.* Published March 31, 2010. [Online] Available from: http://www.archimuse.com/mw2010/papers/ridge/ridge.html [Accessed March 31, 2010].

Russell, R., and Winkworth, K. (2009). *Significance 2.0: a guide to assessing the significance of collections.* Collections Council of Australia. [Online] Available from: http://significance.collectionscouncil.com.au/ [Accessed September 9, 2009].

The Nielsen Company. (2009) *Insights on Casual Games: Analysis of Casual Games for the PC.* [Online] Available from: http://blog.nielsen.com/nielsenwire/wp-content/uploads/2009/09/GamerReport.pdf [Accessed February 23, 2011].

Schell, J. (2008). *The Art of Game Design: A book of lenses*. Morgan Kaufmann

Simon, N. (2010). Principles of Participation in *The Participatory Museum* [Online] Available from http://www.participatorymuseum.org/ [Accessed October 24, 2010]

Snoek, C.G.M., Freiburg, B., Oomen, J., Ordelman, R. (2010). Crowdsourcing rock n' roll multimedia retrieval. In *Proceedings of the international conference on Multimedia*. ACM, New York, NY, USA„pp. 1535-1538. [Online] Available at http://doi.acm.org/10.1145/1873951.1874278 [Accessed February 19, 2011]

Snow, S. and Barnes, P. (2010) Use Game Mechanics to Power Your Business [Online] Available at http://mashable.com/2010/07/13/game-mechanics-business/ [Accessed September 9, 2010]

Tiltfactor (undated). *Metadata games.* [Online] Available from: http://tiltfactor.org/?page_id=%201279 [Accessed 19 November 2010].

Tiltfactor (2010a) *Metadata Crowdmining*, February 25, 2010 [Online] Available from: http://www.tiltfactor.org/?p=1446 [Accessed 30 August 2010]

Tiltfactor (2010b) *Metadata Investigation*, continuing [Online] Available from: http://www.tiltfactor.org/?p=1543 [Accessed 30 August 2010]

Trant, J. (2009) *Tagging, Folksonomy and Art Museums: Results of steve.museum's Research.* In: Archives & Museum Informatics: 2009. [Online] Available from: http://conference.archimuse.com/files/trantSteveResearchReport2008.pdf [Accessed June 10, 2010].

Yee, N. (2006) The Psychology of Massively Multi-User Online Role-Playing Games: Motivations, Emotional Investment, Relationships and Problematic Usage. In: Schroder, R. & Axelsson, A., (eds) *Work and Play: Collaboration and Interaction in Shared Virtual Environments*. London: Springer-Verlag. [Online] Available from: http://www.nickyee.com/pubs/Yee%20-%20MMORPG%20Psychology%20(2006).pdf [Accessed February 18, 2011]

Von Ahn, L., and Dabbish, L. (2004). Labeling Images with a Computer Game in *CHI 2004*, April 24–29, 2004, Vienna, Austria.

Von Ahn, L., and Dabbish, L. (2008). Designing games with a purpose. *Commun. ACM 51, 8* (August 2008), 58-67. [Online] Available from: http://doi.acm.org/10.1145/1378704.1378719

Appendix A Project Definition Report

Game mechanics and interaction models for social good: a case study on crowdsourcing museum collections enhancement

Mia Ridge

Supervisor: Dr Sara Jones

June 2010

Problem

Museum collections site, whether object catalogues or thematic sites with interpretative content, fail to achieve levels of public usage commensurate with the resources taken to create them. Many collections sites lack discoverability or fail to engage the casual or non-specialist visitor who find themselves on a collection page. Research (Trant, 2006) has shown that visitor-created content can improve the accessibility of museum data. How can websites encourage visitors to create content to help others when they may not immediately benefit from it themselves?

Project description

The primary focus of this project is the design and evaluation of interactions applied to the context of an online museum collection in order to encourage members of the public to undertake specific tasks that will help improve the website.

The project will include a design and build component to create game-like interfaces for testing and evaluation, but the main research output is the analysis of museum crowdsourced projects and 'games for social good' to develop potential models for game-like interactions suitable for museum collections, and the subsequent evaluation of the designed interfaces.

Aims and Objectives

This project aims to answer this question: can game-like interactions can be designed to motivate people to undertake tasks on museum websites that will improve the overall quality of the website for other visitors?

More specifically, which elements of game mechanics are effective when applied to interfaces for crowdsource museum collections enhancement?

Objectives

- Design game-like interaction models applicable to cultural heritage content and audiences through research, analysis and creativity workshops

- Build an application and interfaces to create and store user-created content
- Evaluate the effectiveness of game-like interaction models for eliciting useful content

Potential usefulness of the project

This research might eventually inform the cultural heritage sectors need for methods to increase visitor engagement with and discovery of museum sites, particularly museum collections content.

While the number of games for social good and crowdsourced museum sites are increasing, there is little published research on the interface design or user experience elements that lead to successful implementations. An analysis of the game-play mechanisms used to motivate visitors in these contexts will help organisations developing similar projects.

Theory

As this is a relatively new field, terms and definitions are still fluid, and may well change during the project. Games for social good may also be known as games for good, serious games, meaningful games or considered as a form of crowdsourcing. Closely related topics include captology and persuasive design.

Recent projects such as Armchair Revolutionary[15] and earlier projects such as Carnegie Mellon University's 'Games with a purpose'[16] and InterroBang?![17] are indicative of the trend for 'games for social good'. Crowdsourced projects such as the Guardian newspapers examination of MPs expense claims[18], the V&A Museum's image cropping[19], Brooklyn Museum's tagging game[20], the National Library of Australia's collaborative OCR corrections[21]; Chen's (2006) study of the application of Csikszentmihalyi's theory of 'flow' to game design; and Dr Jane McGonigal's ideas about multiplayer games as 'happiness engines'[22] suggest that 'playful interactions' and crowd participation could be applied to help create specific content improvements on museum sites. Game mechanisms may help make tasks that would not traditionally considered fun or relaxing into a compelling experience.

According to Adams and Rollings (2006) the essential elements of a game are rules, goals, play, and pretending. This project will involve adaptations, rather than direct applications of game mechanisms. For example, competitive elements may excluded to keep the build manageable, to reduce the variable factors in testing, and to allow the site to function with a small number of players over time rather than a critical mass of players with periods of intense usage. Other game mechanics such as 'levelling up' and unlocking missions may also be excluded.

---

[15] https://www.armrev.org

[16] http://www.gwap.com/

[17] http://www.playinterrobang.com/

[18] http://mps-expenses.guardian.co.uk/

[19] http://collections.vam.ac.uk/crowdsourcing/

[20] http://www.brooklynmuseum.org/opencollection/tag_game/start.php

[21] http://newspapers.nla.gov.au/ndp/del/home

[22] http://www.futureofmuseums.org/events/lecture/mcgonigal.cfm

Within the terms of this project, the output of a game-like interaction must produce an effect outside the interaction itself - that is, the result of a user's interactions with the site should produce beneficial effects for other site visitors who are not involved in the original interactions. To achieve this, it must generate content to enhance the site for subsequent visitors. Methods to achieve this could include creating trails of related objects, entering tags to describe objects, writing alternative labels or researching objects - these will be defined during the research phase and creativity workshops. Content created may include data not visible on the page, such as relationships between items.

References

Adams, E., and Rollings, R. (2006) *Fundamentals of Game Design*, New Jersey: Pearson Prentice Hall

Chen, Jenova (2006) *Flow in Games* Accessed online at http://www.jenovachen.com/flowingames/abstract.htm

Csikszentmihalyi, Mihaly (1990) *Flow: The Psychology of Optimal Experience,* New York: Harper and Row

Trant, J., Wyman, B. (2006) *Investigating social tagging and folksonomy in art museums with steve.museum*. Accessed online at http://www.archimuse.com/research/www2006-tagging-steve.pdf

Methods and tools

The project is divided into several stages, each with their own methodology and considerations.

Research

The preliminary research process involves a literature review, research into game mechanics and the theory of flow, and research into museum audiences online. It will also include a series of short semi-structured interviews with people involved in creating crowdsourced projects on museum sites or game-like interactions to encourage the completion of set tasks (e.g. games for social good) in order to learn from their reflections on the design process; and analysis of existing sites in both these areas against the theories of game design. This research will define the metrics of the evaluation phase.

Creativity workshop(s)

The results of this research phase sets the parameters for creativity workshops designed to come up with ideas and possible designs for the game-like interfaces to be built. Possible objectives for the creativity workshop include:
- designing methods for building different levels of challenge into the user experience in an environment that does not easily support different levels of challenge when museum-related skills remain at a constant level
- creating experiences that are intrinsically rewarding to enable 'flow'

Build and test

In turn, the creativity workshops will help determine the interfaces to be built and tested in the later part of the project. The build will be iterative, and is planned to involve as many build-test-review-build iterations as will fit in the allocated time in order to test as many variant interaction models as possible and support optimisation of existing designs after evaluation. User recruitment in this phase may be a sample of convenience from the target age group.

The interfaces will be developed in HTML, CSS, Javascript and published on a WordPress platform. This allows a neat separation of functionality and interface design. Session data (date, interface version, tester ID) can be recorded alongside user data. WordPress's template and plug-in based architecture also supports clear versioning between different iterations of the design, allowing reconstruction of earlier versions of the interfaces and split A/B trials if necessary.

Analysis and write-up

Analysis will include the results of user testing and of the user data recorded in the WordPress platform to evaluate the performance of various interface and interaction designs. It may also include log file or Google Analytics analysis of the interfaces, if the platform attracts usage outside the user testing sessions.

Workplan

June, July: research phase and workshop preparation

August: creativity workshop(s) and initial platform development

September, October, November: build and test (rapid prototype and test iterations)

December, January: analysis continues, writing up.

Project Feasibility

Resources

During the build, collections data will be drawn from public-facing and internal APIs published by the Science Museum; this may be supplemented with collections from other museums available as machine-readable data. This project also takes advantage of work already undertaken at the Science Museum on the WordPress platform.

Resources required:
- access to suitable collections data
- computers with internet access for user testing sessions
- a server capable of running the developer interfaces and storing user data
- access to testers

Knowledge and skills

This project will require the knowledge and skills gained throughout the taught courses of this MSc.

It will also require PHP and MySQL development skills, WordPress template and plugin development knowledge, HTML, CSS and Javascript implementation.

Risks

Risks include 'scope creep' in the requirements for the interface. This can be mitigated by an internal 'gate keeping' policy on build requirements, and through the use of Agile development techniques such as user story cards to record and prioritise interface elements and functionality.