Shall We Play a Game?: Crowdsourcing Metadata and Gamification in the Digital Archive

In a Google Tech Talk discussing gamification in digital media given in January of 2011, Dr. Sebastian Deterding - an Assistant Professor at Northeastern University in Boston and a scholar of “gameful design” - references the opening pages of Mark Twain's *The Adventures of Tom Sawyer*, in which the titular hero convinces a gullible group of friends to complete his chore of whitewashing a fence for him. Deterding emphasizes the devious (and effective) strategy underlying Sawyer's gambit: by repeatedly implying that he would rather be whitewashing than out fishing with the other boys, Sawyer builds the illusion that they would also have more fun doing his work. Perhaps it's a cynical view, but it is difficult to look at attempts by various institutions to crowdsource the metadata in their digital repositories and not see at least a little bit of Sawyer's sly logic behind the effort. The benefits in discoverability for digital objects through thorough, accurate metadata tagging are considerable, but when a single collection can contain thousands and thousands of objects, even the most well-funded institutions can be hesitant to devote the staff necessary to perform such a monotonous, unskilled and time-consuming task. By bringing in members of the public to, if you will, whitewash the fence, an archive can create (and pass along) a better described and more searchable collection, at a relatively low cost. The catch, of course, is in luring in the public in the first place: a user's participation in a crowdsourcing project has to be bought with some kind of benefit provided in return. The simplest option available to most archives is to make an argument that is more or less the same as Tom Sawyer's con: “it's fun!”

To be less skeptical about things, most archives do of course genuinely want to provide a meaningful experience for their users when building a crowdsourced metadata project. The ideal is not simply to make a strict business transaction and outsource labor, but, as Trevor Owens has written, to create a new way for the public to connect with the archive and its contents, “to interact with, explore, understand the past.” It is in the interest of the institution (and possibly even in its mission statement) to cultivate this sort of more fruitful relationship with amateur (unpaid) archivists. One suggestion that has arisen for encouraging and accomplishing that goal is the gamification of crowdsourced metadata: that is, the implementation of game design into public, online metadata tagging platforms. But those who run an archive or digital repository can not simply equate the idea of “gaming” with “fun” and assume that gamification will be an effective strategy for their collections. In order to decide whether to implement a gaming aspect to a crowdsourced metadata project, archivists need to both consider what makes for effective gamification, and analyze their capacity to implement those features. By looking at Deterding's three requisites for “meaningful play” - autonomy, mastery and meaning - in the light of existing crowdsourcing projects (from both within the archiving world and without), and distilling those needs into specific challenges of execution in an archival context, we can get a rough idea for successful scenarios for gamification of metadata tagging.

The first of Deterding's gaming requirements we will consider here is the notion of autonomy – the question of whether the user feels free to pursue their own actions in regard to the game. If work is that which one is obliged to do and play is anything that one is not obliged to do, good gaming should always convey the latter, never pressing the user so that they feel mandated towards a particular action. Returning briefly to the Tom Sawyer metaphor, the reason Sawyer's gambit worked was that for his friends, whitewashing the fence was a choice; undertaken as a free action, it became play, when from Sawyer's perspective it was required, disciplinary work. Metadata crowdsourcing projects have the advantage of being completely elective, of course; posing tagging as a volunteer action to the archive's community, rather than as a required task for the archive's employees, automatically shifts the action into the realm of play. What could be simply a matter of human computation (employing humans to solve a problem whose scale and nature a computer can not handle, in this case accurate subject identification and tagging for non-text-based images, audio and video) is turned from labor into contribution. Contrast this with the Tagasauris service (http://www.tagasauris.com/), a start-up founded in 2010 that offers human-assisted image tagging, organization and transcription: for 5¢ a photo, individuals or businesses can upload a collection of images to the Tagasauris platform, where a network of contributors around the world provide a series of subject and descriptive tags, curated by computer algorithms that ensure high searchability for the provided metadata. The key element here is Tagasauris’ “on-demand labor force;” though the details are not readily available, Tagasauris' own phrasing implies these humans in the human computation model actually require some kind of compensation, the costs of which are then transferred on to the user of the service – this is not treating crowdsourcing metadata as play, but work, even if Tagasauris' crowdsourced model severely dilutes the cost over, say, individual archive employees performing the same

---

2 Deterding.
Freedom to participate in the game or project is only part of what Deterding is referring to by autonomy, however; he is also alluding to the user's ability to explore within the game, without a perceivable sense of the game designer controlling and guiding the user's actions (even if they are). Achieving this in practice can require a certain amount of knowledge of game theory and design, which is not necessary to elaborate on here. More important for the moment is the fact that this aspect of "meaningful play" should synchronize entirely with the goals of many digital archives, particularly those attached to cultural and educational institutions. Exploration and discovery within the archive should be encouraged, as in projects like the New York Public Library's collaborative menu transcription project, "What's on the Menu?" which allows users to browse and choose the menu they wish to transcribe based on date or name of the restaurant. "What's on the Menu?" also offers the immediate ability to browse all the project's data that has been reviewed and approved (again by date, name of restaurant, or, in this case, by name or number of dishes). What this project recognizes is that the results of the game are still a part of the game – if the ultimate object of the game is to build a discoverable collection of data, the user should ideally have the option to go back and explore what they (or others in the game) have already accomplished. Even if such a feature is impractical, however, as in explicitly competitive, gamified projects like Tiltfactor Laboratory's One Up, Zen Tag or Pyramid Tag, users can still get a sense of freedom through customizable content settings: the user has the power to decide how random or targeted the content of the images they tag is, still creating the sense that the user is browsing through the repository's offerings even if the designer's algorithm is deciding what specific digital objects to place in front of the user for tagging.

The second of Deterding's requirements is the element of mastery, a sense of progress or achievement for the user provided by the game. Deterding points to "Progress Wars" (http://www.progresswars.com/), a parody "game" developed by Jakob Skjerning, as an example of the problem with many attempts at gamification: instead of offering meaningful challenges and any sense of actual development, many games merely provide an artificial progress bar and arbitrary rankings, confusing numbers going up with achievement. If "fun is just another word for learning," what is really important about a game is the sense of comprehension that arrives from solving some sort of puzzle or challenge (under optimal conditions – as Deterding points out, there are still reasons why, for instance, children don't consider solving math problems in school "fun," but many of these scenarios likely have to do with the sense of obligation described earlier). Effective gamification requires scaffolding, a design that allows for a generally upward flow in challenge and a corresponding increase in learning and comprehension. Such mastery increases the odds that a game will not only capture the user's attention, but hold it.

This is clearly an area in which current metadata tagging gamification projects struggle. For instance, the gaming aspect New York Times' "Madison" platform (http://madison.nytimes.com/), recently launched to encourage crowdsourced tagging and transcription of advertisements from the newspaper's digital archive, is essentially a version of Skjerning's "Progress Wars." For each ad successfully tagged, a "Madison" user is accorded a point; approximately every ten points, they are assigned a new, arbitrary rank ("Reader," "Fledgling Finder," etc.). The task itself never changes (unless the user switches from ad identification or tagging to transcription), and remains essentially the same for the entire time that the user chooses to play. This is a drawback of the unstructured and unpredictable nature of the content provided by the archive: there is no way of controlling which object will appear to the user according to the difficulty in accomplishing the task of description. The user may learn more by viewing more content, but strictly in terms of the game, there is no element of mastery. This is a large potential stumbling block for tagging projects, and likely part of the reason why some institutions still opt not to gamify their platforms, relying on more intrinsic rewards for user motivation.

However, Tiltfactor Lab's Metadata Games project (http://www.metadatagames.org/), sponsored by Dartmouth College, has offered a potential solution to the problem of scaling difficulty by introducing an element of competition into their mobile description app, One Up. In One Up, players score points by submitting single-word tags for an image, and attempt to gain more points than an opponent by applying more accurate tags ("accuracy" in this case being measured by whether a tag they submit matches one submitted by other players in previous games). One Up is also a multi-round game, and as the players move into the second and third rounds of each game, they are penalized for submitting tags that were already submitted by their opponent: in other words, they are encouraged to devise tags that are not only accurate (as determined by the crowd) but unique enough that not every player would necessarily think of them – for instance, once they reach a second or third round, given a picture of a dog, the player would likely be better off submitting the specific type of dog breed as a tag rather than just the word "dog." By creating these more strict, competitive rules for their game structure,
Tiltfactor has crafted challenge through strategy: players can improve their understanding of the game's point system and increase their likelihood of victory while simultaneously increasing the quality of their metadata (obviously a highly desirable result for the archive as well). Like with Madison or just about any other crowdsourced metadata project, the actual task at hand may never change, but the user's method for accomplishing the task can be adjusted and fine-tuned. The two-player setup enhances this experience, allowing that perhaps you don't have to be the best at imagining unique tags, but you just have to be better than your opponent. The user's sense of progress is set in relation to that of other users. Tiltfactor refers to how this Outlier Design model for One Up “incentivizes” accurate but non-obvious tags, but they are essentially referring to the same concept as Deterding's mastery, and their playtest results suggest both increased specificity of metadata and user satisfaction.7

The last and likely the most crucial element of “meaningful play” for Deterding is, of course, meaning. This factor is perhaps the most ephemeral and difficult about which to draw broad conclusions, as it refers to a game's ability to connect to a user's interests – which, by nature, can be highly subjective and ever-shifting. But, in general, creating meaning in a game requires a certain amount of empathy, a capacity to understand the personal goals that a user brings with them when they start to play a game. Even if a game can successfully build a sense of mastery, increasing the challenge it takes to level up or score a certain number of points, the designer must understand that these artificially crafted goals are not an endgame, but a method of expressing or visualizing the underlying motivations of the player.

Familiarity with the personal goals of players requires a corresponding familiarity with the desired user base for the game – it is impossible to determine motivation if the targeted community for the game remains nebulous. Expecting gamification in and of itself to draw in new users fundamentally misunderstands the subtleties of different game types and interfaces (competitive vs. non-competitive, collaborative vs. individual, narrative, puzzle-based, interactive, etc.) and their unique attractions to different players; it also violates the spirit of the Designated Community as defined by the OAIS reference model, followed by many digital repositories: the “subset of Consumers expected to independently understand the archived information in the form in which it is preserved and made available by the [digital archive].”8 The intended users of a tagging game, in other words, should be seen as the same as the intended audience for the contents of the archive; gamification is a method of connecting to a Designated Community, not creating it. Having a clear idea of how a Designated Community wants to access a repository and why is rather integral to a digital archive's general operations, but by extension it can also assist with the construction of a game whose particular design appeals to the Community's inclinations and demands, encouraging participation.

In some cases, crafting meaning around a game may be as simple as the provision of access to an otherwise unavailable and undocumented collection. Related to the user's sense of freedom (see the above section on autonomy) is the personal goal of discovery and exploration: there can be an inherent attraction in offering up materials in a game that would otherwise be very difficult or even impossible for the user of a digital repository to find. The new-ness and randomness of images or other digital objects presented to the user without any particular attempt at curation by the game designers or the archive could appeal to members of the target community who are simply interested in browsing a collection and seeing something that they haven't seen before. The action of then tagging that object with metadata becomes something like astronauts planting a flag on the moon: a way for that user to leave their mark, a conflation of both self-importance and the sense of participating in a larger movement of human discovery and advancement. The sense of personal discovery (the user seeing an object for the first time) is compounded by the knowledge that the player is extending the archive's knowledge. Again drawing from Trevor Owens, it is worth considering that “there is a kind of discovery that happens when one closely reads documents that we already knew were there but no one spent the time to extensively analyze.”9 Assigning metadata can be a mundane task, but it is also an analytical one, and therefore likely to be rewarding by itself for at least some players.

An area where gamified platforms could easily learn about the application of meaning from non-game tagging projects is in the ability to foster a community among the game's user base – not just building connections between the repository and the users, but between the users themselves. Several projects have created active and successful user bases by offering a space for those users to interact with each other, regarding the discoveries they are making in the repository or on similar topics of interest. For instance, the New York Public Libraries' “What's on the Menu?” (http://menus.nypl.org/) project, launched in 2011 and which has now successfully transcribed over 17,000 scanned menus from New York area restaurants dating back to the 1850s, maintained a blog attached to the platform. The blog ran posts not only from librarians and archivists directly involved in collection's preservation, but some chefs and restauranteurs with an interest in the city's culinary history. The project also maintains a fairly active Twitter presence (again, allowing a space for conversation

6 Ibid., p. 8.
7 Ibid., p. 11.
8 Lavoie, p. 6.
9 “Archives as Discovery Zones.” Blog post.
between interested food professionals and amateurs) and regularly updates a publicly available CSV of their data from the project, allowing simultaneous scholarly investigation and debate over the data set at sites like Katie Rawson and Trevor Muñoz’s “Curating Menus” (http://www.curatingmenus.org/). Similarly, Galaxy Zoo (http://www.galaxyzoo.org/) is a scientific tagging site run by the non-profit Citizen Scientist Alliance consortium which offers thousands of images taken from the Hubble and other high-powered space telescopes from around the world to the public for assistance in identifying galaxies, stars and other space objects; critically, Galaxy Zoo, offers a public forum where users of the platform can post images they have found that are interesting or confusing, sparking a large amount of debate and discussion among a base of contributors that has included over 150,000 people. The site has been active since 2007 and received millions of classifications, and spawned a large series of spin-off projects (the “Zooniverse”) based on scientific image and audiovisual data sets.

Contrast these robust online public spaces to Madison, which emphasizes the aid contributors will provide to researchers, scholars and businesses by adding data to the New York Times archive, but offers no explicit space to interact with these beneficiaries. The Madison platform does offer links for users to share individual images on social media, but there is no particular guarantee that sharing images in that way will get them in front of other interested users. What Galaxy Zoo and “What's on the Menu?” understand are some of the personal goals that lie behind online community and sharing: an opportunity for socialization, for education through discussion, for enhanced reputation, recognition and respect from peers. These are powerful motivators for participation in a crowdsourcing project, or indeed most any sort of online venture; the precedent for online community in gaming, for instance, is abundantly evident in the popularity of Massively Multiplayer Online games like World of Warcraft and the endless forums, wikis and apps that go with them. Spaces for public interaction could offer an opportunity for a repository's target community to collaboratively make meaning out of the contents of a tagging game.\(^\text{10}\)

One last method for constructing meaning in a game is to consider the possibility of narrative in and around the game. The nature of the task at hand will of course make it difficult to implement narrative into a metadata tagging game the same way users might be used to getting a story from, say, a video game. The entire point of a crowdsourced metadata project is that the content in question is under-described, and it is difficult to think of a way in which a game designer would be able to build a story out of content with no context. But there is ample opportunity in an archival context for meta-narrative, in which the description, community and publicity that surrounds a crowdsourcing platform casts the user in the role of protagonist in a story of social progress and discovery. For example, ProPublica, an online investigative journalism non-profit, in 2012 launched “Free the Files” (https://projects.propublica.org/free-the-files/), a crowdsourcing project in which users are tasked with sifting through reams of publicly released documents from the FCC to log political ad campaign expenditures in a number of critical congressional districts from around the country. Assuming that ethically dubious or outright illegal behavior on the part of interest groups and super PACs is available but buried and obfuscated within these dense documents, ProPublica casts the users of “Free the Files” as, essentially pro bono staff members – noble citizen journalists tracking down important stories of fraud or deceit. Likewise, Galaxy Zoo often refers to its contributors as “citizen scientists” and even titles its “About” page “The Story So Far” - again, quite consciously placing its users in a well-defined and noble role, which allows for the glorification of the individual in the midst of a project that relies on the contributions of thousands. The idea of the amateur “citizen archivist” helping to advance world knowledge or uncover unknown secrets could offer similar results for many digital repositories; in fact, the National Archives are already employing this meta-narrative with their Citizen Archivist Dashboard (http://www.archives.gov/citizen-archivist/). Though it's obviously not an element specific to gaming, archives could consider how this sort of larger construction of story around the user can supplement and enhance the sense of meaning within their desired game.

Having looked at the broadest, conceptual necessities for meaningful play and the ways they could affect strategy for building a crowdsourced metadata tagging game, it is worth considering some of the more specific challenges faced in implementing such a tool, and the situations in an archival environment where gaming might be applicable. First and foremost are issues that overlap with the needs of any digital repository: you can't build a metadata tagging game without content to tag. A strong foundation should already exist in terms of storage, the availability of and ease of retrieving access copies should be such that a crowdsourcing/gaming platform can be built on top of the storage system's front end, and the content needs to be copyright-cleared at least to the degree that the archive feels comfortable providing public access. The institution's server needs to be able to handle the demand for access to the repository.\(^\text{11}\) These are hopefully concerns that...

---

\(^\text{10}\) Metadata Games does offer a searchable database of all the materials tagged by users using their various games; however, it does not offer any sort of comment or public interaction section.

\(^\text{11}\) Author's note: though I had no difficult accessing Metadata Games' browser-based games like Zen Tag and NexTag, their One Up app frequently encounters an “error in the server node” that renders the app inoperable. My earlier description of the app was based on Tiltfactor's documentation of the game's theoretical function, but the institution obviously still has some issues with practical use/application.
will already be solved before the consideration of gamification even arises, but the the archive should not ignore how front-end gamification could potentially affect back-end operations, or vice versa.

The actual content of the collection concerned will also affect gamification. As a practical concern, the format of the materials can create difficulties in both game design and proper tagging; there is most likely a reason that until recently, most of the major crowdsourcing metadata projects discussed here are image and text-based files. For any number of reasons (difficulties in organizing and describing a wider variety of acceptable file formats, storage space, institutional bias, etc.) audiovisual formats have proven slower to make their way into digital repositories, and present unique challenges in providing access in platforms like a tagging game. However, strong strides have been made in leveling the playing field in this regard, as “Waisda?”, a tagging game developed by the Netherlands Institute for Sound and Vision in 2009, and Metadata Games' NexTag are capable of tagging audio and video files.

More broadly speaking, the intellectual content of the collection can prove a challenge if, unlike many of the projects described here, there is not a readily available hook or obvious reason for a broad base of users to be interested in what they might discover. Restaurant menus and images of space taken from the Hubble have some obvious attraction, thanks to the relatively sharp focus of the collection and a built-in community around their subjects; the New York Times, meanwhile, realized they could build publicity based on the popularity of the television show “Mad Men” by initially using Madison to focus only on their advertisements from the 1960s, before eventually opening up other eras in the future. Other collections may have less broad or at least less obvious appeal, severely narrowing the number of users who might be interested in playing a game based on that content. This could be as much of an issue for holding the user's attention as it is for attracting them in the first place – even in a collection where the player thinks they might discover something, they might come to the conclusion that they there is in fact nothing valuable (for them) to discover at all.

Implementing a metadata tagging game also requires a consideration of folksonomy. Folksonomy refers to a system of classification derived from collaborative tagging, and in this context regards, essentially, the kind of metadata that the archive hopes to receive in return from its user base using the tagging game. A broad folksonomy is accomplished by including a wide variety of users employing a correspondingly wide variety of vocabularies with their tags (thus generating a higher quantity of metadata), as opposed to a narrow folksonomy, in which a more limited number of users generates tags using a more select and specific vocabulary. The general intent behind a crowdsourcing metadata project would seem to favor a broader folksonomy, but it is possible to design a tagging platform and/or game that guides a user towards a narrower vocabulary in order to limit what could conceivably become an overwhelming and inconsistent amount of metadata input. Madison's subject tagging interface, for instance, only allows for selection from a number of pre-selected categories rather than a free text field. The level of description an archive hopes to receive from a tagging game folksonomy will likely depend on the repository's already-established system for capturing, logging and storing metadata; the back-end design of the game itself must be tailored to ensure the data logged by users can be verified, organized and filtered easily into the repository.

Finally, it is good to consider Deterding's requirements for meaningful play at a conceptual level, but understanding how they are practically applied in-game requires knowledge of game design, an area that archivists are not necessarily familiar with on a professional level. Creating a game that is functional, accessible and intuitive to the desired users, and “fun” as defined here requires expertise, resources, and time. The game has to be designed, built, extensively tested (with test users taken from or equivalent to the target community) and integrated into the front end of the repository, all steps at which the archive will likely find unexpected results that require tweaking and reconsideration of the game's goals and execution. Even institutions with large, under-described digital collections already served safely in a repository may be wary of this undertaking, whether for financial, staffing or other concerns. It is worth, however, noting some potential avenues for cutting these costs, including the potential availability of open source software. All of the software developed by the Metadata Games team is open source, available for no licensing fee, allowing for other institutions to simply customize Tiltfactor's already-developed games to their own collections. With the baseline software and game development has at least already been completed, other archives can keep their attention on other challenges of implementation. Though it has not yet been released, members of the New York Times R&D Lab have indicated in the press that they plan to make Hive, the modular platform underlying Madison, available as open source software that other institutions and companies could also use for their crowdsourcing needs; and the Netherlands Institute for Sound and Vision have made the software behind “Waisda?” openly available on GitHub. Meanwhile, archives that have some amount of digital objects that they would like to be tagged, but do not think they have enough material to merit the development of a game or crowdsourcing platform of its own, can currently contact the Metadata Games group about having their materials simply submitted to be included in One Up, Zen Tag or any of the project's already-established games. Building a tagging game does not require acting in a vacuum and starting over from scratch – significant headway has been made with the

concept, and developers can build on those foundations to improve metadata gamification.

A factor of gaming that has generally gone undiscussed relative to metadata tagging and current tagging games is the question of intrinsic versus extrinsic motivation. Relying on intrinsic motivators (e.g. a sense of civilian duty, a desire to help, interest in material, a sense of community and most of the other considerations outlined previously in the section on “meaning”) may be something of a necessity when it comes to non-profit or academic institutions; and in any case, the argument of work as opposed to play and the improved sense of meaningful play without obligation previously discussed may stand as a considerable argument against the use of any kind of extrinsic compensation for participating in metadata tagging games. Considering the potential investment already made by an institution into developing a gaming platform, combined with the precedent of successful volunteer crowdsourcing projects like Galaxy Zoo, even a corporate-based archive (like the New York Times Digital Archive) might be hesitant to offer tangible rewards for completing a certain (likely extremely high) number of tags. But beyond Tagasaurus, which in any case is explicitly billed as a service and functions along a business rather than an archival model, it would seem that the sample size of major online crowdsourcing metadata projects that have offered extrinsic motivation is essentially zero. How might a free tote bag at 10,000 tags or a personal tour of an archive for the highest-scoring player of the year affect implementation of and participation in a metadata game? For now, that question remains unanswered.

Keeping Deterding's requirements and some of the above challenges in mind, it is possible to summarize some of the key issues posed to an institution considering gamification as a solution for gathering metadata for their digital collections:

- How focused is the nature of your collection? Does it lend itself to a well-defined, unique platform, or is it more broad and scattershot?
- Consider the line between crowdsourcing and gaming: is the target community for your repository interested in a game?
- Can you offer all three of Deterding's elements of “meaningful play?” Do you have a game design that will let users a) feel free to participate in and explore the contents of the platform, b) achieve some measure of personal accomplishment or learning, and c) connect to the specific contents that you are offering?
- Can you upload to an existing platform, or do you need to design and build your own?

It is impossible to make any broad declarations regarding the worth of gamification in metadata tagging. The variables involved in institutional preparedness, ability and targeted communities make it an option that may likely be unnecessary for many or even most archives, but it remains an intriguing avenue for further exploration and testing. The benefits of a whitewashed fence merit proper consideration and investigation.

Bibliography/Webography


