

INSTITUT FÜR INFORMATIK
der Ludwig-Maximilians-Universität München

KUNSTMAKLER

A GWAP with a motivating
narrative for collecting deep
semantics tags

Kilian Bosse

Bachelorarbeit

| | |
|-----------------|--|
| Aufgabensteller | Prof. Dr. François Bry |
| Betreuer | Prof. Dr. François Bry, Martin Bogner |
| Abgabe am | 16.11.2020 |

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst habe und keine anderen als die angegebenen Hilfsmittel verwendet habe.

München, den 16.11.2020

Kilian Bosse

Abstract

This thesis presents *Kunstmakler*, a *Game With a Purpose* (GWAP) for the Artigo platform. Artigo a project dedicated to annotating artworks with browser-games in order to fuel a keyword based search engine. The specific goal of the game is to collect annotations that are not often gathered by other games, which tend to collect mostly surface-level descriptions. This is integrated with a more gameplay-focused approach to design and the novel inclusion of a narrative context for the experience, an idea not present in other GWAPs. For this, the current state of Games With a Purpose, gamification and game design are analysed, especially on the Artigo platform, to firmly establish the niche *Kunstmakler* fills. The design decisions and the constraints they are based on are thoroughly discussed, such as the weighing of gameplay versus data-gathering, avoiding bias in the data gathered and creating an enjoyable experience for the player. The implementation is briefly described as well as an assessment of potential shortcomings. At the end, further possibilities to expand the - currently rather minimalistic - game are presented.

Zusammenfassung

Diese Arbeit stellt *Kunstmakler* vor, ein *Game With A Purpose* (GWAP) auf der Artigo Plattform. Artigo ist ein Projekt um Kunstwerke mithilfe von Browserspielen mit beschreibenden Schlagwörtern zu versehen, um eine Suchmaschine für diese Kunst zu ermöglichen. Das Ziel des vorgestellten Spieles ist es Schlagwörter zu sammeln, welche von den anderen Spielen der Plattform seltener erhoben werden. Dabei handelt es sich vorwiegend um nicht-oberflächliche Beschreibungen. Dieses Ziel wird kombiniert mit einem erhöhten Fokus auf das Spielerlebnis und einem Narrativ als Kontext des Spiels, eine Idee die im GWAP Feld nicht verbreitet ist. Zum Anfang werden GWAPs im Allgemeinen, Gamification und Game Design untersucht, vor allem im Bezug auf das Artigo Projekt, um die Nische, die das Spiel auf der Plattform füllen soll konkret zu bestimmen. Die Designentscheidungen sowie die zugrundeliegenden Umstände werden ausführlich diskutiert, zum Beispiel die Abwägung zwischen Datenerhebung und Spielerlebnis, Ansätze gegen die Verfälschung der erhobenen Daten und ein unterhaltsames Erlebnis für den Spieler zu schaffen. Die Implementierung wird kurz beschrieben, gefolgt von einer Einschätzung zu vermuteten oder möglichen Problemen. Am Ende werden Möglichkeiten beschrieben die derzeit eher minimalistische Umsetzung zu erweitern oder darauf aufzubauen.

Acknowledgements

First, I would like to thank Prof. Dr. François Bry for his enthusiasm for the Artigo platform and the great advice. His knowledge the art world and GWAPs, which he readily shared, greatly helped me tying different ideas together.

Secondly, I have to thank my research supervisor, Martin Bogner, without whose endless assistance, help and feedback this thesis would probably never have come to be.

Lastly I would like to thank my family and especially Julia for motivating me every step of the way, and proof-reading even the roughest drafts.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction and Related Work | 3 |
| 1.1 | Overview | 3 |
| 1.2 | Introduction | 3 |
| 1.3 | Games With A Purpose | 4 |
| 1.4 | Deep Semantics Tags | 4 |
| 1.5 | Casual Games | 6 |
| 1.6 | Gamification | 7 |
| 2 | Concept | 9 |
| 2.1 | Scenario and Concept | 9 |
| 2.2 | Exemplary Play-through | 9 |
| 3 | Design Considerations and Decisions | 11 |
| 3.1 | Weighing Gameplay Against the Data Collected | 11 |
| 3.2 | Premise and Narrative of the Game | 11 |
| 3.3 | Selection of the Artworks for the Game | 12 |
| 3.4 | Tag Selection | 13 |
| 3.5 | User Interface | 13 |
| 4 | Implementation | 15 |
| 4.1 | Frontend and Backend-Communication | 15 |
| 4.2 | Back-End | 16 |
| 4.3 | Tag and Artwork Selection | 16 |
| 5 | Challenges and Problems | 19 |
| 5.1 | Are the Taggings Generated Actually Deep | 19 |
| 5.2 | Shallow Image Selection Limiting the Selection of Playable Tags | 20 |
| 5.3 | Data Integrity | 20 |
| 6 | User Feedback | 23 |
| 7 | Perspectives | 25 |
| 7.1 | Expanding the Gameplay | 25 |
| 7.2 | Expanding the Game Behind the Scenes | 26 |
| 7.3 | Evaluating the Game's Success | 27 |
| | Bibliography | 29 |

CHAPTER 1

Introduction and Related Work

1.1 Overview

This thesis is structured as follows: First the history and context behind “Games with a Purpose”, the Artigo project, their successes and shortcomings will be discussed along with problems common in the field, as well as an overview of game design, cumulating in the motivations for this thesis. After that the concept for the “Kunstmakler” Game will be presented, followed by the design choices and the considerations that led to these decisions. The implementation section consists of a short overview of the project structure and highlights of noteworthy solutions. The feedback section describes and evaluates user feedback gathered regarding the implemented project. Perspectives is a collection of possible extensions for this thesis project, ideas cut due to scope constraints and potential follow-up analysis.

1.2 Introduction

‘Human Computation’ is not a new idea. In fact *computer* was a job title first, and then the thing was named after the activity. This however was largely forgotten in the public consciousness as the job was fully taken over by the aptly named machine. The reason for the swift and thorough takeover was in large parts due to the efficiency with which electrical computers could perform the (at first small) set of tasks they were designed for: unambiguous, simple calculations. Outside of this specific field of problems however, the human mind prevailed, especially for more complex or nuanced tasks. Especially describing images is a field where computers are only slowly catching up [21], and this rise in competence was originally fuelled by humans showing how to actually do it: manually labelling images from which neural networks could then learn. However, getting people to do such computation is time intensive and therefore too costly to commission on a large scale. So money as an incentive was not a desirable approach.

1.3 Games With A Purpose

The answer to this was presenting and implementing the task so that not money, but fun was the motivation for the *computers*: making the work enjoyable as a video-game. In 2003 Luis von Ahn presented the ESP Game: a way to generate annotations for images by letting people play a small browser-game. With this concept, the ESP Game spawned the genre of “Games With a Purpose” (GWAP). In the original implementation, two players were shown an image and asked to enter words which describe the image (referred to as tags in this thesis). Points were only rewarded when a word was entered by both players, which also served as the validation method for the gathered data.

The idea was soon licensed and implemented as the Google Image Labeler, which ran until 2011 with only small deviations from the core concept.¹ As a private enterprise, few results of this project were officially published, but the use of this data for training the algorithm behind Google’s Image Search [17] and the improvements and results of that service imply great success in gathering annotated images and utilizing them. Closely inspired by the original ESP Game, the Artigo Project was started to adapt the idea of labelling general images to tagging artworks instead, with the goal of providing a keyword-based search engine for art. To generate the keyword-data the search was based on, a number of different GWAPs were developed with different types or directions of tags in mind. Francois Bry[21] classified three major categories of games:

- Description Games, like the ESP Game, for surface-level tags, like visible objects;
- Diversification Games, like Karido,² highlight the differences between similarly tagged Artworks;
- Integration Games serve the purpose of gathering connections and relations between tags to allow for a better evaluation of the data gathered by other games.

Description and Diversification games on the Artigo platform function identically regarding the gathered data: Taggings, as a relation between an artwork and a tag are saved, where a tag is a keyword that should describe the artwork in some way. As opposed to the original ESP Game, validation of the tags is not handled in a game session (where the two players in one round had to agree), but outside of any individual game. A tag for a certain artwork is considered as validated, when it has been entered by multiple different players often enough, which is counted across different games.[2] Despite the difference, this method of data validation still falls into the same category Luis von Ahn originally put the ESP Game in: “output agreement game”, since the net output of both (in the original game), or multiple (on Artigo) players has to match. [19, p. 61]. Other types envisioned by von Ahn was the input agreement game, where players would put out an “input” and then had to change or adapt it until they agreed that their inputs matched.

1.4 Deep Semantics Tags

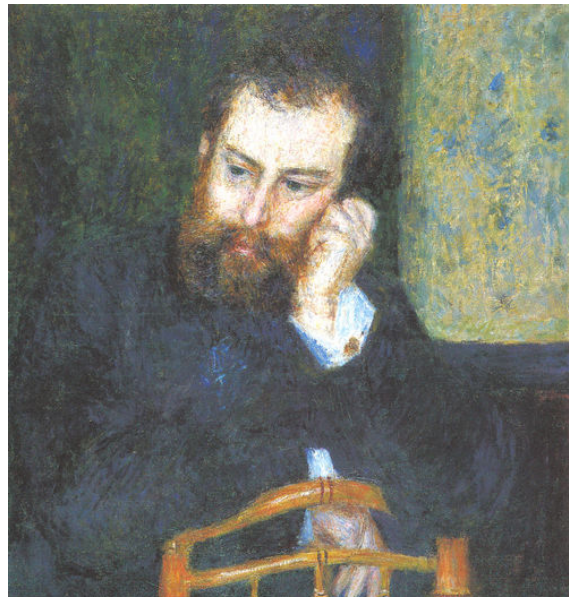
With the main game of the ecosystem being a rather straightforward adaptation of the original ESP Game, some of its shortcomings have carried over as well. One of these is the tendency for the gathered tags to be rather superficial[18]; since the design of the game requires both players to agree the most obvious keywords are the most incentivised results. [17] [4] In the end, this pushes players to resort to mostly object recognition and extracting other basic information like colour scheme and text contained in the image, generally: the

¹at crowdsource.google.com Google has since relaunched its crowdsourcing effort, but the activities differ far from the original ESP Game and are hardly gamified besides an achievement tracker

²<http://www.artigo.org/karido>

most surface-level characteristics of the artwork. The Karido game has been added to the Artigo ecosystem to counteract this issue by focussing on the differences between similarly tagged artworks (thereby implicitly blacklisting the most common words). While this does remove the incentive to enter the most obvious tags, there is still a focus on using more surface level factors distinguishing the images in question, and still primarily incentivises “surface semantic tags” [4], like small objects contained within one image but not the other. And while these surface semantics tags contain valid information, they do not represent the full range of keywords that might be searched, or that the Artigo dataset could be usefully analysed for. While some keywords can be considered deep or surface semantics depending on the context, more abstract concepts tend to fall more clearly into the category of deep semantics or rich tags.

For Figure 1.1 as an example, the major tags according to the Artigo database are “man”, “beard”, “dark”, “blue”, “brown” and “armrest”, all surface level information. The obvious melancholy is far off in the ranking, and the sadness that mister Sisley’s gaze shows has not even received the amount of taggings necessary to be verified. While the surface level tags are perfectly valid, they are also only helpful when one is searching for such surface level characteristics. A problem arises when the deeper semantic tags are barely ever entered and therefore cannot be searched for or otherwise analysed. An exemplary search for images portraying either calmness, relaxation (or the related adjectives)



return a total of 14 images, while “sleep” and “sleeping” return over 50, several of which could perfectly be tagged with the aforementioned richer tags. The issue intensifies as the concepts desired become more abstract: “democracy” returns a total of 2 images, one of which contains the word as writing and the other in direct reference to a painting of a German parliamentary session (Figure 1.2) (Sisley not shown).³

Systemically generating these “deeper semantics tags” is far more difficult than having players gather surface level details, partially because of their more subjective nature. A man with a beard is almost universally recognizable, while it is up to each person whether they individually see or feel democracy portrayed in a painting of a parliamentary session, a sentiment that is often also dependent on the image’s original context. Multiple smaller approaches have been taken to increase the likelihood of deeper semantic tags being entered generally across Artigo’s games, for example priming players by including captions for the images on the front page to encourage pondering the images presented [4, p.5]. Systemically gathering deeper semantic tags was also attempted with the game “Sentiment”, which asked players directly what emotion an artwork conveys. While this definitely does generate deep semantic tags, it is also limited to purely the emotions conveyed, which is one single cluster of deep semantic tags. This thesis presents an approach to systematically generate deeper semantics tags by inverting the concept of the ESP game: instead of being presented with an image and proposing tags for it, the player will be presented with a keyword, and is tasked to find fitting images for this tag. This approach of (some-

³the democracy example was performed in German with the tag “Demokratie”, as that language’s dataset is far more extensive than the English one [2]

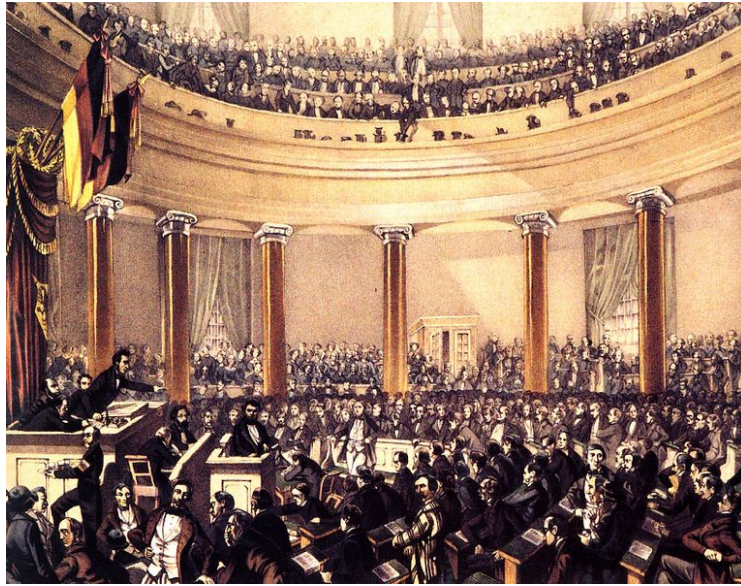


Figure 1.2: Ludwig von Elliot - Innere Ansicht der Paulskirche in Frankfurt, 1848. From artigo.org

what) directly asking the player what to do is called scripting and was proposed by Bry and Wieser as a good way to get the desired deeper semantics tags. Formally: “In pedagogical psychology, scripting denotes techniques to bring learners to perform as desired” [4, p. 6] While originally proposed for the game *Sentiment*, the method behind it directly transfers to this game as well. By replacing free form tag generation, or letting the players mind wander to whatever words they deem fit for the image, players are instead directly asked to contribute a specific set of taggings. In *Sentiment*’s case this was a plain question of what emotions an artwork conveys, but the blunt nature of the question comes across rather crude and changes the perceived activity from a game something akin to filling out questionnaires – an activity generally not considered fun. Therefore Wieser proposed performing this scripting via the context of the activity, which was then executed in several smaller low-key suggestions scattered across the Artigo platform (mainly the home-page) discussed in the previous paragraph. This thesis however executes the idea of scripting via the games context as one of the main design fundamentals in order to not merely increase the amount of deep taggings, but to solely gather them.

1.5 Casual Games

The terms “Casual game” or “hyper casual game” suffer from a clear definition or meaning, which does not stop them from being thrown around in the video game industry and especially self proclaimed “hyper casual game” companies are growing at a rapid pace [8] For “casual games”, Annakaisa Kultima et al. Attempted to come up with a clear definition and at least managed to compile several attributes indicative of the genre: simple gameplay and UI necessitating little to no tutorial or explanations, a forgiving but endless or replayable game loop as well as letting the player almost instantly play the game. While this characterization is now 11 years old, the attributes do not appear to have changed much compared to the current common understanding of the term, and are also nearly identical to that of “hyper casual games” nowadays [22, 23, 15], which is sometimes proclaimed to

be nothing more than an industry buzzword.[8]. These characteristics also apply to most GWAPs to a strong degree, to the point where non-casual Games with a Purpose seem to be the exception⁴, especially the Artigo platform's games all show these characteristics to a rather high degree. It is not a coincidence that they clearly follow this genre's design principles, even though the term gained popularity after most of Artigo's games were designed and therefore does not show up in any of the earlier papers as a design goal. The reason is likely one of the biggest benefits of casual games: the mass market appeal that results from these characteristics: they are easy to get into and do not require the time commitment of core and hardcore games. As of 2020, casual games are the most popular genre of video games overall, and have an audience that extends beyond the typical gamers [1]. A lot of GWAPs focus on problems that are either small by themselves or can be readily divided into subtasks fitting for the required short game loop as well as often relying on intuitive judgements instead of deep, focused thinking. The result is that gamifying these problems in such a way that anyone can pick up the game and start playing is a very sensible approach since the more players the game can attract, the more computation can be done for the original problem.

1.6 Gamification

Gamification is the underlying concept of all Games with a Purpose: Let the problem solving feel so much like a game that people will want to do it without additional extrinsic motivation. Or as Werbach & Hunter formally defined it in their 2012 book "For the Win: How Game Thinking Can Revolutionize Your Business": Gamification is "the use of game elements and game-design techniques in non-game contexts" [20, p.26]⁵ Wherein Game elements are the objects and things that make up the game, for chess it would be the pieces, the board but also the rules and possible actions of the pieces. These elements can make up a game, but can also be transferred to other areas, like problem solving for GWAPs or a scoreboard and point rewards for warehouse workers, which then is gamification. Game-design techniques are methods and approaches of how to effectively combine game elements into an engaging experience. Scoreboards and achievements could be slapped onto every interaction, but actually creating a cohesive experience requires game design as well [20, p. 26-30]. Most Games with a Purpose only use gamification somewhat lightly and only thinly veil the problem that is to be solved, relying partially on other motivating factors rather than the pure gameplay experience. They are recognizable as problems with game elements added, which makes sense since they are usually designed by computer scientists trying to make games, and not game developers attempting to produce scientific work. Another reason for this light gamification is that more complex gameplay can introduce biases into the generated data or distract from solving the actual problem. And while understandable, it is hindering their popularity, since the fun playing them is a major factor in motivating users to play Games With A Purpose (and thereby solve the original problem). While casual games generally lack the depth of 'regular' games and narratives are on first glance antithetical to and too deep for casual gaming, light narrative elements are found in a not insignificant number of successful casual games,⁶ giving context and an implied greater meaning to the simple gameplay as well as aiding players' immersion.[16, p. 40] Despite this aspect of game design fitting well with casual game design principles,

⁴One notable exception is the FoldIt project, which deviates from almost all casual game design principles and has still managed to acquire a core player-base that is actively playing the game. Homepage: <https://fold.it/>

⁵While there are many, usually slightly differing, definitions for the concept of gamification, this is the one used for this thesis

⁶these top-lists were used as a somewhat random slice of the casual gaming scene[25, 11, 24] as well as some examples being the games "Candy Crush Saga" and spin-offs thereof like "Farm Heroes Saga"

the topic is completely absent in Games With a Purpose, partially due to the aforementioned heavy focus on the problem to be solved. This thesis will attempt to integrate core GWAP problem solving with a light narrative framework by means of giving text-based context to the gameplay and isomorphing the actual problem solving to look and feel like story-appropriate activities. Grace & Jamieson [6, p. 658] recommend this abstracting and reflavoring of actual problem solving as an approach to designing a GWAP.

With this introduction of a narrative framework for the game mechanics and overall increased focus on gamification over the data gathering, the Kunstmakler game will hopefully deliver a more engaging as well as different experience; With the goal of attracting its own audience and not purely competing with the other games of the Artigo platform for a player-base. As mentioned above, narratives and casual game design sometimes clash, and while present in the industry, most casual games do not feature many narrative elements, so this GWAP might only attract a niche subgroup of casual gamers. In the categorization of A. Lindley[12], this game caters to the immersionist players almost exclusively, while the other games on the Artigo platform appeal far more to the achievers or socializers by how they are played and designed. While these categorizations are not mutually exclusive in a player, this difference in the target player type supports the secondary goal of not directly competing with other Artigo games for its player-base. Whether there is a substantial amount of casual gamers intrigued by a narrative has not been scientifically analysed but due to the single-player nature of this game, no large audience is required for a smooth gameplay experience; and even if just a comparatively small group of players is attracted by this approach for a GWAP, as long as they are not cannibalized from the other Artigo games, it is still a success for the platform. A 2014 analysis also counted games contributing only few annotations (tags) as a success, if those are not tags entered in other games [3, p. 6], which is highly likely, since Kunstmakler was designed to contribute rarely entered deeper semantics tags to the platform.

2.1 Scenario and Concept

This chapter will lay out the implemented end result of this thesis in order to let the design decisions and other considerations directly relate to the end result. Condensed as much as possible, the game can be described as follows: The player is in the role of an art broker (in German: “Kunstmakler”, hence the title), who knows the theme of an upcoming exhibition and is visiting auctions in order to buy suitable artworks. In the game, the auction is represented by a random sequence of artworks, the theme is a singular word and the bidding is modelled in a single-player, two step process against random numbers in the second phase.

Conceptually this inverts the flow of the original Artigo / ESP Game, because instead of finding fitting tags for an image, the player is tasked with finding suitable images for a given tag. This way we don’t gather the tags that first come to mind, but instead any tag desired. Any features beyond the core concept of selecting fitting (such as the budget, the second round of bidding etc.) images are primarily for gamification purposes in order to increase player enjoyment and therefore retention.

2.2 Exemplary Play-through

The game opens with a short text, introducing the player to the premise of the game and presents the theme for the upcoming exhibition (the tag to be played). After that individual images are presented one after another (the number of images is capped at 60 although the player is not expected to reach this limit), which the player can either save for the later auction, note down as potentially interesting or decline them, removing them from the showcase in the actual auction. Once multiple images have been selected, the player can move on to the second stage, the actual bidding. In this second screen 2.1, the player is presented with all his selected images at the same time (plus artworks marked as potentially interesting) and an estimated final bid for each image. The player can then enter his own bid for each image, while keeping within his total budget (shown as a live tracker in a sidebar). Once the player is content with the entered bids, they can submit the values, progressing the game to the final screen. Here the player is presented the images that he

managed to buy as well as artworks that can now be found in the Artigo Database under the keyword which is the theme of this rounds exhibition.

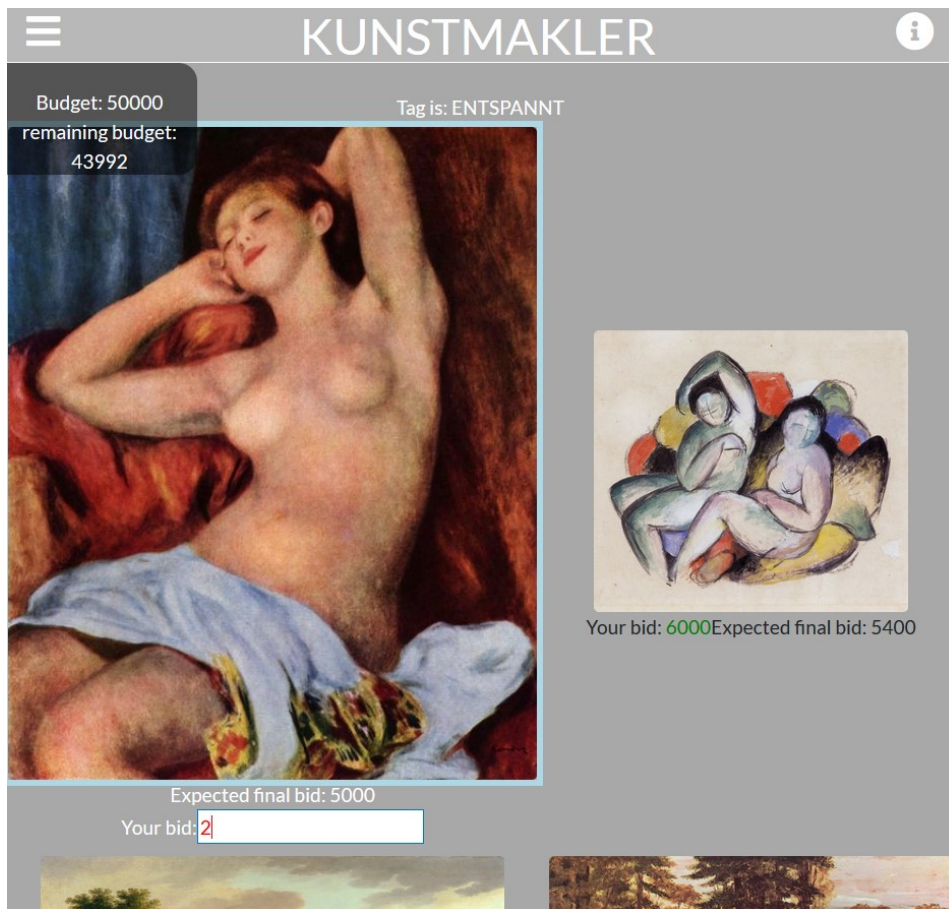


Figure 2.1: Screenshot from the third game screen, the actual bidding. The artwork on the left is currently selected and therefore enlarged. The button to submit the bids and more images are below the area shown.

Design Considerations and Decisions

This chapter presents the factors that had to be considered when designing the Kunstmakler Game With A Purpose as well as the decisions these considerations resulted in. It will mostly go through the game from more abstract concepts influencing the design to small factors affecting concrete parts of the game flow.

3.1 Weighing Gameplay Against the Data Collected

The first major decision regarding the design of Kunstmakler was to go against most other GWAPs, and make concessions in regard to data generation if the gameplay experience stands to benefit. While the acquisition of quality deep semantics taggings is not the absolute focus any more, it remains a major factor to be considered in all aspects. One of the results of this decision was to add gameplay elements that do not primarily serve data collection, such as almost the entire second stage (the actual bidding screen), whose sole contribution data-wise is the confirmation of images the player flagged with “maybe” in the first stage. Also the entire money-related system on that screen has no purpose other than to tie the gameplay closer to the narrative, as well as the intro screen presenting said narrative and to add complexity to otherwise overly simple gameplay. They take up the player’s time in which they are not contributing to the Artigo database, but the overall more rounded and therefore more enjoyable gameplay experience should make up for the reduced net productive output per session or per playtime by letting players stick longer with the game and therefore in the end generate an adequate or even greater amount of data.

3.2 Premise and Narrative of the Game

The narrative element of the game was introduced to unify the game elements, mechanics necessary to solve the underlying problem (selecting images fitting for the tag), and other constraints into a concise gameplay experience. The biggest factors for this were the content of the game – artworks from the Artigo database – and the necessity for the player to make a selection thereof considering the current tag. For these constraints the player’s role of an art broker looking for art for an upcoming, themed exhibition was a good match. The

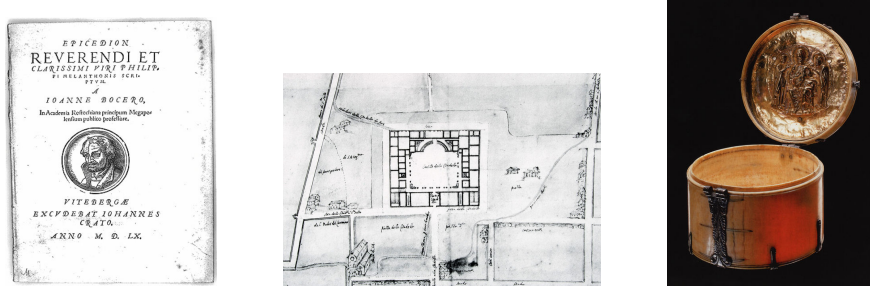


Figure 3.1: Example images from the Artigo database which are not fit for the narrative of the game. From left to right: Johannes Crato - Bildnismedaillon des Philipp Melanchthon, 1560; M. A. Dordoni - Grundriss und Umgebung des Palazzo Farnese in Piacenza, 1558/1600; Anonym - Elfenbeinpyxis, 1100-1300

auctioning was added to enrich the gameplay aspect (while also allowing for secondary data to be raised) while the specifying of a given tag integrates naturally into the story, since a lot of exhibitions tend to have an overall theme or concept. It also directly lends itself to more abstract concepts as proposed themes, which are the kind of tags this thesis is focused on gathering data for.

3.3 Selection of the Artworks for the Game

A good selection of the artworks “auctioned” to the player is essential for the success of the game. How ‘good’ a selection is, is defined by how well most of the artworks fit for the given tag as well as the game as a whole. From the data-gathering perspective a good selection contributes to the amount of data gathered, since a declined image generates no tagging and is simply skipped instead.¹

For the gameplay, a good selection of images is useful in facilitating smooth gameplay by reducing frustration which would arise if the player had to click through seemingly endless unfitting images before finding a few actually fitting artworks. However, a perfect selection would also lead to boring gameplay by having the player always pick the artwork without much consideration or thought. This could also lead to players mindlessly selecting all images (including the few not appropriate ones), which would generate bad data. Additionally, the Artigo database includes Images, which do not fit the premise of the game and would break the immersion if they were shown alongside context-appropriate artworks. Examples of these are tobacco boxes, photos of vases, title pages of medieval books and similar. These should be excluded independent of any currently played tag. When gathering data in the form of a GWAP, care has to be taken to not introduce biases in any direction, or to keep unavoidable biases to an absolute minimum. One example of a potential systemic bias would be that images which are added to the selection for a given tag may eventually be verified simply due to them being shown for it in this game. Given that the preselection is likely not perfect, some images could be flagged for a tag simply by chance, at which point the confirmation bias could lead to the tag being verified. Measures to counteract this bias have to be implemented. One such countermeasure is in removing artworks for the playlist of a tag if they have been declined multiple times.

¹The declining of an image is used to potentially exclude this artwork for this tag in future games, but not evaluated for the general Artigo ecosystem as ‘negative taggings’ have no broader use on the platform.

3.4 Tag Selection

Selecting appropriate tags to be presented as the game-rounds theme follows similar constraints as the selection of the artworks: tags that are too superficial or boring could break the – in a browser game already fragile – immersion and lower the player’s enjoyment, while also not fulfilling the data-generating goal of this thesis: to gather deep semantics taggings rather than superficial annotations. Since the players cannot freely enter words as they like, any game-rounds with a surface level tag will not yield the desired data and therefore miss that goal; as well as being very similar to the not-gamey newer reCAPTCHAs where the user has to select all images with a certain object . Since user know these as a mild inconvenience or annoyance, this association would not be beneficial in regard to motivation. Several approaches of generating a list of such fitting tags were considered, but due to scope constraints, most of them can only be found in the Perspectives7 section. The approach chosen in the end was the simplest possible: manually pre-writing a short list of tags fitting. The script-based interface for adding new tags is fully implemented and simple to use, so any user-interactive approach of generating new tags or a motivated administrator could immediately start expanding the selection of playable tags with little effort. ²

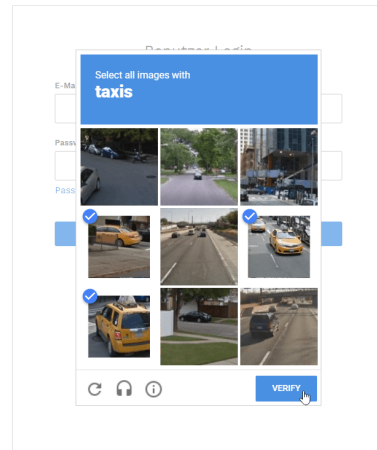


Figure 3.2: Example of the Google ReCaptcha v2, a widely disliked method of bot prevention where users also select the most fitting images out of a selection

3.5 User Interface

This thesis is in regard to computer science and appropriates some aspects of game design, but the topics of user interface and user experience are not the focus of this work and as such the interface built and presented in the screenshots of the example play-through 2.2 are mere prototype-placeholders which, while functional, are not as appealing as desired. This lack of polish in the UI and UX department is a problem with the Artigo platform in general and is likely hindering it from achieving its full potential. Some considerations however are implemented in the prototype and should be present in any polished version as well.

First the question of how many images to show at the same time. The original ESP or Artigo Game only ever present the player one per session to think about, while Karido utilizes a 3 by 3 grid totalling 9 images to compare at the same time. And these two different characteristics both come into play in Kunstmakler. In the first round, the player should decide about each artwork individually, the previous image perfectly fitting the given tag or not should not influence the decision for the current piece, since in the Artigo database, such comparative context is not evaluated and this bias could introduce unfitting taggings (bad data). While it cannot be totally avoided that a player for example may tend towards agreeing to the first barely passable piece after already declining 10 artworks in a row or falling to similar psychological effects, only showing one piece at a time minimizes the subconscious comparing between the artworks, leading to taggings that are cleaner in regard to how they are mainly evaluated (in an aggregate across games and sessions without accounting for any further context). In the second round however, players should compare images: the only data gathered is regarding those pieces they flagged as “maybe” during

²Exemplatory tags include: “love”, “peace”, “democracy”, “melancholy”, “majestic”, “anger” and others

the first round and so the player now needs the context of what – according to themselves – constitutes a fitting artwork. To give the players this context, in the second stage all images are presented next to each other and without any time or turn limit pressuring hasty decisions. This difference can also be explained to the player with in-narrative logic: in the first stage the player is browsing through the catalogue or gallery of the auction house, noting down artworks to bid on. For this an art broker has no limit as to how many selection he wants to make (hence they can immediately jump to the second stage after selecting only a few or up to 50 pieces of art), there is no need to directly compare them to each other, they are either interesting or not. In actual auction however, one has to take the available budget into account, and every image has to be worth the money that cannot be spent on others.

For another detail, the trade-off of user experience and immersion versus data collection was decided against the user experience: Up until the very last screen, the player does not see any metadata for the artworks (e.g. artist name, title of the image, date of creation), despite them being readily available in most cases. This majorly contrasts the narrative framework of it being an art auction, because such information is always provided at real world auctions if possible, in order to enable an informed purchase decision. This game with a purpose however is specifically not about informed (regarding metadata) decisions, but rather about intuitive and subjective judgments whether the specific piece of art fits for the theme of the game-round. Potential words in the title of an artwork are already considered in the Artigo search engine, so additional taggings because of this information would not show any new images if searched for and the taggings would not provide much value to the platform. For information regarding the artist or the current owner of the painting, well renowned names could lead to the player attributing inherent value to the image, which while possibly correct in the real world, would go against the goal of selecting images fitting for the theme given. This would introduce a potentially significant bias in the data gathered on such artworks, which is why this type of information is only given for images presented at the end, where no further data is gathered. The specific content of the end-screen is two faceted: For one, the game should come to a conclusion, the player should know which artworks he managed to give the top bid for, giving the player a perceived success and reward for playing (partly by showing the full image's metadata now). Whether to rely (partially) on some randomness to determine which images the player actually manages the top bid on or to just use the highest bids was heavily debated. On one side, "loosing" without any influence can be frustrating and discourage further play. On the other side, random chances to "win", especially in their implementation as loot-boxes, currently enjoy great popularity in the video game industry, generally for their addictive nature which aids in player retention. The current implementation decided in favour of the latter, though A/B testing on a larger scale could overturn this decision and changing it in the implementation would be trivial.

The second facet is a completely different approach at rewarding the player: Any images that can now be found in the Artigo search are highlighted as well. This idea from Stefanie Kloss's thesis future works section highlights the players contribution to the Artigo project as a whole and reinforces that they are not just playing a game but are also part of something bigger, and lending meaning to the time spent on the round, a significant motivating factor in video-games [13, chapter 6, p. 95].

This chapter gives a brief rundown of the implementation of the game and highlights noteworthy solutions.

4.1 Frontend and Backend-Communication

The frontend of the game is a fairly regular React App, written in JavaScript with some HTML and CSS. The basis of the code was originally written as part of Stefanie Kloss' bachelor thesis for the mobile version of the Karido game called "Dico" [10]. The Kunstmakler game however has support for multiple languages, implemented by replacing every string with a pointer to a file holding an enum for each entry which then contain the actual strings for each language. Her project was also extended by a secondary entry-point into the app which then directly loads custom CSS and instantly requests a new Kunstmakler game-session from the back-end. Communication with said backend is managed by a single Class "Connection.js" which serves as a wrapper for the popular socket.io library. The communication always consists of a JSON-formatted string containing all currently relevant pieces of data for the game, where the front-end merely adds some values and then sends the entire object back to be changed for the next stage of the game.

When the website is called, it opens with Dico's "MainScreen" class, which then loads a header bar and Kunstmakler's wrapper class "KunstmaklerGame" for the body of the screen. This class functions by rendering the game screen appropriate to the session data received from the back-end. Upon initialization, this class requests a new game-session from the back-end (A) and renders a placeholder. Once session-data is received (B) (with the current game-state as "intro"), the introduction screen is rendered. When the player clicks "next", game data contained in the

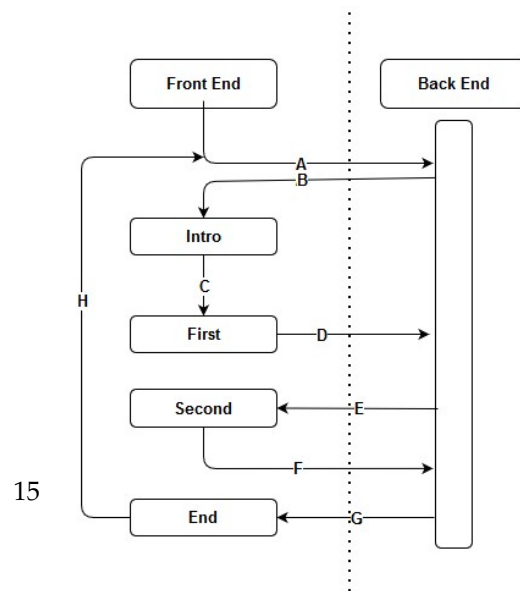


Figure 4.1: Flowchart of a play-through

state ¹ of the `KunstmaklerGame` instance is changed to indicate the first selection screen. The corresponding screen is then rendered (in a React app, a change to the state directly triggers a re-render from the top down where necessary, which in this case leads to the Intro screen being replaced by the first selection one) without further communication with the back-end (C). This is possible because all data required for the first selection screen is already included in the first game state received from the back-end. The first selection screen keeps the game state object received intact and merely adds a value for each image the player has judged. The game state object always contains the following attributes: `gameSessionId` (integer), `tag` (JSON object consisting of the `tagid` in the database, name and language of the tag), the budget to be played with, the round-state (one state for each screen of the game, intro, first- and second bis as well as the end-screen), and an array of the `artresource-data`. This array contains another object for each image that could be played, including the `tempid` that is used by the browser to fetch the actual image data through Artigo's obfuscator. To this object, the judgement of the player is added as the attribute `"firstBid"` with the values 0 for declined, 1 for maybe and 2 for chosen. Once the player chooses to commit to his selection, this JSON string is sent back to the back-end (D), which returns the session data changed for the second screen (E) (`artworks` array trimmed to those that were selected plus some 'maybes' each with an `"expected final bid"` given, as well as round-state on `"second"`). The second and third screens function similarly. The end-screen can invoke a new round simply by requesting one from the back-end (H / A), once the round-data is delivered (B), it automatically overwrites the old data and immediately the intro screen is rendered anew.

4.2 Back-End

On the back-end side, the actual `Kunstmakler` code is embedded into a nodeJS based rewrite of the Artigo platforms codebase that was written by Nicola Greth as part of her Master thesis regarding automatic semantic categorization of tags gathered *g*. This codebase already included game-session classes for all of the different games on the Artigo platform, even though they all are based on a structure of multiple game-rounds in one game-session, which had to be overwritten when inheriting from the base class. The back-end works by simply waiting for connections to come in from the separate client frontend (or Dico's / `Kunstmakler`'s "mobile frontend") requesting a game-session for a specific game type. With `Kunstmakler` being the only single-player game a `KunstmaklerGameSession` is immediately created without waiting for other players or looking for already open game sessions. In the initialization process this class gathers a playable tag from the database, as well as artworks that will be played (the tags are already in their own database table with a second table containing artworks likely appropriate for each tag), this session data is then sent to the requesting front end starting the round. Once the data from the first selection round comes back, any artworks selected have the tagging (relation between tag and artwork with the metadata of time and player) persisted.

4.3 Tag and Artwork Selection

The process of selecting the images to be shown in game has three stages and starts with a general preselection independent of any playable tags. In order to address the problem of images not fitting into the general narrative, all images in the database are compared

¹a React concept to keep all relevant data in one place

against a whitelist of tags which suggest an artwork fitting for the game, and only images which have one of these tags confirmed are selected. This selection is persisted in its own database table, which is then used for a tag-specific secondary selection. The criteria for this selection is whether an artwork has already been tagged with the tag in question, but not often enough for this tag to be validated for the image². This way the images selected have at least once been described with the tag by a human in a different game, and are therefore likely to fit. Images with more taggings are even more likely to fit, but the data gathered for playing with them would add far less value, since they are already confirmed. The implementation of both selection steps is quite simplistic yet still produces adequate results. More complex approaches were considered but would shift the focus of this thesis to data mining instead of gamification and games with a purpose. When a game is started with a certain tag, 40 images are chosen randomly from this second selection at runtime. The only constraints are, that these image cannot have been declined more than 3 times in the first stage of the game to avoid potential confirmation bias in the gathered data. Additionally, images where the tag has been confirmed have a lower priority for this selection, in order to gather data on images where a confirmation could be achieved, which contributes more to the end goal of the Artigo project: making artworks be searchable with keywords. After this, 10 other images from the general whitelist selection are added to the playlist to purposely dilute the otherwise too good selection. These 10 random images will statistically not fit for the played tag, ensuring actual decisions and breaking potential monotony as laid out in Design Decisions and Considerations 3.3

The process of selecting the tags to be added to the game is the simplest viable option and therefore fully manual. Once a fitting annotation is found, the nodeJS command `"kunstmakler newtag *tagName* *language: de/en/fr* *type: noun/adjective/adverb*"` will persist the tag in the appropriate table and will run the aforementioned artwork selection for this tag. If enough images were found as likely fitting, the tag will be flagged as playable and immediately has a chance of being served to the next player starting a game.

²Originally explained here: 1.4 This threshold is currently a single digit number[2]

Challenges and Problems

5.1 Are the Taggings Generated Actually Deep

Since the players do not come up with the tags themselves, but are instead presented with a singular tag, the question arises if the annotations (which are merely words) raised from this game are still deep semantics tags. For example, “liberty” is an abstract concept and seemingly of deeper meaning, but if an image simply contains the lettering “liberty” 5.1 players would very likely select that image resulting in a tagging, which clearly constitutes a surface semantics tagging, since it is merely reading the word. Other scenarios of tags played resulting in rather surface level taggings could also be possible. However, such cases are unlikely to appear often in Kunstmakler since most images containing text are filtered out in the first stage of artwork selection ¹. Those that make it through that filter – or in cases where the tagging would be superficial not because of writing – should then not show up for the tags corresponding to the written words, since in Artigo’s main games such obvious descriptions are usually tagged first and therefore the tag should already be confirmed, removing it from play in Kunstmakler. The few that do end up in the game can be accepted since they are not bad taggings, they merely miss the goal of this game, but are still valid for the Artigo platform in general and as such constitute only a lost opportunity to generate rich tags.



Figure 5.1: William Hogarth - John Wilkes (First State), 1763 From artigo.org

¹While not actively filtering for text in the images, the blacklisted tags overlap greatly with those artworks containing text

5.2 Shallow Image Selection

Limiting the Selection of Playable Tags

One stumbling block for the game could be the very simplistic artwork selection which could cause tags added to the game to not be playable (with the partial database dump used for development this were 2 out of 12 tags added). This is considered acceptable, since there is no limit to the amount of tags added, if too many are flagged as not playable, new ones could be supplied instead. But this problem should not come up since the Artigo project has been running for over 10 years, and as such there is an extensive collection of images and taggings to draw from – a problem however for any attempts at transferring this thesis to other projects or contexts which lack such data pools to draw from. The drawback of potentially unplayable tags is also outweighed by the quality of the selection, especially considering its simplicity. The fact that at least one human player on another game on the platform had to see the artwork and then had to come up with the tag on their own results in a – subjectively- very fitting selection, to the point where ‘wrong’ random images had to be added to the selection presented. Since the artwork selection is not the focus of this thesis and merely the means to enable the rest of the game to work, this implementation suffices for its purpose.

5.3 Data Integrity

The integrity of the data levied from a GWAP is of utmost importance. If a game with a purpose were to go viral and become a smash hit, it would still be almost worthless scientifically, if the data generated by the players could not be trusted or meaningfully analysed. On the Artigo platform, this is the case for taggings which relate a tag to an artwork that is not described by the phrase. Since tags are descriptions this would be annotations that do not describe the image accurately. Because the principle of the platform relies on the subjective descriptions of art, objectively bad taggings are hard to define, but there are clear sources of outright bad or merely not-so-good taggings (meaning users that search for the tagging in the database do not gain from a certain artwork showing up for the term). One source of clearly bad data are malicious actors, or trolls. These are players that purposely “play” the games wrong, adding false data. The frequency or occurrence of such players has not been analysed on the Artigo platform, but since a troll’s enjoyment (and thereby motivation) mostly stems from social feedback [9] and the platform limited social interactive elements, this potential polluting factor is not likely to occur in a relevant frequency, outside trolls are unlikely due to the low publicity of the platform as a whole. Especially as a single-player game with a likely niche focus, Kunstmakler is an even more unlikely target. Secondly, even if trolls target this game, the tags are predetermined and any artwork shown has at least been tagged with it once, so the potential damage is quite limited. The random “wrong” images mixed in cannot be tagged, further minimizing the potential damage. These known (or likely) false inputs could theoretically be used in tandem with known confirmed images to determine potential malicious actors, but this added complexity was deemed not necessary as of now. Another more benign cause of unclean data are biases baked into the game itself. The human mind is not a purely logical computer and instead prone to numerous subconscious biases² and as such, outsourcing computation to sentient beings carries the risk of such biases tainting the computed outcome (in this case, the annotations collected). While the topic of human biases is incredibly complex and obscure for people outside the field and as such designing a completely bias-free game is near

²see https://en.wikipedia.org/wiki/List_of_cognitive_biases

impossible, each aspect of the game has been designed with the goal of avoiding obvious sources of bias.

CHAPTER 6

User Feedback

Due to the current COVID-19 pandemic, gathering feedback from users has proven more challenging than usual, which is part of the reason why the user feedback for the *Kunstmakler* game has a somewhat small sample size of 7. The methods for gathering user feedback was loosely based on the “Game Design Workshop” [5, chapter 9] and advice from Tobias Biber¹ based on Peter Neumann’s work on market analysis [14]. The testing was conducted in form of a small questionnaire regarding background information deemed relevant such as age, general interest in art, and gaming habits. Afterwards the participants were asked to play the game without further input but under observation. Afterwards a semi-structured interview was conducted, with the pre-written questions regarding the general impression of the game, the perception of the narrative context, general enjoyment and whether they would play or recommend the game to others once it was published. The users were told that the interface itself was still a work in progress and to focus on the gameplay itself. In the unstructured part of the interview, a non-directive approach was taken in order to let the users focus more on the topics at the top of their head. Testing was conducted with 7 individuals aged 20-55, with overwhelmingly academic background and higher art affinity, somewhat representing the target demographic, though feedback from not-targeted individuals was raised as well.

The findings are roughly as expected, including the more critical passages of this thesis. This includes rather negative overall sentiment regarding the user interface, lower appeal to core gamers and non-art-enthusiastic individuals. The feedback on the narrative ranged from annoyance at the additional step before the actual gameplay to praise for how it “ties everything together”, but on average a neutral stance on the topic. Since this part of the thesis was introduced as a way to put *Kunstmakler* into a specific niche and out of competition with Artigo’s other games this a welcomed result despite the absence of widespread enthusiasm, since players enjoying the niche approach were found. General enjoyment was average as well, most users being either unfamiliar or unmotivated for the format of browser-games citing them as outdated, but could see the game as an entertaining past-time with a more polished flow on mobile devices. The same general sentiment was found when asked whether the users could see themselves play or recommend the game after an official release: the users either stated to not spend much time in front of a computer or preferred more complex games when they did. Impromptu tests for use with a phone-

¹Graduate in marketing and market analysis

sized screen resulted in frustration due to the layout being optimized for the screen sizes of laptops or large tablets.

In summary: the game fulfils the expectations as a prototype for a GWAP filling the niche for more story-interested players.

CHAPTER 7

Perspectives

With the limited scope of a bachelor thesis, some features originally planned had to be cut and several interesting ideas could not be explored in the detail they deserve. Furthermore some work could be done based on this thesis once the game has been running on live servers for some time. These ideas and opportunities for future knowledge gain are presented here.

7.1 Expanding the Gameplay

A full-blown game (with a purpose) using a narrative could extend the rather shallow attempt of this game at providing an immersive experience. While such a game would directly go against the grain of casual gaming, the success story of FoldIt shows that GWAPs do not necessarily have to follow this path. While such game should likely be developed from scratch, the narrative and problem solved by the Kunstmakler game lends itself to such a feature-rich experience as well. A number of features could be added to the current game to expand it into a more immersive experience while driving it slightly further away from the casual game it currently is. One of these extra features could be a set of galleries participating in the themed exhibition event that starts the story. These could each have a secondary set of preferences (like only classical art for one gallery), and their own budget, adding another layer of interactivity to the currently quite short gameplay and adding to the storyline of each play-through. These galleries could be further expanded by a social aspect where at the end the exhibition event is shown across multiple galleries, each fitted with the player's artworks as well as those other players selected with potentially the name of the providing art broker (that player's username) presented. More depth to the gameplay could also come from a budget persistent across multiple sessions, where some initial funding is supplied for each round and some more made back from loaning the art to the galleries. A small step from there would be a player's own gallery where they can present artworks they bought to other players, or just for their own viewing pleasure. Another addition could be dynamic bidding against real players or merely simulated ones, for a more auction-like feel that is somewhat lacking in the current implementation. This could be further expanded into a full virtual economy centred around the artworks with revenue from the exhibitions being spent on a wider market and in the extreme even trading between players. This would probably be a stand-alone game in its own right and would be very

complex in implementation as well as in gathering useful data from it.

Smaller expansions could be in the enhanced gamification of the already implemented gameplay, such as achievements for special goals achieved during a session, such as being the top bid on 8 images while remaining 50% under the budget, perfectly maxing out the budget, getting all artworks bid on etc. Achievements are a simple way to add a progression mechanic outside of the core gameplay loop to increase re-playability and present other opportunities for the player's success. [13, p. 21 and 237+] The game could also utilize themed rounds like only showing landscape paintings in one session to spice up the experience and increase re-playability.

7.2 Expanding the Game Behind the Scenes

Outside of the gameplay directly, further improvements could be made. One of these would be to directly integrate the game's tag list into Artigo's search engine. If someone searches for a tag that yields only few results, a pop-up could ask them whether they want to propose their search term for the Kunstmakler game. This could automate the gathering of new tags to be played (maybe with administrative oversight as necessary) as well as integrate the game further with the platform as a whole. Gathering new tags to be played with could also be automated by semantically evaluating tags for being abstract concepts, emotions or similar clusters and adding whole groups of words instead of singular tags to the game. Evaluating statistically or with machine learning whether tags are 'rich' could also be used in conjunction with feedback from the players whether playing a round with a tag proposed by such a system in order to fine tune this avenue of tag selection. Besides from selecting new tags, the selection of artworks for the game is another huge source of potential. As stated in the section about design decisions, the artwork selection implemented as part of this thesis is intentionally close to the minimum viable option. The reason for this is that the current approach suffices for this project, potential for improving the process is endless, and alternative approaches could make up an entire thesis by themselves. For the small improvements in the first step, selecting images generally suitable for the Kunstmakler game as a whole, the ex- or including of certainly tagged artworks could be fine tuned and expended, as well as combining the current including list (only artworks tagged with one of multiple tags are valid for the game) with further excluding filters, removing some images that slip through the first list. Expanding both lists and statistically weighing or counting the taggings for each tag (currently one tagging is enough to qualify an image for the game) could all improve the output of this still simple approach. For the second step of selection (finding artworks appropriate for each tag in the game) more complex approaches were considered, currently, pieces of art are selected, if they have already been tagged with the tag in question, but not often enough to count as confirmed. This relies on Artigo's other games to provide a good base to work off and for some tags still fails to find a sufficient number of images to enable the tag. One alternative could be using semantically close tags to be searched for as well. These similar tags could be provided manually, but also sourced automatically from games like Tag-A-Tag [4], Nicola Greth's work semantically clustering tags [7] or other data about relation between tags or between words in general (for example by using word embeddings from other sources like word2vec trained on general texts). Using such a selection could – with the necessary metadata saved in the database – then also allow for such data to be validated in return when players play on images selected in such a way and accept or reject them. The previous improvements already mentioned increasing or improving the data gathered from each play-through, but changes focusing solely on this topic could also yield great results. For one, the actual bid on the artworks could be further analysed. From the amount bid on each (already confirmed) image could result in data regarding certainty or an order of fitness regarding the tag in

question. Such weighing of taggings persisted is currently only done in the Dico game and not evaluated at all in the Artigo project as a whole, meaning that such an extension would have to be driven by a meaningful purpose for this type of data and would require more modifications across the multiple parts of the platform. The more complex bidding system mentioned above could also be modified to generate such data in a more informative manner. A smaller, capsuled improvement to quantity of the data gathered could be to supply either the back-end or just the player with the same tag in multiple languages. This is interesting, because the German version is by far the most extensive, with French and English lacking in comparison [3], and with this extra, a play-through in any language would yield the tagging data for all languages if the translated tags have been supplied. This would be more reliably feasible for manually entered tags, as different connotations across languages and multiple meanings for the same word in one language would have to be checked very carefully as to not introduce bad data, and as such direct translations, however sourced, could prove problematic and their output would have to be closely monitored and analysed, especially for automatically generated translation pairs. Games for generating such tag-translations were originally envisioned for the Artigo platform, and if implemented, the data from those could eventually be used as well for this type of cross-language tagging, with feedback for those games, if the identical round is played by players of a different language, the difference between their choices could be used to analyse their output as well. An improvement more noticeable could be actual optimisation for players on mobile devices. While the front-end code is loosely based on a mobile-optimized version of Karido, named *Dico*, not much of the changes to accommodate this platform have carried over to the current implementation of Kunstmakler. As such anything from small optimisations in layout to larger changes such as caching sessions for complete offline play (which could work due to the single-player nature with the data gathered being sent back later) or reimplementing the frontend as a stand-alone app could help the game succeed in this ever-growing category of gaming which has a strong tendency towards casual gaming [1] and would therefore perfectly fit this game, as also mentioned by users questioned about Kunstmakler. The current implementation as a react app would also allow for both layouts to work from the same code base with different layouts dependent on screen size applying automatically. Another small improvement would be the detection for bad actors, which was forgone due to the risk being considered rather low. It could still prove valuable even to determine the occasional player merely clicking through the steps to see the game and invalidate their data. The required known false artworks¹ are already in the game and finding known positives for each tag should be trivial for most tags so the requisites are almost present already.

7.3 Evaluating the Game's Success

One last piece of work regarding this thesis still remains: large scale testing and feedback as well as evaluating the game's results once it is running on the live Artigo version. Points to be analysed should include the user interface for quick adjustments to the look and feel of the game, but more importantly the acceptance and enjoyment of the general player. As stated in the design considerations, this game targets a different niche than Artigo's other games and as such universal acclaim is not to be expected, but whether there is a significant player-base that does enjoy this niche would be worth examining. Furthermore the quality of the levied taggings should be scrutinized, especially regarding a possible confirmation bias: after an artwork gets tagged just once with a certain annotation, it is eligible to be played with that tag, does this initial tagging lead to confirmed taggings in cases where

¹this refers to the random images mixed under those which were deemed as likely fitting for the tag by the second artwork selection stage. Statistically random images are highly unlikely to fit to the tag given for the round

it should not have. The small testing and feedback rounds included in this thesis do not suggest such problem, but with such a small sample size no conclusion should be drawn from it.

Bibliography

- [1] Entertainment Software Association, *2020 Essential Facts About the Video Game Industry*, https://www.theesa.com/wp-content/uploads/2020/07/Final-Edited-2020-ESA_Essential_facts.pdf, last accessed: 2020-11-08.
- [2] Martin Bogner, *Conversation with Martin Bogner, research assistant with the Artigo Project*, personal communication, "During regular meetings as part of writing this bachelor thesis the items cited were discussed. He is a scientific employee at the LMUs Institute for Computer Science which hosts the Artigo Project".
- [3] François Bry and Clemens Schefels, *An analysis of the artigo gaming ecosystem with a purpose*, <https://www.en.pms.ifi.lmu.de/publications/PMS-FB/PMS-FB-2016-6/PMS-FB-2016-6-paper.pdf>, 2016.
- [4] François Bry and Christoph Wieser, *Squaring and Scripting the ESP Game: Trimming a GWAP to Deep Semantics*, *Serious Games Development and Applications* (Berlin, Heidelberg) (Minhua Ma, Manuel Fradinho Oliveira, Jannicke Baalsrud Hauge, Heiko Duin, and Klaus-Dieter Thoben, eds.), *Lecture Notes in Computer Science*, Springer, 2012, pp. 183–192.
- [5] Tracy Fullerton, Christopher Swain, and Steven Hoffman, *Game design workshop: a playcentric approach to creating innovative games*, 2nd ed., Elsevier Morgan Kaufmann, Amsterdam ; Boston, 2008, OCLC: ocn173659700.
- [6] Lindsay Grace and Peter Jamieson, *Gaming with purpose: Heuristic understanding of ubiquitous game development and design for human computation*, *Handbook of digital games* (Marios C Angelides, Harry Agius, and Grace, eds.), 2014, OCLC: 1159672916, pp. 645–680.
- [7] Nicola Greth, *Automatic semantic categorization of image annotations generated by games with a purpose*, masters thesis, Ludwig-Maximilians-Universität, Munich, Germany, 2019.
- [8] Omer Kaplan, *The truth about hypercasual games*, <https://venturebeat.com/2019/03/24/the-truth-about-hypercasual-games/>, March 2019, last accessed: 2020-11-07.
- [9] Ben Kirman, Conor Lineham, and Shaun Lawson, *Exploring mischief and mayhem in social computing or: how we learned to stop worrying and love the trolls*, *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12* (Austin, Texas, USA), ACM Press, 2012, p. 121.

- [10] Stefanie Kloß, *Designing an asynchronous image labelling game for mobile devices*, bachelor thesis, Ludwig-Maximilians-Universität, Munich, Germany, 2020.
- [11] Mark Lagace and Jordan Palmer, *Enjoy your commute or down-time with one of these casual games for Android*, <https://www.androidcentral.com/best-casual-games-android>, September 2020, last accessed: 2020-11-08; Archive-Link: <https://web.archive.org/web/20201108143914/https://www.androidcentral.com/best-casual-games-android>.
- [12] C. Lindley, *Story and narrative structures in computer games*, <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.797&rep=rep1&type=pdf>, 2005, Gotland University, Sweden.
- [13] Jane McGonigal, *Reality is broken: why games make us better and how they can change the world*, Penguin Press, New York, 2011.
- [14] Peter Neumann, *Handbuch der psychologischen Marktforschung*, Huber, 2013.
- [15] Su Rocks, *What are hyper-casual games?* - Quora, <https://www.quora.com/What-are-hyper-casual-games>, last accessed: on 2020-11-07.
- [16] Michael Sailer, *Die Wirkung von Gamification auf Motivation und Leistung: empirische Studien im Kontext manueller Arbeitsprozesse*, Research, Springer, Wiesbaden, 2016, OCLC: 951120637.
- [17] B. Steinmayr, C. Wieser, F. Kneißl, and F. Bry, *Karido: A GWAP for telling artworks apart*, 2011 16th International Conference on Computer Games (CGAMES), July 2011, pp. 193–200.
- [18] Bartholomäus Steinmayr, *Designing image labeling games for more informative tags*, Diplomarbeit, Ludwig-Maximilians-Universität, Munich, Germany, 2011.
- [19] Luis von Ahn and Laura Dabbish, *Designing games with a purpose*, Communications of the ACM **51** (2008), no. 8, 58–67.
- [20] Kevin Werbach and Dan Hunter, *For the win: how game thinking can revolutionize your business*, Wharton Digital Press, Philadelphia, 2012.
- [21] C. Wieser, François Bry, Alexandre Berard, and Richard Lagrange, *Artigo: Building an artwork search engine with games and higher-order latent semantic analysis*, AAAI 2013, 2013.
- [22] Wikipedia, *Casual game*, https://en.wikipedia.org/w/index.php?title=Casual_game&oldid=975622628, August 2020, last accessed: 2020-11-07, Page Version ID: 975622628.
- [23] ———, *Hyper-casual game*, https://en.wikipedia.org/w/index.php?title=Hyper-casual_game&oldid=980170394, September 2020, last accessed: 2020-11-07, Page Version ID: 980170394.
- [24] Kevin Wolstenholme, *Top Hyper Casual Games - June 2020*, <https://risinghighacademy.com/the-best-12-hyper-casual-games-june-2020/>, June 2020, last accessed: 23.08.2020; Archive Link: <https://web.archive.org/web/20201020032434/https://risinghighacademy.com/the-best-12-hyper-casual-games-june-2020/>.

- [25] ———, *Top Hyper Casual Games - March 2020*, <https://risinghighacademy.com/the-top-12-hyper-casual-games-april-2020/>, April 2020, last accessed: 2020-08-08; Archive-Link: <https://web.archive.org/web/20201020022609/https://risinghighacademy.com/the-top-12-hyper-casual-games-april-2020/>.