Virtual Editions in the LdoD Archive using Crowdsourcing and Gamification

Gonçalo Medeiros Serapicos Montalvão Marques
goncalo.m.marques@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisboa, Portugal

October 2018

Abstract

Today's Digital Age has changed our way of life, influenced our relationships, our culture and much more. An important breakthrough to this age is the World Wide Web, which in turn caused the emergence of new fields, one of them being the Digital Humanities. Crowdsourcing has emerged with the Web and has become an important sourcing alternative in business as well as in Digital Humanities. Besides crowdsourcing, another process that has risen is Gamification, which focuses in increasing user engagement. This thesis seeks to harness the crowdsourcing and gamification potential, from the Digital Humanities perspective, specifically in the context of the *LdoD Archive*. The *LdoD Archive* is a collaborative digital archive of the *Book of Disquiet* by Fernando Pessoa. It allows users to read and explore the book in a digital way, such as through the creation of Virtual Editions. Therefore the goal of this thesis is to enhance the user experience in the *LdoD Archive* using crowdsourcing and gamification techniques.

Keywords: Crowdsourcing; Gamification; Fernando Pessoa; *LdoD Archive*; Digital Humanities;

1. Introduction

In today's age of our society we can access, learn, engage, communicate, create and much more online. From an organizational perspective there is an opportunity to reap the benefits of that potential that exists in every Web user. Crowdsourcing is a way to achieve that, an example is Wikipedia, which is a collaborative online encyclopedia where users can read, create, extend and review information, and it's all done by volunteers. The emergence of Web platforms and sites, resulted on the surge of crowdsourcing, making it both a business alternative as well as useful option in Galleries, Libraries, Archives and Museums (GLAM's) industry.

Although the definition is very recent, 2006 [7], the notion exists from years ago, presumably the very first crowdsourcing project was The Longitude Prize designed by the British government in 1714, that sought a way to measure a ship's longitudinal position and offered a prize to the best solution.

Besides the surge of crowdsourcing, another concept has seen rise in interest which is gamification. Despite the fact that the term was introduced in the 2000's [4], only after 2010, there was a real growth in the use of the concept. In this thesis, we will also study the notion of gamification, through examples, and try to successfully implement a crowdsourcing activity integrated with gamification techniques within the context of a Digital Humanities project.

Fernando Pessoa is hailed as one of the finest Portuguese authors, best known for his countless poems, heteronyms and his *magnum opus The Book of Disquiet* (in portuguese, *LdoD*). The book was published posthumously and is signed under Bernardo Soares, one of Pessoa's heteronyms.

Pessoa died in 1935, leaving behind multiple manuscripts of his work in the form of fragments. These fragments revealed the existence of the *LdoD*, yet the book was found incomplete, without a particular order and with at least two heteronyms as authors of the work. After almost 50 years of research, study and transcription of said fragments, the book was published in 1982 by Jacinto Prado Coelho. However due to the absence of a specific ordering for the texts, three more editions appeared, Teresa Sobral Cunha in 1990-1991, Richard Zenith in 1998 and Jerónimo Pizzaro in 2010, each one with different and diverse interpretations of the book.

These editions diverge on many things such as fragment organization, chronological order, heteronym attribution for the fragments and even spelling. This heterogeneity coupled with the absence of certainty about the author's intentions generates a greater need for the reader to be able

to read and understand each edition and then decide for himself what makes sense to him. As such, the *LdoD Archive*[1] appears as an important tool enabling the construction of multiple reading paths while simultaneously promoting the work of the author. The *Archive* provides tools for reading and comparing editions; creating virtual editions and annotations; and sharing each edition with other users. After understanding the unique characteristics of the *Book of Disquiet* as well as its relation with the *LdoD Archive*, then the question arises **How can crowdsourcing and gamification techniques be used in order to enhance the collaboration in the context of the LdoD Archive?**

2. Related Work

In this section we present some background concepts regarding the *LdoD Archive*; Digital Humanities; Crowdsourcing and Gamification that are of interest in order to understand the work developed. Furthermore, we present the case of ESP game which serves as a baseline for this work.

2.1. The *LdoD Archive*

The *LdoD Archive* is a collaborative digital archive based on the *magnum opus* by Fernando Pessoa, *Book of Disquiet*. The archive follows two main principles: representation and simulation. The history of writing and editing the Book can be seen as the representation principle. Additionally the possibility for users to read, edit and write in the digital medium, embodies the simulation aspect. The archive has multiple goals, it is noteworthy the goal to develop a virtual environment with multiple purposes, such as leisure leisure reading, scholarly research and literary creation; enable the possibility of comparing the four main editions of the *Book*; to encourage different ways to read the *Book of Disquiet* and finally to promote the use of software tools that can be used to change the current analysis and interpretation predominantly based on printed literature.

From a more concrete point of view, the *Archive* has six features:

1. Reading: allows users to read the fragments according to each edition, and knowing where each fragment appears in the other editions. This feature is coupled with a recommendation system, which suggests the user what the next fragment to read given the set of fragments previously read and the following criteria: heteronym, date, text and taxonomy.

2. Documents: this section of the *LdoD Archive* has important pieces of history, mainly witnesses which are elucidations of the fragments made up of interpretations and further-

more of printed texts by Pessoa, alongside some peculiar and interesting metadata. Besides this, it includes encoded fragments.

3. Editions: functionality that enables the viewing of different editions, such as the four main published editions, a virtual edition made by the team behind the *LdoD Archive* and many other virtual editions that can be created by users of the platform which can make it public and available for others.

4. Search: includes two types of searches, a simple and an advanced one. The simple search uses the title as criterion, inquiring in the expert editions and witnesses. In turn, the advanced search allows for the use of a rich set of possible combinations according to several criteria, such as edition, heteronym, manuscripts, taxonomies, text and date.

5. Virtual: one of the most interesting features, enabling the reader to create virtual editions and taxonomies, as well as sharing it with the community by making them public.

6. Writing: finally, this feature is not yet implemented, however its purpose will be to allow the writing of new texts that will be based on the fragments.

2.1.1 Virtual Editions in the *LdoD Archive*

For our work, we expand on the concept of Virtual Editions as it will be a key point in this thesis.

Firstly, as we have seen, the *Book of Disquiet* is comprised by multiple fragments, which if rearranged in different orders (since it does not exist an authorial definitive order) can lead to different interpretations of the book, which as stated originated four main consensual editions among scholars. However, this absence of a natural and obligatory order for the fragments allows readers to have any interpretation and different reading experiences. From this, the *LdoD Archive* created the Virtual functionality allowing users to create their own editions of the book.

So, a virtual edition consists of a selection of fragments made by one or more users. This is a unique entity because it allows its virtual editors to choose the fragments they want to include, their ordering, and their annotation through notes and categories, and a virtual edition can be public or private.

The interesting part of virtual editions, besides its uniqueness, is the collaboration aspect, because a virtual edition can have multiple editors, i.e, users select fragments, categorize and annotate them.

This feature is available at https://ldod.uc.pt/virtualeditions and the only requirement to create and join virtual editions is to be a registered user of the platform.

2.2. Digital Humanities

Digital Humanities, is a field that combines the traditional humanities with the digital environment. The genesis of the Digital Humanities, as David Berry introduces in [1], is very humble, in way that initially it wasn't even a field, simply serving as support for the research of other fields, and being originally called "humanities computing". As it became more relevant, a switch to Digital Humanities occurred and *meant to signal that the field had emerged from the low-prestige status of a support service into a genuinely intellectual endeavour with its own professional practices, rigorous standards, and exciting theoretical explorations* [5]. A good definition is provided by Presner in [8] stating that it's *an umbrella term for a wide array of practices for creating, applying, interpreting, interrogating, and hacking both new and old information technologies.*

From these definitions, we can draw a parallel and understand that the *LdoD Archive* fits perfectly in this area of study.

2.3. Crowdsourcing

With the development of the Web, new areas of interest have emerged, one of them being crowdsourcing. The term crowdsourcing first appeared in an issue of *Wired* magazine [7] and was later defined as: *the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call*. This definition also states that *The crucial prerequisite is the use of the open call* [6].

In the paper *Towards and Integrated Crowdsourcing definition* [3] the authors discovered that forty definitions were published between 2006 and 2011. The authors then synthesized that research into a single definition, where we highlight the following characteristics: *participative online activity* proposed *via a flexible open call to a group of individuals of varying knowledge, heterogeneity, and number*, and the *undertaking of the task (...) always entails mutual benefit*. This definition is our starting point, because it synthesizes many other definitions into one, however we will try to mold and adapt this definition to our specific purpose and context.

2.4. Gamification

Another considerable component of our research is gamification. As we will see this phenomenon has risen as a key factor to motivate user across multiple contexts. For instance, nowadays many applications have turned to gamification. "Todoist" is an example: it's an application for personal and professional use, where users manage tasks among other things. However the key notion is that for each task completed, the user is awarded points, introducing a game element and motivating users to keep on finishing the tasks. Gamification can then be defined as *the use of game design elements in non-game contexts* [2]. The term is relatively recent, having appeared in 2008, but only gained traction in 2010. In [2], Deterding provides a definition for each term. From there we synthesize each term as follows:

- a *game* follows rules and has defined goals;

- *elements* are characteristics encountered in most games and that have a significant impact in the gameplay;

- the *design* is characterized by *varying levels of abstraction* and five designs were identified (ordered from concrete to abstract):

    - *interface design patterns*, for instance badges;

    - *design patterns and mechanics*, time constraint is an example;

    - *design principles and heuristics*, for example clear goals;

    - *models*, a fantasy model for instance;

    - *design methods*, such as playtesting.

- *non-game contexts*, which is not clearly limited by the author, because *there are no specific usage contexts*, however it excludes *the use of game design elements as part of designing a game, since that would simply be game design*.

The relevance of gamification to this thesis is due to the fact that gamification presents itself as a potential solution in order to create, support and boost the interaction between users and the archive, and even among users themselves.

2.5. The ESP Game

The ESP game[2] [9] was a game developed in order to produce meaningful and relevant metadata information. This is a prime example of an idea that harnessed the potential of the crowd allied with the gamification of an otherwise boring process. As such it is relevant to this thesis to try replicate some of the principles applied as well as avoid the deficiencies and drawbacks of its implementation.

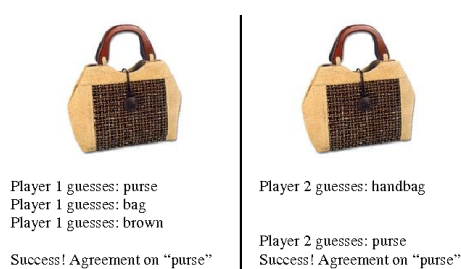[2]https://en.wikipedia.org/wiki/ESP_game

The motivation for the game was the issue regarding the labeling of images on the Web. There is a high level of importance on having rigorous imaging metadata because it is useful for site accessibility (consider applications for visually impaired people); blocking of content not appropriate (graphic images containing pornography, violence and others); overall good search imaging and also for computer vision research (large databases of labelled images used in machine learning).

This categorization had two main ways of being made, computer computations, which have shortcomings such as relying on images descriptions resulting in inaccurate results or manual labelling which is a tedious and costly task.

Therefore, the goal for the creators of this game was to create a way to address the metadata issue while producing meaningfully results in a way that people had fun and were entertained in the process.

The game consists of two random online players from the pool of users in the game and it is made to be played in pairs only.

The flow of the game is very straightforward. Both users have the same image showed simultaneously to them and their goal is to "agree on an image", this process consists during a limited time users must insert labels until each has submitted the same string and then they move on to the next image. However, the participants can pass or opt out of challenging images, in that case a message will appear on the screen of the other partner and a image will only be passed if both have agreed on passing. The users do not know one another, are not able to communicate with nor see the guesses of its counterpart.

Player 1 guesses: purse
Player 1 guesses: bag
Player 1 guesses: brown

Success! Agreement on "purse"

Player 2 guesses: handbag

Player 2 guesses: purse
Success! Agreement on "purse"

Figure 1: Example of players agreeing on a guess [9].

Agreeing on a image results on points, large bonus points are also attributed if they agree on 15 images. The interface also shows the number of images agreed upon. User's decisions are reinforced and they are pushed to keep playing due to point attribution.

Since the participants cannot communicate the only strategy to follow is to type an obvious word related to the image, specially because the system asks players to "think like each other" not to describe the image.

Lastly, in 2006 the ESP game was purchased by Google, with the purpose of being applied to their own online image results, renamed Google Image Labeler[3]. Google's own version was online between 2006 and 2011, and was relaunched again in 2016.

The relaunch maintained the same name, nonetheless it does not resemble any of the game origins. Nowadays it is mostly a classification service, where the user is presented with a category, selects one and can start playing. From there a set of image are presented and the player must indicate if the image belongs to that category.

3. The *LdoD* Classification Game

In this section we introduce the game that was designed for the *LdoD Archive* as well as its rules, flow and game interface.

3.1. Game

The *LdoD* Classification Game is a game that aims to use the *Archive's* Virtual Editions and its ability to classify and categorize fragments as way to create crowdsourced tagged Virtual Editions in a fun and entertaining way.

A game is simply an online, real-time and synchronous environment, where a set of users meets and has the goal of categorizing a fragment of a selected virtual edition, resulting in an enriched virtual edition with new community categories, made in a collaborative and interactive way.

The game follows a mix of two different crowdsourcing processes. It consists in a crowdcreating process whose goal is to create artifacts based on a diversity of heterogeneous contributions and in a crowdvoting process where the *wisdom of crowds* is harnessed to filter the artifacts into one.

Nonetheless, how does this relate to the game itself? Well the game consists of three different stages, one in which each user individually suggests tags (crowdcreating), another where the player sees the suggestions of the other participants and can choose between maintaining his or her own suggestion or to changing the tag (voting other suggestion) and finally a more collaborative one, where all users see the votes and corresponding tags and as group vote in the best tag available for the fragment (a somewhat crowdvoting process).

The game can also be split into two modes: classic and custom. A classic consists of playing an instant game, i.e, the user starts playing with anyone at any date (if other people are available) and classify a fragment at random from a public virtual edition. This mode is designed to be maintained

---

[3] https://crowdsource.google.com/imagelabeler

and developed by the *LdoD Archive* organization by creating multiple games with the sole purpose of entertaining users. On the other hand, the custom game consists of a more private, personalized environment that can be co-related directly to the teacher-students scenario. Its allows for the design of a game for all registered users or only members of the same virtual edition (ideal for the classroom scenario); plan a specific time and date and play together.

3.2. Interface

As stated, the game takes inspiration from the ESP Game therefore the interface will have some similarities.

The game is initiated in a new console, which is opened from the main site, accessible only to the registered users of the *LdoD Archive*. As expected the users will have to log in on the game website in order to participate. The game platform has an About section where users will have the opportunity to re-read game rules, objectives and know more information about the game.

As mentioned above, the game can be split into different stages, hereby called rounds, that have different specific objectives in which the player must comply to. The fragment used in the game is split into paragraphs and their size dictates the time available for each round and determines when the game should switch to a different stage.

The round follow a cycled approach, round one occurs for the first paragraph, round two occurs for the first paragraph. When it ends we switch to round one again however with the second paragraph of the fragment and so forth, until all of the paragraphs have been analysed in these two stages. Subsequently, we reach the final stage, round three, where the fragment text is used in full.

3.2.1 Round 1: Individually submit tags

In this round, the participants goal is to submit a tag for each paragraph of the fragment.

The user will have the first paragraph of the fragment to be placed at the center of the screen and bellow a category submission area. In the bottom part of the page, the user can see a progress bar showing visually how many paragraphs are still missing. At the top of the screen, in the center the user has a steps interface showing their current round and which round follows next. Above that and still in the center they can see the time remaining in the current round - note that the time is variable according to the size of the text. At the top left, the users can see the number of people participating online.

Summarizing, the player must in the given time, read the text carefully, think and submit one tag they believe is appropriate. The allowed input is equal to the input available in the *LdoD Archive*, which is anything above one character.

The time and occurrence in which a tag was suggested also plays a factor having impact in the calculation of scores. Let us consider player A suggests a category of *dreams*, also in the same stage a player B submits ten seconds later this same tag. Player A is considered the author of the tag and player B is a co-author or voter of the tag submitted of player A, this is important regarding the players scoring.

After the time elapses we move to the next step, which is the round 2 for the same paragraph. Note that the user returns to the round 1 after round 2 but with the next paragraph of the fragment.

3.2.2 Round 2: Choose a tag

In round 2, the interface is very similar, showing again the same paragraph of the fragment, but now bellow, it will have the categories that have been suggested by all participants in the earlier round.

The goal in this stage is to once again find the best category available although now the user has the ability to choose from a range of options, as such the user can now choose/*vote* only in one category that the user agrees to be the most suitable for the paragraph that was analysed, the user can even choose the category that he/she submitted.

At this point the game becomes more interesting because the players now must simultaneously find a category they seem fit while also trying to think which category will be chosen by the other participants. For round 2, independently of the size of the paragraph and since the players had earlier read and analyse the text, this round timer is of 30 seconds.

After the time expires, we move again to round 1 if there are remaining paragraphs to be analysed or otherwise we move to the final round.

3.2.3 Round 3: Review and decide

After round one and two finish for all paragraphs, the challenge reaches the review deciding round.

The interface changes, the screen contains the complete fragment at the bottom part of the screen while in the top part there is a voting area containing the categories that won each voting. Additionally, we have a timer, and the indication that the game is iis final round.

This round is the final crowdvoting process, users can read if necessary the fragment as whole and the players must vote on one of the tags in order to elect the top tag of the fragment, which will be the final one and the one that gets to to be included in the virtual edition. However this stage

is more dynamic and iterative due to the fact that users now can switch their vote until the time expires whereas in the earlier stage they could only vote once on one tag.

Accordingly to that they can also see the points of each tag and the current top tag in real time. Nonetheless the rule maintains that they can only vote in one category, which will be the option they leave selected when the time expires.

This presents the challenge of trying to simultaneously chose the best available tag whilst trying to predict the winner tag. between categories. The score of a participant follows the formula presented bellow:

$$Score = s + s_{RWT} + v_{RWT} + s_{GWT} + v_{GWT} + c,$$
$$where:$$

s : submit a tag = 1 point;

s_{RWT} : submit a **round** winning tag = 5 points;

v_{RWT} : vote on a **round** winning tag = 2 points;

s_{GWT} : submit the **game** winning tag = 10 points;

v_{GWT} : vote on the **game** winning tag = 5 points;

c : for each vote change = **minus** 1 point;

(1)

From this we can see that, submitting a classification earlier is key (it allows to be considered the author of a tag); it is crucial to avoid making unnecessary vote changes and of course trying to find a tag that fits while trying to think in agreement with the other participants, resulting in the *wisdom of crowds*.

The player that suggested the winning tag will have their name credited as the author.

4. Implementation

The implementation can be split in two parts server-side and client-side. Each is presented in the following sections.

4.1. Server-side implementation

The server side was built upon the existing architecture of the *LdoD Archive*. The *Archive* is based on Java, the Fénix Framework[4], the Spring Framework[5] and JSP.

As expected, Java is used for the business logic while the JSP is used for the presentation logic of the *Archive*.

In order to have a persistent domain model the *Archive* uses the Fénix Framework, which is an open-source ORM developed at IST. The framework uses it's own language DML and overall the use of this framework simplifies the support of transactional behavior and persistence.
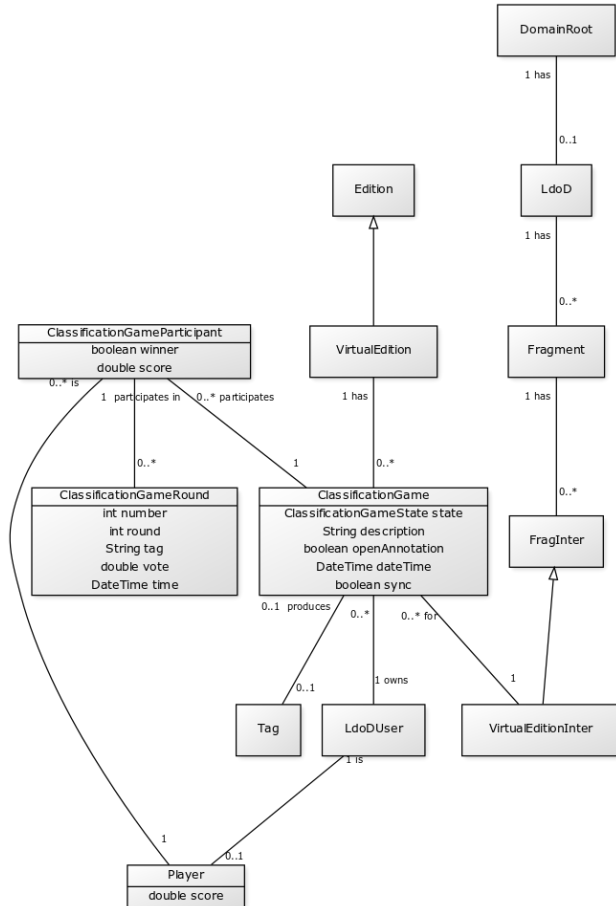
Figure 2: Enriched UML Class Diagram of the *LdoD Archive* with Classification Game entities

On the other hand, the *Archive* uses the Spring framework, which provides many other modules for Java based application that also make everything easier, such as Spring MVC.

4.1.1 Domain Model

The new domain model is simply an extension of the existing *LdoD Archive* domain model. This model is presented as a UML Class Diagram in figure 2, notice that this diagram is simplified as a way to make it easier to understand and follow, showing only the classes that matter for the game interaction and without their attributes.

As we can observe from the figure, the new domain model has four new entities: ClassificationGame, ClassificationGameParticipant, ClassificationGameRound and Player. Player is a very simple entity that only holds *score* as an attribute and has two relations one with a LdoDUser and another with the ClassificationGameParticipant. The Player semantic is evident: an LdoDUser can be a Player or not (0..1) and a Player can only be associated with an LdoDUser; also a Player can be a Participant in many games or none at all and the

inverse is also true. This allows to create a game without initially defining who plays and giving the flexibility to allow for adding more players later, but only if the game hasn't started yet.

On the other hand, ClassificationGame is a bit more complex having mainly a *description*, a *datetime* corresponding to the moment in which a game should be deployed, a *sync* variable that signalizes if the game must be synced and finally has a state machine with the following states **CREATED, OPEN, STARTED, TAGGING, VOTING, REVIEWING, ABORTED, FINISHED**. The state machine and the sync variable are the key concepts for the synchronization (explained in section 4.1.4). Note that any ClassificationGame complies with the rules of the Virtual Edition that is associated with (such as public/private editions condition the users that can play the game).

Lastly, the ClassificationGameRound allows to keep a history of interactions between the participant and the game, having information like in which rounds the user played (attribute *round*), what did he submit; vote; at what time and in which paragraph (identified by the variable *number*).

4.1.2 REST API

This section describes the RESTful API that was developed to support the communication between the server (the *LdoD Archive*) and our client (the Classification Game). The application developed uses aforementioned Spring Framework, which allows the definition of one or more Java classes to act as Controllers and handle requests.

Defining a controller in Spring is easy as annotating it with the @Controller annotation, or in our case @RestController. Four main controllers are of interest to us: the Authentication Controller; the API User Controller; the API VirtualEdition Controller and the ClassificationGameController.

The AuthenticationController is only responsible to authenticate users from outside applications that connect to the server. It is the only endpoint that is truly open to everyone since it is not covered be a JWT filter that checks the token validity.

The APIUserController as the name suggests, returns a DTO of an user (with information such as username, virtual editions, games).

The VirtualEditionController deals with requests for fragments and virtual editions indexes.

The Classification Game controller is explained in the following section.

4.1.3 Classification Game Controller

This controller it is perhaps the most important from all of the above, it manages all the game interactions in two different ways: REST requests and real-time communications using STOMP over WebSocket[6] while also saving data to the database.

Regarding the Restful requests we have three methods:

- ActiveGames: Returns a list of active games for the specified user, if the requesting entity is the user. The result is a list of games that did not yet started that are available to everyone and games in which the user is a editor of the virtual edition associated with the game;

- End: Method responsible to terminate a game, saving it to the database and return the game winner information;

- Leaderboard: Returns a leaderboard to the client.

Concerning STOMP, it has identical characteristics to HTTP. Uses TCP along with providing a set of commands, yet we will focus on the most relevant to the scope of this thesis which are: CONNECT/DISCONNECT; SUBSCRIBE/UNSUBSCRIBE and SEND. The choice to use STOMP is by virtue of Spring supporting it. The usage is very simple, one must simply define the appropriate websocket URL and define which topics are available for client subscription. After that a broker is instantiated that can send to the subscribed users of a topic.

Regarding the server STOMP related handler methods, the communications is very straightforward the clients subscribe to topics such as *connect*, *tag*, *vote*, *review*. After that clients any relevant data (for *tag* the word that was submitted; for *vote* what was the word that was voted) and every communication includes an *userId* and a *gameId* which ensures that clients and server communicate only with the specific users of a game and not the whole universe of other players. All these methods mentioned, do a transaction to save into the database data such as tags; votes; rounds and scores.

4.1.4 Synchronization

The main concept for the synchronization of the players and the game is to make the game clients alert the server each time they have completed something. From there, the server wait some time for the clients and afterwards informs them to proceed the game. In the time, that the server is waiting the clients stop and do not nothing while holding for a command to continue.

---

[6]https://stomp.github.io/stomp-specification-1.2.html.
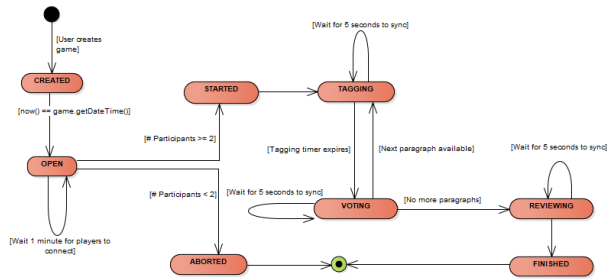
Figure 3: State machine that illustrates the synchronization flow of a game.

The synchronization is done by instances of the Game Runner class. The server has a scheduler that executes a every minute a method that finds which games are available creates a new thread of Game Runner for each available game. The *getGames* method simply reads from the database which games are in a **CREATED** state and that are scheduled to start in the next two minutes.

The Game Runner class acts as a decider for the game synchronization steps. Firstly, if the games has entered the *OPEN* state, then keeps checks the current time and the game starting time and when it has reached one minute after the starting time it starts (this one minute windows allows users to connect). After this, the GameRunner checks if the game can start (i.e, the number of users that connected is at least 2). If the game proceeds, the GameRunner is in a loop until the game has not finished, and periodically checks if the game must be synced (this is triggered by the ClassificationGameController that it puts the sync variable to true when the clients request it). Then the GameRunner, waits some time for the users to reach this stage and after that it sends to the clients a command to continue the game.

Figure 3 summarizes the synchronization process.

4.2. Client-side implementation

The client for the game was built using ReactJS which is an open-source JS framework maintained by Facebook and it is usually used to develop UIs.

The framework main concept revolves around Components, which can be compared as Java classes. The game has four main components that focus on getting data from the server and deciding what to do next: App, Game, Virtual Edition and Fragment. While we also have four other components mostly focused on game mechanics and presentation: Paragraph, Tag, Vote, Review.

Our main component is the App component responsible for authenticating a user; fetching active games for the logged in user and registering the paths for the other components.

Afterwards, we have the Game component that

is responsible for connecting and registering users to a game; deciding if the game starts or aborts accordingly to the command sent by the server and rendering the component Virtual Edition. A Virtual Edition component fetches the fragment to use and passes it to the Fragment component, which in turn splits the fragment text into paragraphs and passes it along to the Paragraph component.

A Paragraph can then render different components according to each stage of the game: Tag (Round 1), Vote (Round 2) and Review which renders a Vote because reviewing also includes voting (Round 3).

Each of these last three mentioned components communicate with the server through the use of websockets

These components are very similar in their behaviour and communication. Firstly, as aforementioned the Paragraph is the component which determines everything about these components: the Paragraph holds the text for a paragraph of the Fragment selected to use in the game and the life cycle of a Paragraph is determined by the time attributed.

Longer paragraphs correspond to more time to read; submit a tag and vote. So for each Paragraph for the specified time: renders a Tag (which simply allows for text input; saves data submitted and transmits to the server); following this a Paragraph requests a server sync; moves to render a Vote component (similarly allows for voting input; save vote date and relays to the server), this is repeated for all paragraphs until when we reach the final paragraph after which we render a Review component that receives the fragment text, requests the server which tags are to be voted and finally renders a Vote component which acts similarly as before but now allows vote changes. Subsequently, when the time expires the Paragraph component uses a callback to alert the Virtual Edition component that the games has ended; which in turn alerts the server and the server replies with the game winner.

To note, the Virtual Edition component could be removed and the Game would communicate directly with the Fragment because games at the moment are only for one fragment. However, this allows for a easier way to extend the game. In another iteration of the game if we would want to have multiple fragments to be used it would be a simple as making the Virtual Edition fetch them all and control the render of each Fragment in turns.

5. Results & discussion

The accomplishment of any product frequently relies on assessment of the general population that will use it. In this case, if a game is alluring, enter-

taining and addictive it will probably be exceedingly utilized.

The developed game endeavors to make people collaborate with each other in a fun manner in order to produce new artifacts (Virtual Editions with new tags) whilst promoting knowledge about a literary work by Fernando Pessoa.

As the fundamental objective of the created application is to be a text based game, it was asked to colleagues; family members; existing users of the *LdoD Archive* to test the prototype, in order to better understand its positives and also negatives.

After testing the game prototype, participants were requested to fill a little questionnaire that assess and profiles the user. Part one of the questionnaire focuses on group age; gender; acquaintance with the work of Fernando Pessoa and gaming habits. While the second part targeted characterization of the game itself such as its ease of understanding, rules and entertainment.

The questionnaire had sixteen responses. Thus, every interpretation we take from this must be weighted and cannot be extrapolated without having more responses.

Regarding the profiling of the inquired we can say that nine of the survey respondents are men and seven women; the majority of the participants have an age between 19 and 24 and all of the inquired have a higher education degree (the majority, 10 a graduation). When questioned about knowing about the *Book of Disquiet* only two said that they did not had heard of its existence. Furthermore half of the people had read the book and also half selected *I am a normal reader of Fernando Person knowing only mandatory works* as the option that best described them. Five considered themselves as an above average reader and three as experts. Finally, only one person never plays games and a majority (five people) play games daily.

Analysing this section of the questionnaire, firstly its obvious that it is a very small sample size. The majority of users are very young; know some thing about Fernando Pessoa and regularly play games.

The next section, focused on finding about the game experience the players had. In a scale from 1 to 5 (*very easy, easy, normal, hard, very hard*) users had to rate the game regarding the difficulty of understanding it (objectives and rules). The results were: seven choose rating *very easy*, two choose rating *easy*, three *normal*, three *hard* and one found it *very hard*. Regarding the entertainment that the game provided to the participants, from 1 - *not fun* to 5 - *very fun*, two people found it *not fun* while a majority (8) found it *fun* and six other found it *normal*.

The majority (eleven people) would consider playing again and would recommended the game to another person. Two people said they would not play again and five responded as *maybe* regarding recommending the game. The question if the game made it interesting to read and learn more about the *Book of Disquiet* and *Fernando Pessoa*, was very divisive with seven people saying *yes, a lot*, other four *not at all* and the rest in between with *a bit*.

So from this, considering this small sample size we can say that, for now, the game seems to be easy to follow; it has not reached a high level of entertainment; it has the potential to entice players to play again and could also attract other users to join. On the other hand, it does not seem to push and promote the interest on Fernando Pessoa and the book.

After this, the questions were of open format and not mandatory. The questions were what did the user like the most and the least; what did the user think it should be improved and any other suggestions.

The most relevant points brought up for each of these questions were:

- most liked: compare what other participants wrote; originality; reading and knowing fragments of the book in a fun manner;

- least liked: time pressure;

- improvements: interface related changes (text size, formatting and UI in general); weekly rankings, increase time, in the end showing what tags other users in other games had chosen for that fragment;

- general suggestions: ranks and social achievements; increase times; provide suggestions of classifications in round 1; explain better the game (to the general audience and how a game must be created and its rules).

In conclusion, noting the small sample size this questionnaire confirms our need to keep having feedback of our players in order to improve. The game as is has some potential, but it must follow the user base suggestions in order to reach its goals which at the moment it does not.

6. Conclusions & Future Work
6.1. Achievements

The accomplishment of any social oriented tool is of course very dependent of its user base and their interactions with one another. As such the platform success will be a constantly evolving accordingly to the users suggestions and improvements.

However, there are already some visible results. Firstly, a foundation for an API, that includes adequate JWT login; RESTful methods and Websockets communications was developed and can be

now extended to anything that it may seem relevant and interesting in the future. Additionally, this thesis allowed to further expand the existing features of the *LdoD Archive* and integrate them in a new environment.

Beyond this, our prototype allowed and paved way to integrate more outside environments with the current platform, since the new *LdoD Archive* now serves two interfaces one from JSP and other from the React game.

Regarding the game, despite that fact that the number of users to test the game at the time of the writing of this thesis is undoubtedly low, we can already take some interesting points. The game has some solid synchronization without having any reported incidents. The majority of the people found it the game very easy to follow; a majority voted as it entertaining and considered that the game sparked some interest regarding the *Book of Disquiet*.

6.2. Future Work

The future work on this solution should be based mostly on user feedback focusing on many different aspects in order to achieve the best possible engagement. Such as: point distribution; changes to the current game flow; fine tune the available time for each paragraph in order to reach an optimal time; introduce thresholds and taboo words for saturated words; UI modifications for having a modern game like experience; although the game is responsive for mobile it could also be expanded to a mobile app which is a platform very appealing for quick games; new game modes and mainly include community suggestions that seem to be in the interest of a large majority of the playing user base.

An interesting work for the future would be to instead of having only one tag by fragment, one for each paragraph. That is, the paragraph-by-paragraph analysis that the players do identifies important semantic information that is then lost if the final classification is built only by a tag for the whole snippet. The current game version is mainly accurate only for shorter texts that are well synthesized with one label only. While fragments that are longer than one paragraph would be better featured with multiple labels. From the point of view of the score of the players one could create a weighting factor that would add up the points of those players whose labels were chosen - which would allow comparing a player whose label received the most votes, with another player with two labels voted, but only one of them exceeded the threshold to be considered relevant. When there were identical scores for different labels, both would be adopted. This could also reinforce the collaborative aspect of the game by allowing multiple *winners*.

On the other hand, the *Archive* website could have an reformed UI that revamp existing elements and introduce other such as profiles; leaderboards; social badges and groups; and other elements to make a bridge between the game and the website.

Furthermore, the game should be internationalized, as it is a very important factor, making it a primary step to be be considered because it enables to reach a more broad crowd.

In order to keep track of the project evolution, the source code is publicly available in a GitHub repository[7] and the game can be accessed through the link present in the webpage[8] or directly at the game homepage[9].

References

[1] D. M. Berry. Introduction: Understanding the Digital Humanities. In *Understanding digital humanities*, pages 1–20. Springer, 2012.

[2] S. Deterding, D. Dixon, R. Khaled, and L. Nacke. From game design elements to gamefulness. *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11*, page 9, 2011.

[3] E. Estellés-Arolas and F. González-Ladrón-de Guevara. Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2):189–200, 2012.

[4] F. Groh. Gamification: State of the Art Definition and Utilization. *Research Trends in Media Informatics*, pages 39–46, 2012.

[5] N. K. Hayles. How we think: Transforming power and digital technologies. In *Understanding digital humanities*, pages 42–66. Springer, 2012.

[6] J. Howe. Crowdsourcing : A Definition. (June), 2006.

[7] J. Howe. The Rise of Crowdsourcing. *Wired Magazine*, 14(06):1–5, 2006.

[8] T. Presner. Digital Humanities 2.0: A Report on Knowledge. pages 1–12, 2010.

[9] L. von Ahn and L. Dabbish. Labeling images with a computer game. *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*, pages 319–326, 2004.

---

[7] https://github.com/socialsoftware/edition
[8] https://ldod.uc.pt/virtualeditions
[9] https://ldod.uc.pt/classification-game