****

**MSC Coverpage Assignment**

|  |  |
| --- | --- |
| **Name of Supervisor:** | **Roy Erkens** |

|  |  |
| --- | --- |
| Course code: | **BTR3000** |
| Course Title: | **BTR** |
| Internal Advisor: | **Chris Pawley** |
|  |  |
| Student Name: | **Molly Brown** |
| Student ID nr.: | **I6128745** |
| Academic Year: | **2019-20** |
| Number of Words: | **9860** |
| Title Assignment: | **Computer application teaching plant identification: development of a matrix key and memorisation game.** |
|  |
|  |  |

Date: Signature:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Computer application teaching plant identification: development of a matrix key and memorisation game.**

**Abstract**

This paper demonstrates a new approach combining matrix keys and novel educational game to improve learning within botany. The introduction discusses the need for plant identification skills by non botanists, the use of technology in botanical education, and game design. The literature review compares different approaches in these fields. The methodology describes a novel approach: to create a matrix key and to develop a game based on the principles of method of loci. The results present the software, which is designed to make accurate plant identification quick and fun. The layout of the game was developed specifically using a familiar generic software layout, and is intended primarily for undergraduate students. The discussion describes the utility of the program, how it could be further developed, and alternative uses of this thesis research.

**Introduction**

Plant Identification had been an area of study since antiquity. Texts about Ayurvedic medicine describing medicinal herbs, date from 7th century BC), and the Greek philosopher Theophrastus (371-287BC) wrote ‘Historia Plantarum’ (‘Enquiry into plants’) which described many aspects of plants, such as their cultivation and medicinal uses (Mark, 2019; "Theophrastus (371 - 287 BC)," 2011).

A detailed knowledge of plants has always been necessary for human survival, as they are a major source of food, medicine, and materials. In modern history, technological advancements such as the agricultural revolution (1700’s) has meant that the need for plant identification has lessened ("Agricultural revolution," 2015). This, as well as the urbanisation and development of cities (replacing nature with infrastructure) has meant that people are less likely to interact with nature.

Today, there can be a high entry barrier for people interested in understanding botany. Developments in botanical research have created a large field of study, which require knowledge of complex scientific terminology to understand its literature. Although descriptions of this terminology are widespread (within textbooks and scientific dictionaries), accessing and understanding in these formats is often considered an arduous task for students (Jacquemart, Lhoir, Binard, & Descamps, 2016). Students today, and increasingly in the future, have been brought up using computers and regard the internet as their primary source of information (Hebert, 2018). Educational computer games are being increasingly adopted within pre university schooling (Hebert, 2018). The latent adoption by universities of this method, can be attributed to the vast range of subject areas and the overall complexity of the information.

Ultimately, utilising computer database creation and gamification provides students with more interesting and engaging ways to access, interact and learn new botanical information.

Plant identification matrixes

Taxonomic data tends to be displayed as either di/polychotomous keys or matrix keys. Dichotomous keys are more common, due to ease of use in print, minimising paper requirement. With increased accessibility to computers, computer matrix keys have become a viable alternative. Winston focuses on dichotomous keys as matrix keys were largely underdeveloped (Winston, 1999). Despite this, many stated design rules for keys were immutable. Relevant design considerations include a list of potential flaws of keys: bad or missing information, information that doesn’t cover all possible species, the exclusion of information on hybrids, and attempting identification with fragmented specimen. This tends to result in two possible outcomes, misidentification, or lack of identification (Hagedorn, 2007; Winston, 1999).

The advent of DELTA (1980) is the first widespread computerised plant identification key (Winston, 1999). It is a tool intended for professional botanists to aggregate a global dataset of taxonomic information, and be able to read and query it. It was once widely used. This is evidenced by the program containing many datasets, and its utilisation is still available to this day. However, many programs superseded this, offering newer and more technologically advanced programs. The criticisms of these new programs were that they could not encapsulate all of the professional scientists and other interested parties, resulting in largely segmented datasets rather than the global dataset envisioned. The issue is a common problem with social media platforms, it is only successful if a large enough body of people invest their time into it. Presently, Biodiversity Information Standards (TDWG) is attempting to combat this segmentation by determining industry standards for software to use ("Biodiversity Information Standards (TDWG)," 2019; Hagedorn, 2007).

Object Oriented Database Model

Databases allow the storage of data for utilisation within a computer program. Relational databases use tables and keys to store information. With this you are able to read, write, update and delete data, often using SQL (Structured Query Language) statements. Matrix keys are a form of relational database. Object oriented Database Model (OODBM) utilises both principles of a relational database and object oriented programming (OOP) (Zaiane, 1998).

Principles of OOP are described as objects and classes (Zaiane, 1998). They act like the hierarchical structure present in phylogeny and taxonomy, where plant kingdoms, orders, families, and geneses are all classes which are defined by their characteristics. Plant species are considered as objects, each of which is a part of a class (genus). Within OOP everything is an object, so you can also consider plant families as objects (as well as a class) of a plant order. Within OOP there are 4 main concepts: Inheritance, Abstraction, Encapsulation and Polymorphism. These concepts are useful when describing plant taxonomic structure.

* Inheritance: A class can inherit aspects defined within another class.
* Abstraction: Hiding information for the end user.
* Encapsulation: Bind properties with functions, it allows the creation of private variables.
* Polymorphism: methods can have different impacts

Educational approach

Educational database tools do not need to be large and complex. Current botanical teaching methods often incorporate the flora of their respective local areas (Attigala, De Silva, & Clark, 2016; Erkens, 2019; Silva, Pinho, Lopes, Nogueira, & Silveira, 2011; Thomas & Fellowes, 2017). Similar to the time of the ancient Indians and Greeks (Mark, 2019; "Theophrastus (371 - 287 BC)," 2011), linking plant identification to both their surroundings and to their uses, provides important contextual information for students (Silva et al., 2011). Moreover, plant identification requires specimens, and local flora is in abundance.

Understanding the hierarchical structure of plants (taxonomy) and biota is a skillset taught to students, and the ability to describe the divisions of more general categories (kingdom, phylum, order, families) creates a transferable skill for students when confronted with other geographical areas. Therefore, creating an educational database which restricts the geographical location and the specificity of identifiable characteristics, allows the creation of a smaller, more user friendly tool.

Game development

Game development can take a variety of approaches. This thesis uses ("Extra Credits: Games in Education," 2013) as its source, to describe the components necessary in game development. Some principles of game development are listed below, and these are further explored in this thesis.

* Simplicity: There is a general recommendation not to attempt to combine too many genres and styles or unnecessary features.
* Meaningful decisions: Players need to be stakeholders in the outcome of the game. Player choices should have an impact throughout the game so that there is ongoing engagement.
* Rules: These should be clear and simple.
* Variable and quantifiable outcomes: The game outcome should have value to the player
* Negotiable outcomes: the real life impact of the game can be influenced by the user.

One approach is to make a simple game, as they are quick to learn and reduce the user entry barrier. It also helps to combat the issue of game literacy, by using concepts familiar to the game users before they even play the game. Ideas for pre-familiarisation might include certain controls performing the same action in multiple games or commands being displayed in a familiar areas of the screen. Examples include menus being placed in the top left corner, which is commonly seen in many software packages.

Another aspect of a game, is the inclusion of meaningful decisions. This engages users during game play. Successful education applications, such as Duolingo, uses counters to achieve this ("Learn a language for free. Forever.," 2019). Repeated use of the application adds to a counter and a 'streak' to be created and maintained. This results in gaining more points to update an avatar, and a place on a leaderboard. This can be meaningful in the real world, as you can add people you know to your leaderboard. The use of counters is often used in gamification, however it is not always successful ("Extra Credits: Games in Education," 2013; Khan, Ahmad, & Malik, 2017). If it compromises the simplicity of the game, or does not have a 'meaningful' effect, it can be unnecessary and detrimental.

Deciding the outcome of the game is also important. Outcomes need to be variable, quantifiable, valorised and negotiable (Burgun, 2012; Juul, 2018). The first three can act as motivational tools. This is because by having a worse or losing outcome, players can be encouraged to play again to achieve a better result. This is an important aspect of gamification, the concept of low stakes (negotiable outcomes) and a high repetition of gameplay. The process of repetition is a common learning tool, being most beneficial when the same information is presented in different formats (Khan et al., 2017; Kiesler, Kraut, Koedinger, Aleven, & Mclaren, 2011; Pinola, 2019). Again, the concept of meaningful decisions creates a meaningful outcome for players, helping to encourage this repetition strategy (Juul, 2018). The concept of low stakes needs to be balanced with the amount of player effort necessary for the game. If there is no player effort required, users will not be engaged. However, conversely, too much effort will dissuade users from repetition (Juul, 2018).

Memorisation is an important aspect of learning. Games are effective in facilitating memorisation by the use of repetition, but the process in which information is repeated can vary widely (Pinola, 2019). For the purposes of education, conceptual learning is preferred over rote learning (Pinola, 2019). By this, students can recall not only facts, but also related facts or concepts. Plant family information can be taught using both methods. Their names use a largely unfamiliar list of vocabulary, not too dissimilar to learning nouns in a foreign language, which often includes rote learning. But their descriptions can be placed in a context of plants students might already be familiar with, or described within the taxonomical hierarchy.

**Literature Review**

Providing a comprehensive review is not in the remit of this paper. Rather, it is selecting a few distinct case studies and analysing their success in the context of the implementation of my program. The approaches being considered are: (Kissi & Dreesmann, 2018), iNaturalist (Seltzer, 2019b), BioBlitz , (Jacquemart et al., 2016; Silva et al., 2011) and (Leggett & Kirchoff, 2011). These papers were selected to show a variety of modern educational approaches to plant biology which also incorporate of the use of computers, mobiles or other technologies. These papers were also chosen because of their unique methodologies.

Description of the approaches chosen:

The Kissi & Dreesmann (2018) approach uses a tablet mobile application, to educate students between aged 10-16, about plant diversity, endangerment, protection, morphology, ecology and systematics. This application was used during a school excursion to a botanical garden. The ‘hunt’ element of the application is an example of gamification.

iNaturalist teacher’s guide (Seltzer, 2019b) describes how its two mobile applications (iNaturalist and Seek by iNaturalist) can be utilised by teachers. It is aimed at a wide age range, from children to adults. It recommends use of traditional teaching methods (field guides, sketching, and hands-on exploration) in conjunction with the application. They recommend the use of BioBlitz in conjunction with the two applications, and provide multiple examples of successful implementation of iNaturalist in educational settings (Seltzer, 2019b). BioBlitz was developed in conjunction with iNaturalist. It provides a structure in which educational botanical excursions can be conducted, and it encourages scientific curiosity via a series of questions. It promotes citizen science, and the use of data gathered in research applications. It also provides practical information, i.e. logistics, budgeting, etc.

Jacquemart et al. (2016) uses an online dichotomous key of Belgian flora, containing multimedia and hyperlinked texts. It is generated for undergraduate students, but is considered applicable to high school students, as well as amateur botanists. The research derives much of its foundation from Silva et al. (2011), who developed both an online and CD-ROM multimedia application to identify Portuguese plants. Her work was aimed at high school students.

Leggett & Kirchoff (2011) provides a literature review of plant identification applications. The paper shows a diverse range of approaches, however many of the available applications have been developed by only a few companies. Their work has an emphasis on the use of images within these applications, with the additional papers (Baskauf & Kirchoff, 2008; Kirchoff, Remington, Fu, & Sadri, 2008), also discusses the same subject.

These papers give insight into different considerations of computer and mobile use in the field of plant biology. All papers were optimistic in the notion that, used correctly, technology can be beneficial to the learning experience (Jacquemart et al., 2016; Kissi & Dreesmann, 2018; Silva et al., 2011; Thomas & Fellowes, 2017; Leggett & Kirchoff, 2011). Kissi & Dreesmann (2018) associates reduced anxiety with the use of technology, with students having existing familiarity with computers. iNaturalist (Seltzer, 2019b) is a commercial mobile application, so it is likely to have a biased opinion of it’s positive utilisation. However, within its teachers guide and with the use of BioBlitz, it still reiterates the need for traditional teaching methods, and considers the use of the application as supplementary. Jacquemart et al. (2016) uses student feedback to illustrate the effectiveness of their tool, specifically, and states that the use of multimedia tools has been proven to be effective (Silva et al., 2011). The conclusion from these texts suggest that the discussion ‘is technology useful in education’ is now superseded by ‘what is the best way to apply technology in education’. There tends to be a lack of consensus of the ‘best way’ to apply technology, and it is likely that each of the approaches discussed are optimal to different subsets of students. Research directly comparing these approaches, and perhaps the approach taken within this paper, would be a useful resource for educators in this field. The closest literature review is (Leggett & Kirchoff, 2011). However, without the involvement of participants, it is harder to compare, when the users and the use case scenario of the applications, are so diverse. Furthermore, these approaches often target a large age range of students, and potentially a more focussed approach may produce better results, rather than trying to develop tools that are accessible to a wide audience.

The use of computers is not always considered beneficial. Thomas & Fellowes (2017) show that there is no significant advantage to using a mobile app for plant identification over an identification book. They also state the limited access to smartphones may present a problem when utilising mobile applications. Only 71% of the participants (undergraduate students at Reading University, UK, taking an ecology course) owned a smartphone, meaning that if mobiles were used in a formal education setting, appropriate devices would need to be made available to students. The number of students with smartphones is high (90% in UK (Thomas & Fellowes, 2017)) however, the prevalence of smartphones is unlikely to be as high globally.

The papers also encourage excursions as part of plant biology education (Jacquemart et al., 2016; Kissi & Dreesmann, 2018; Seltzer, 2019b; Silva et al., 2011). This approach is not a modern one, and has been used since the origins of formal plant identification ("Theophrastus (371 - 287 BC)," 2011). However, this approach has been ratified in scientific literature and study. In addition to, or possibly in lieu of, this approach and the use of images and multimedia (photography, illustrations, videos…) is advised by (Jacquemart et al., 2016; Silva et al., 2011), (Baskauf & Kirchoff, 2008; Kirchoff et al., 2008; Leggett & Kirchoff, 2011).

Jacquemart et al. (2016) proposes that the use of multimedia is integral to her website’s design. It is based on the evidence provided in (Silva et al., 2011), who created a comparable online multimedia dichotomous key tool. They consider the use of images as beneficial, when being implemented within illustrative and descriptive glossaries, but they also state updating and maintaining a visual library, as an additional benefit of an online medium. This could be considered a requirement, and thus a burden to the longevity of the tools application. Applying this notion to the work of (Jacquemart et al., 2016), the success of this illustrated multimedia approach is somewhat limited. The paper discusses the expansion of the website into English and Dutch. However, there has been no evidence of this taking place and for an accurate comparison of it to other techniques, a common language between the papers would be very beneficial. This, more broadly speaking, is one of the limitations of development of educational techniques, particularly in pre-university education. Tools are developed in local language, as these are the languages used throughout the rest of the curriculum. Writing educational papers in English (as is the case with (Jacquemart et al., 2016; Kissi & Dreesmann, 2018; Silva et al., 2011) ) does help to mitigate this, and perhaps if one approach was clearly preferable to the others, time and resources can then be allocated to the translation of these programs into multiple languages.

Leggett & Kirchoff (2011) considers correct use of imagery an integral part to plant identification guides, so much so, that he has developed a style guide to plant photography for mobile applications. He states that that there are two benefits. First, that it is more useful for novice users who are unfamiliar with technical terminology, and second, that it is more useful for experienced botanists who are trained to identify plants by sight. A contrasting point to Kirchoff’s (Baskauf & Kirchoff, 2008; Kirchoff et al., 2008; Leggett & Kirchoff, 2011) importance of imagery, is that the understanding of technical scientific botanical papers relies on an understanding of botanical terminology (Plitzner et al., 2017). Thus, when an application is placed in an educational setting, the comprehension of botanical terminology can be more important than the use of plant identification tools for a lay user. Kissi & Dreesmann (2018) specifically aimed for an improvement in the understanding of vocabulary, but they were not successful in their results. Instead, their results suggested the use of directly repeating strategies (a technique used in language learning) and a spiral curriculum (a technique where topics are revisited throughout a curriculum) to improve results.

Contextual importance is a noteworthy part of several of the papers. Silva et al. (2011) adds descriptions of plant medicinal uses, and which plants can be safely eaten. This aligns with the historical need for botanical knowledge, but also potentially aligns with a newer approach in medicine where ancient medicinal remedies are tested using scientific rigor, and their active chemicals formally investigated. Another contextual element discussed is the biodiversity and conservation of plants: (Kissi & Dreesmann, 2018; Seltzer, 2019b). This topic, along with climate change, has been increasingly in the news ("International Year of Biodiversity," 2010; "Taking Action for Biodiversity,"). The inclusion of these elements is reported as positive, within the papers discussed.

Looking at modern approaches to education in other fields may provide further insight as to what features make an effective educational tool, specifically those utilising gamification. These games included in this papers analyses are: Reality ends here (Watson, 2012), Pokémon GO ("Pokémon GO," 2019), Duolingo ("Learn a language for free. Forever.," 2019).

Reality ends here is a game developed at University of Sothern California (USC) School of Cinematic Arts (Watson, 2012). It was a voluntary and secret game, with card gameplay, creative projects (e.g. making videos), and provided access to industry people. It was very successful with students, with a high uptake of players. Its aimed to develop the creativity in Film students, as well as make first year students get to know the other subtends within their degree.

Pokémon is mentioned in the report of (Balmford, Clegg, Coulson, & Taylor, 2002) as being a more effective method of educating children about information of (fictitious) species compared with plant identification. It is a brand, mainly a card game and tv show but has expanded into computer games. One of the most recent additions is, Pokémon GO, a mobile application, sending users around the local area collect these species. It was very successful (with 12.7million downloads on the Google Play ("Pokémon GO," 2019) and more on iOS ("Pokémon GO," 2019)), and because of its notoriety it has appeared in pedological writing, discussing its impact on children (Althoff, White, & Horvitz, 2016).

Duolingo ("Learn a language for free. Forever.," 2019) is a language learning website and mobile application, included because of the similarities in language learning and plant identification learning. They use a direct translation method. They have points system and a competition board. They have achievement badges.

(Kissi & Dreesmann, 2018) has been described above. But it has gamification element of the inclusion for a treasure hunt within it.

Reality ends here and Pokémon GO both are used in multiple locations ("Pokémon GO," 2018; Watson, 2012). Pokémon GO relies on GPS to know where the player is and unlock features. Whereas, Reality ends here has people who can pass on information to progress within the game. This location based game is a new feature in gaming. It has benefits, namely the increased interaction with other players compared to traditional computer games. However, this also increases the risk for players to go to potentially dangerous areas, a problem more so apparent in games with a young userbase. Pinola (2019) indicates that exercise improves memorisation, however this is not conclusive. This increased mobility has been encouraged in American paediatrics reports with regards to Pokémon GO.

Reality ends here and my thesis has the same userbase of undergraduate student, although they are within different degrees and country (Watson, 2012). The similarities in user base may mean that the tools employed in the game to persuade and encourage students to play (creative tasks and access to industry contacts) are age appropriate and possibly effective across both fields. Additionally, the fact it was an optional game, with such a large user base, indicates it’s an alluring game design.

Duolingo uses gamification point systems effectively. This is a major concept within gamification (Khan et al., 2017). Using point systems has been recommended in classrooms as a more motivational process than grading (Hebert, 2018). With its success in this application, (the largest language learning application), and its relative ease of implementation it is a good tool to incorporate within a game. Although, according to ("Extra Credits: Games in Education," 2013), point systems need to be intentional, and should not be added to every game if they serve no purpose.

**Aims & Objectives**

The research statement of this thesis is to develop an identification key and game within an educational computer application that can improve undergraduate taxonomic knowledge. It will combine the efforts of traditional teaching with identification keys into an under-utilised matrix style, with a novel game approach. The goals are that students will be able to recall plant families and their characteristics and able to understand language commonly used in plant taxonomy. Also, an aim is that the application functions are user friendly, by creating a comprehensible user interface. Additionally, the application should be enjoyable to use such that it might be used outside a curricular requirement.

**Methodology**

Dataset generation

The generation of a dataset of the plant families in the Netherlands was conducted using two sources: Heukels’ Flora van Nederland (Van der Meijden, 2005) to identify the families present, and Plants of the World encyclopaedia (Christenhusz, Fay, & Chase, 2017) to define the characteristics of the families. It was conducted in this way as Heukels’ provides a comprehensive list of plant families in the Netherlands (Van der Meijden, 2005), while the encyclopaedia had descriptions of plant families globally and thus was able to provide information for descriptions (Christenhusz et al., 2017). This approach could result in characteristics being included that may be present in the family globally but not in the Dutch species. However, this factor would only produce a false positive.

As a learning tool a false positive is not always a bad result, as when utilising plant identification keys, the final determination of which plant should include a confirmation or verification phase (Hagedorn, 2007). The confirmation phase of plant identification is a process in which the specimen is identified using information available for the identified class. This requires other sources of information beyond the application and all characteristics can be assessed, rather than the limited number in identification keys (Hagedorn, 2007). False positives can also be mitigated by the use of data indicating the prevalence of the family in the area. Both of these strategies encourage the use of additional sources of information, which is beneficial as identification keys do not contain all the information available about a plant family, rather it is a trade off between providing enough information to lead to distinct identifications, while still few enough categories not to confuse users.

Character ranking

Character ranking is the process where certain characteristics for the plant families were chosen for inclusion within the matrix key. Christenhusz et al. (2017) included a comprehensive list of all the characteristics of the families. But good identification key design, a minimal number of characteristics are chosen, while still enabling correct identification. This paper developed a selection criterium to determine which characteristics to include:

• Were they common in the top 50 plant list? (A booklet of common Dutch plant families created for students to use during undergraduate field excursions. (Erkens, 2019))

• Are they identifiable easily distinguishable by eye? Rejecting features which are often require using magnifying glasses or optical microscopes to be seen.

• Do they surpass a minimum threshold of characteristics to make the family identifiable? This means that they have a unique list of features described. Also, more than 4 features will be described for each plant family.

The first 25 of the 50 plants (Erkens, 2019) acted as a ‘training dataset’, despite analysis being done by hand, to determine these characters. This approach was chosen as it was simple and time efficient. As its use was within an educational tool, a smaller more simplistic key is more beneficial.

Development Software

The database would need to be object orientated and thus an object orientated programming language, specifically Java, was utilised. It was created within the software editor IntelliJ IDEA. Both java and IntelliJ IDEA are commonly used.

The first stage in database design was creating a method which could access the information store on the Excel sheet. Excel was used as it is a simple tool, one biologist are likely already familiar with. Access shares many similarities with Excel in its use of macros, so it is a tool which could be converted to access for future projects if they require more specialised database software. A generic version of this method was created that could loop through every cell to read the file, format and store the data as to ultimately be used in other methods. This method was modified to access specific sections of the spreadsheet. For use within the java program, the plugin from Apache software was used. This software is commonly used to read Microsoft documents in Java code. This meant that there was good support for it online and can be maintained by many java programmers.

Next was the creation of the General User Interface (GUI). This was done using JavaFX. JavaFX and Java Swing are the two main in built tools for GUI within Java. JavaFX is the more modern version. This means that there is more support, and is more likely to be maintained. Additionally, designs can be more specialised with the tool. In designing the GUI, common application design (with toolbars in the top and left) were considered. Mobile applications use a different layout and thus use in mobile would thus require modification.

Layout is a major part of creating a usable GUI. The approach taken was the utilisation of JavaFX BorderPane within most of the program (Fig 1). This has 5 sections: top, left, middle, right and bottom. This structure is comparable to many other applications. The left was used for ‘menu’ items, often ComboBoxes, with examples below. The right was used for outputs, And the bottom contained a ‘Back to Main’ button throughout.

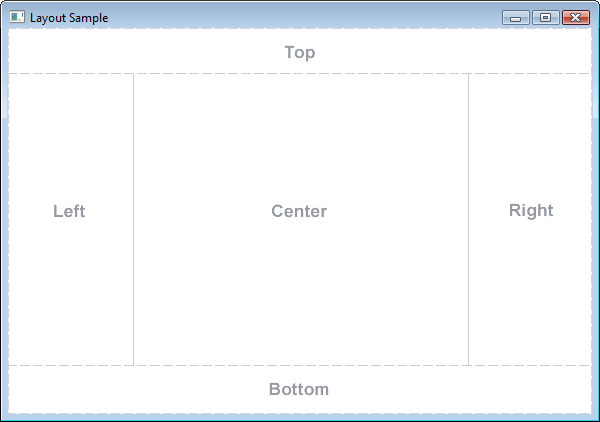


Figure 1. BorderPane layout ("Class BorderPane,")

Database design

The program has four parts: a factsheet, a plant identification program, game point calculator, and a game example. The factsheet displays the information of one plant family, for the data within the matrix as text. The plant Identification program allows users to select characteristics and will then print the what plant family it could be based on the input. The game point calculator and the game example are partial aspects of the game, created for proof of concept. The game will be described in more detail below.

Factsheet

The fact sheet was made as an improvement to reading the matrix within the Excel sheet. One way it did this was removing the absent characteristics, and the other entries when reading the matrix. This more compact presentation is better for displaying information. Secondly it used text instead of the ‘tick box’ format of the matrix. This is also for easier comprehension.

The fact sheet utilised layout guide specified by having the ComboBox with the list of plant families on the left.

Plant identification tool

The plant identification tool was created to allow users to identify plants without need of understanding the matrix, by use of a GUI. Plant identification used the GUI layout by having the characteristic type box on the left, the characteristics checkboxes in the middle, and the output on the right.

The design has checkboxes which can be selected and unselected. The design is as follows:

1. User selects a characteristic
2. This characteristic gets added to an array of characteristics
3. When there is less than 4 characteristics, a message to input more characteristics appears.

The tool required 4 characteristics to be selected to minimise the number of possible results produced, and as character ranking also required a minimum of 4 characteristics this number was chosen.

1. With 4 characteristics are selected, the probability modifier is used, and a list of plant families appears in place of the input more characteristics message.

Another design feature created was a probability modifier. It attempts to improve user experience and likelihood of correct identification. Generation of a modifier was implemented for characteristics by selecting characteristics that users were likely already familiar with, then adding them to an array, and giving the values of the characteristics a weighting of 0.7 and the characters not within the array a weight of 0.5. Weighting is based on subjective knowledge of suspected common features. A list of characteristic descriptions, along with their lookup rate within merriam-webster dictionary (Dictionary, 2002), started to be compiled with aim to develop weighting based upon its values.

Game design

The creative nature of game design does not, in the initial stages, lend itself to the reductive method of selecting an optimal approach from a variety of references. Instead a unique approach is required. Considerations when developing this game an include the genre, as different genres have different benefits.

Educational games need both the benefits of game and the presentation of educational information in a way which students can learn. Repeat exposure to information in a variety of formats is known to increase the chances of student learning the information. Investigating memorisation approaches allowed for the development of the game: Method of Loci game. The other game created is: picture-game.

Game Ideas & Prototyping

Game ideas were generated after considering research into current educational tools and aspects of game designs from the literature review. Games that already have are already commonly used in education such as flip cards, crosswords, fill in the blanks were not used as they did not represent a novel idea, despite proving successful in (Stagg & Donkin, 2013). Ultimately two idea were developed: MSP memory palace and Spot the Difference.

Spot the Difference

Use a ‘standard’ plant illustration containing all the characteristics of the matrix. Select a family and the image will modify to that families characteristics. But one character is wrong… Users attempt to identify the incorrect feature. With some point system which reward identifying the incorrect feature sooner. Attach real images of the family and discuss simplified drawing vs real life.

Ultimately this idea was not chosen. This was because it had seemingly less grounding in literature. Also, many papers discussing the use of images in plant identification and warn against the use of model plants, as they are not representative of what is found in nature. Additionally, it is not the best use the tools developed within the matrix key.

MSP memory palace

A game where you create your own story based on a room you are familiar with. You memorise facts in specific places which then allows you to recall the fact when you think of the room later. The better your story the more points you get. Rate your peer’s stories to give them points. You can also gain points for describing all the families, describing rare families or being the first to describe a family.

But how to make a good story:

* Make it ridiculous, you remember the crazy thing in life not the mundane
* Why say it once when you can say it twice, maybe you thought of two great sentences reinforcing info about your plant family weave them both in or your detailed description.

This game took the memory technique of method of loci, and used gamification to make it a game. MSP memory palace - use photos and develop a memory palace for two of the buildings rooms (teaching users the concept of a memory palace), encourage them to make their own palace (creative memorisation).

This paper compared memorisation techniques and chose to base game design on the method of loci (Yates, 1966):

The method of Loci is a technique where someone memorises the layout of a location (room, building, street) which has numerous discrete loci (points of interest). A list of items can be memorised by ‘walking’ through the location, identifying a discrete loci and thinking of an association between the loci and one of the items. Memory retrieval works by imagining the room and its loci then remembering the association with the item.

Fassbender & Heiden (2006) have used this memory strategy to develop a virtual reality (VR) experience. VR is more of an immersive experience than the MSP memory palace. However, its strategy is very different to the game invented and is much closer to the method of loci description. VR is not a technology available to most schools or students, but the ‘MSP memory palace’ idea could be applied within VR instead of the use of pictures.

MSP memory palace introduces a social aspect to the method of loci by introducing sharing features and a voting system. This was because, in Reality Ends Here (Watson, 2012), they found that a social game was beneficial. It allowed first year students to get to know one another. Additionally, teamwork and collaborative efforts are often important skills both within university and starting careers.

There was also potential with the MSP memory palace to use places across the city. This would often require walking or biking to different locations. There has been some evidence aerobic exercise is linked to better memorisation, so the inclusion of this feature could improve the game’s effectiveness (Pinola, 2019). Pokémon GO showed that people are willing to use location based games, and there have been doctors encouraging this trend ("Pokémon GO," 2018). Further features of Pokémon GO, such as achievements for going to said location are also beneficial. A common game of treasure hunting could be incorporated to make it more enticing to users.

The game is also useful in teaching the strategy of ‘method of loci’, which is a learning skill that could be used for many other subjects.

Developing MSP Memory Palace prototype

Two aspect of the game MSP memory place were developed to see if there was initial interest in the game. One part is the example of a room, with one plant family being assigned to a room and different objects within being described as characteristics of that family. This could be described in video of audio formats, but for the prototype a slide show with text was used. This was because it was accessible for people without cameras and audio devices as the same images can be used for different palaces for different people. The images include within the game is one of the buildings of Maastricht Science Program (Kapoenstraat 2), as test users were likely familiar with this area. Because of this format the BorderPane layout of the other parts of the program were not used. Instead it was just the slide show with a next button to progress through the slides.

The other part of the game that was developed was the game point calculator. This had a TextField so users can put in the specific family they are rating. With this, checkboxes with the characteristics of the family appear so that users have the family information to rate. Other checkboxes are permanently there to encourage some more fun aspects of the game. There include points for rhymes, jokes and collaborating on a memory palace. The number of checked boxes are then counted and displayed. The layout is with the checkboxes on the left, the text field in the middle, the points counter on the top and the back to main on the bottom.

Questionnaire

For the questionnaire a version of the program was released containing 10 plant families. This was because the data within those families needed to be vetted to ensure it was correct. Additionally, it could focus the questions asked, when assessing botanical knowledge. Certain characteristics were deleted and other e.g. sap (which was a characteristic in 8 of the 10 families) could be expanded. These families were chosen using a random number generator. One of the generated numbers was excluded as it produced a value which would have resulted in 4 of the 10 families being in the same order, and thus was more likely to create a bias for people who were familiar with that order. Additionally, the frequency of the family (using data from (Vries, Roeke, & Lemmens, 2018)) was assessed to ensure a variety of common and uncommon families were present (Fig 2).

The questionnaire starts with an assessment of botanical knowledge. This is done firstly by determining whether their study interests are biology based. But to give a more specific botanical assessment, the most common species of two of the ten families in the survey, are portrayed as an image and the user are asked to identify them to the best of their ability.

After this a series of multiple-choice questions are asked. This is to allow the user to provide quick feedback (and making them more inclined to complete the survey). Additionally, it provokes the participants to be aware of different aspects of the program so that they can provide more long form feedback if they feel it necessary.

Figure 2 Graph shows the pictorial distribution of the plant families included in the questionnaire.

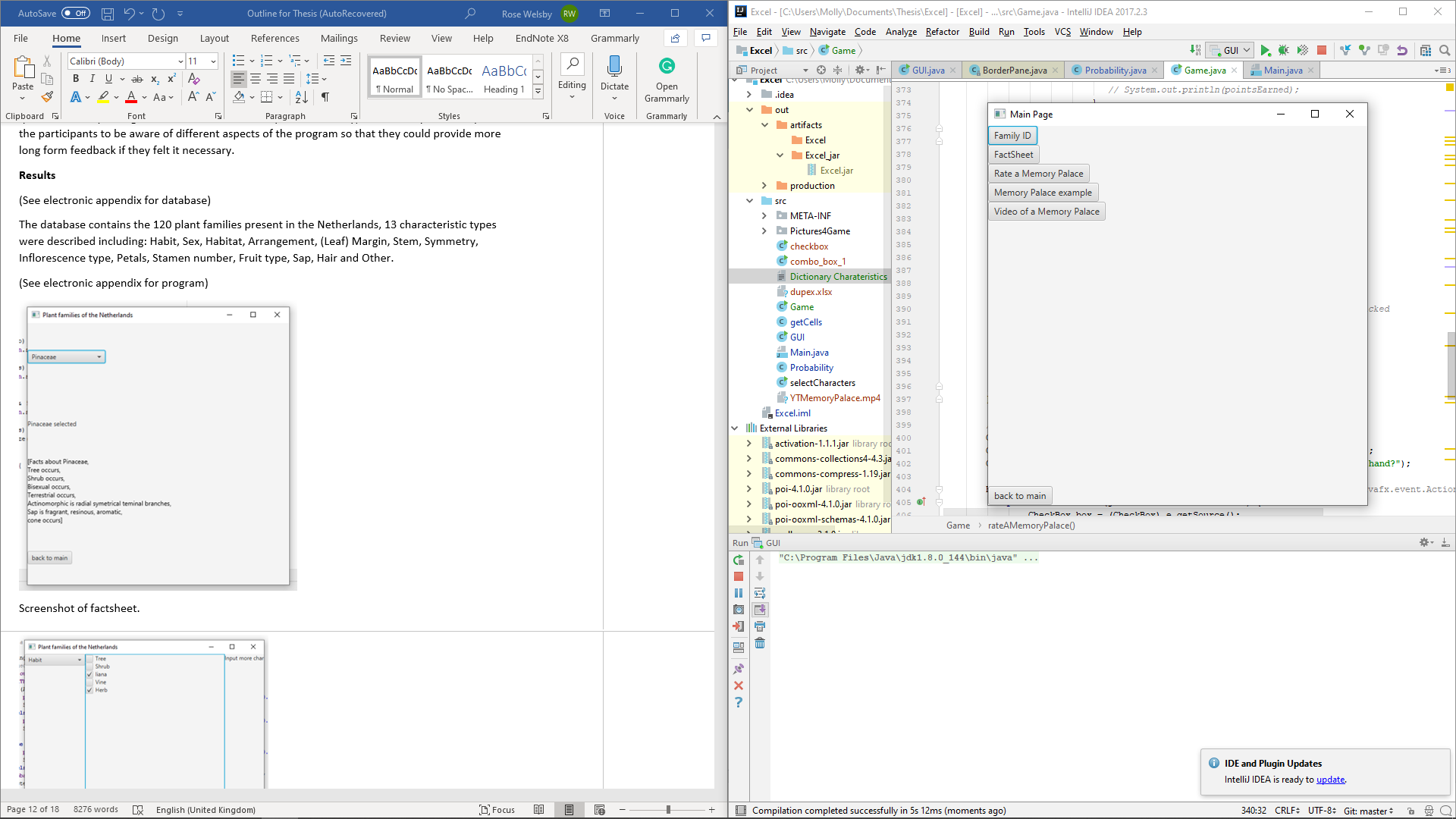
Data obtained from: (Vries et al., 2018)

**Results**

*(See electronic appendix for matrix)*

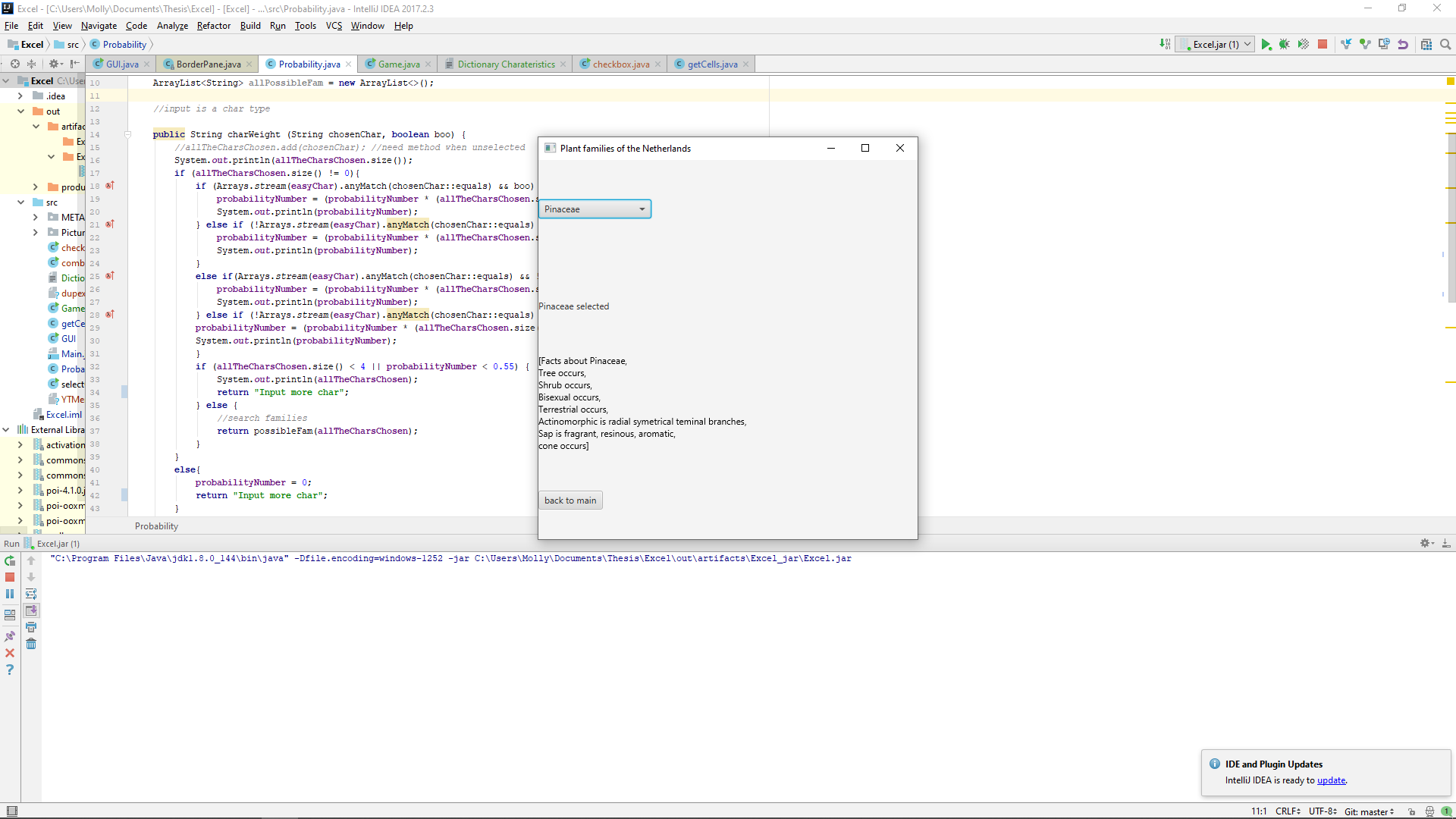
The database contains the 120 plant families present in the Netherlands, 13 characteristic types were described including: Habit, Sex, Habitat, Arrangement, (Leaf) Margin, Stem, Symmetry, Inflorescence type, Petals, Stamen number, Fruit type, Sap, Hair and Other.

*(See electronic appendix for program)*



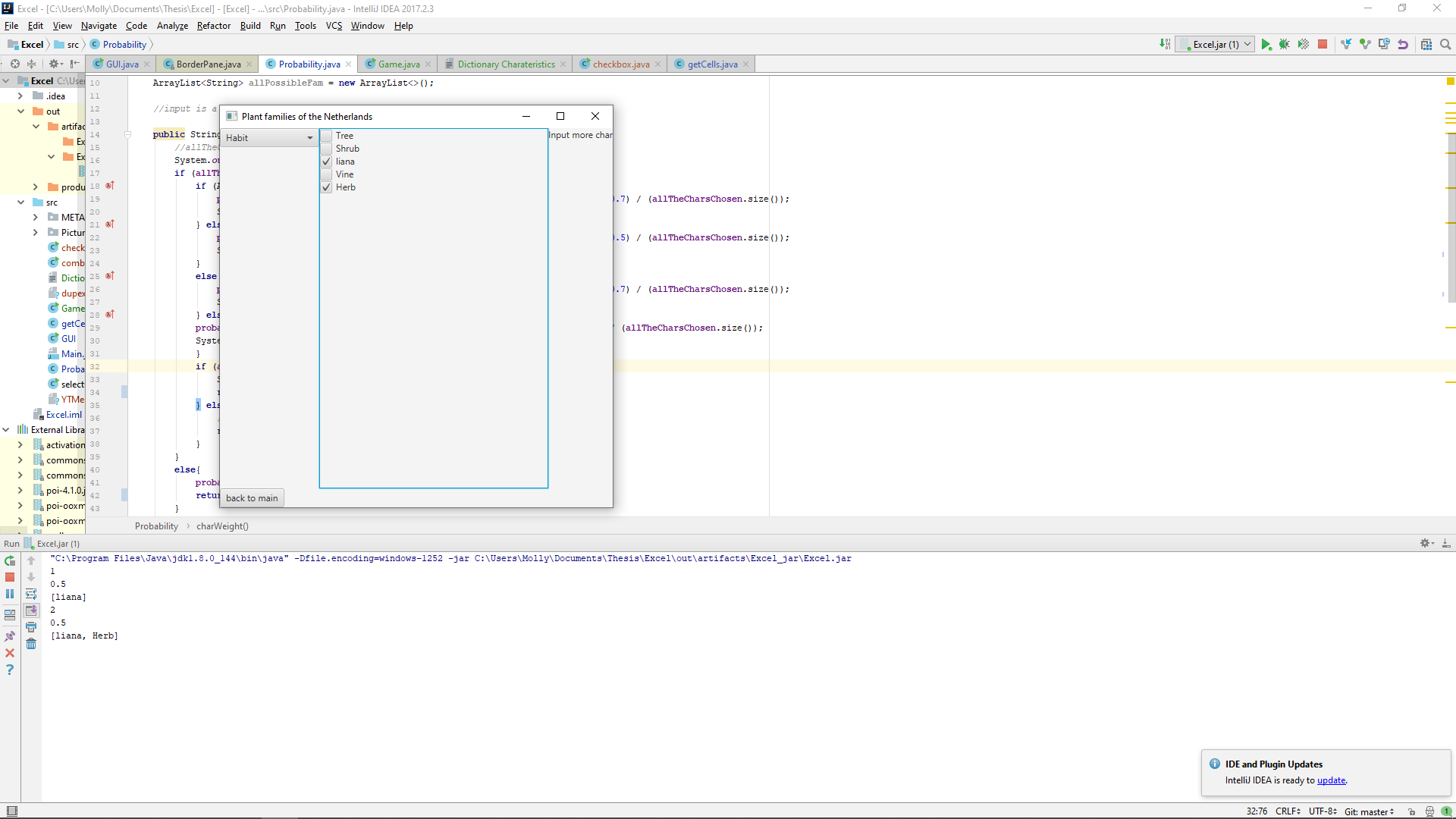
*Figure 3 Screenshot of the main menu*

The computer program was developed such that there was a main menu and 5 sub menus: Family ID, FactSheet, Rate a Memory Palace, Memory Palace example, Video of a Memory Palace.



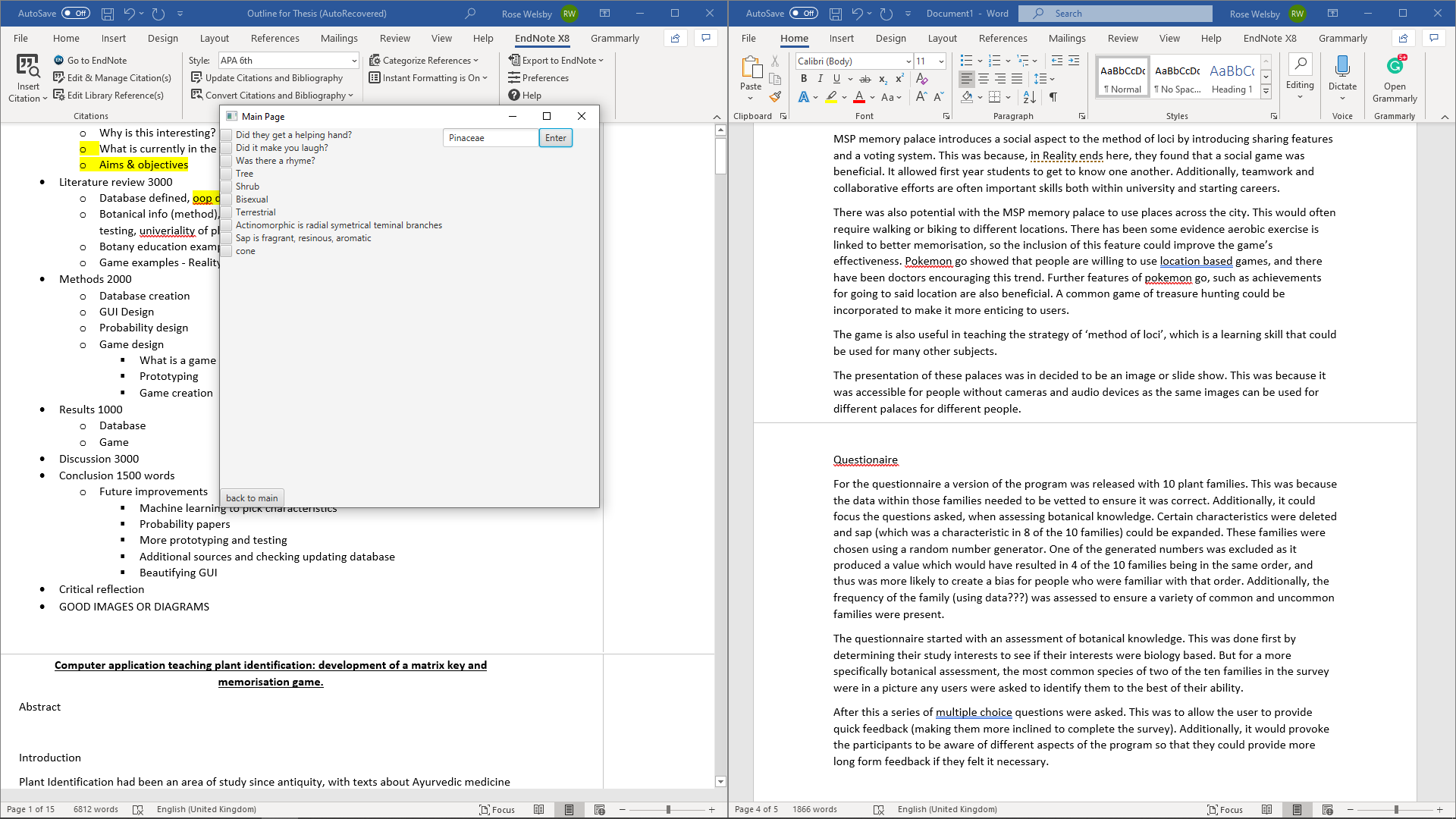
*Figure 4 Screenshot of factsheet.*

Factsheet displays the information from one family of the spreadsheet.



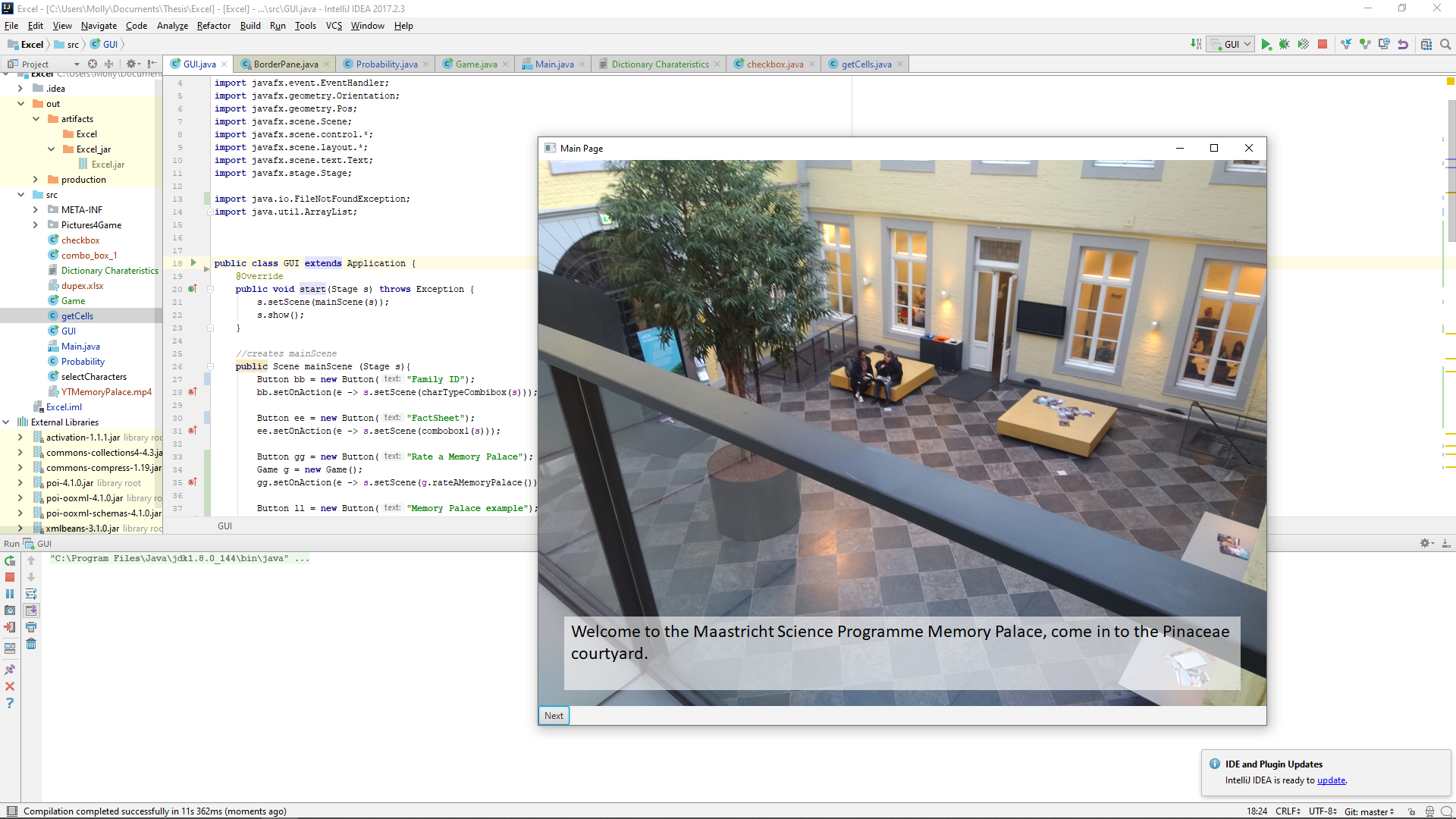
*Figure 5 Screenshot of family ID*

Plant ID: allowed users to fill in characteristics, and then printed out the possible families it could be. Upon selection, checkboxes of the characteristics would be presented and users would select all applicable. Their selection would be weighted using a probability modifier and the most likely plant families are displayed.



*Figure 6 Screenshot of game points allocation system.*

Rate a memory palace: provided a checkbox form for users to fill in and tallied a score of their memory palace.



*Figure 7 Screenshot of slide within the game example*

Memory palace example: a slide show showing an example of the ‘MSP memory palace’ game.

*(see electronic appendix) - Video*

Video explaining a memory palace: A video taken from YouTube describing what the ‘method of loci – memory palace’ concept is, to act as an additional example.

Questionnaire

*(See electronic appendix)*

The questionnaire contained the following questions:

* What topic do you (mainly) study?
* Identify Plant 1
* Identify Plant 2
* Factsheet
  + Speed
  + Usability
  + Layout
  + Comments
* Family ID
  + Did you find the correct family?
  + Speed
  + Usability
  + Layout
  + Comments
* Game
  + Have you used this memorisation technique before?
  + What did you think of the game concept?
  + How well was the game explained?
  + Create your own game using information from a fact sheet
  + Did you enjoy the game?
  + Would you use this technique again?
  + Do you think it would be better filmed and/or narrated?
  + Comments?
* Game rating
  + How many points did I get?
  + How easy was the point system to use?
  + Did you enjoy rating the memory palace
  + Comments
* Other Comments

**Discussion**

Users

The main users are intended to be high school and undergraduate users. Using this demographic for the questionnaire will yield meaningful results for the success of this program.

Quick recall of plant families is necessary for determining plant species as many of the resources are developed by botanists who have this assumed family level knowledge. This knowledge is seldom universal, as researchers often specialise in genus or smaller categories of plants. Without the undergraduate level reaffirmed there might be oversight in the projects taking place. Also, universities often attract researchers from over the world, familiarity with Dutch flora may be beneficial when joining a Dutch university.

Data use

The creation of the matrix provides a useful dataset for use in other research projects about Dutch flora. Although the information was not unique (being aggregated from (Christenhusz, Fay, & Chase, 2017; Van der Meijden, 2005)), the information is not available in a matrix format. This format is considered the optimal when developing identification key software. With programming, a method could be developed which displays this information as a di/polychotomous key, as well as developing other matrix keys. The reverse cannot be done, as dichotomous keys do not provide less information.

Creation of ‘sub’ matrixes is also possible (where not all the plant families are shown, as was conducted in the questionnaire). This would be useful for students or educators where you only wish to learn about the most prevalent families in the area, as is the case with the 50 families’ booklet resource. Learning a subset of families is common in education, as diversity is so high, and some families have few species which are also rarely occurring in an area. Also, people don’t tend to remember so much detail, and the process of learning and applying some data is more important than just memorisation.

Additionally, the data can be used in non-programming research. It is a source of plant family information. For data analysis you could compare the relationships between different plant characteristics. Hypotheses such as two characteristics often occur together, or Dutch plant families have a higher occurrence of a certain characteristic could be investigated. Analysis like this is much faster with the use of spreadsheets rather than paper. Additionally, botanists are more likely to be familiar with spreadsheets than databases and SQL.

Matrix

This paper provides a detailed methodology on how to develop matrixes. It offers a strategy on how to select which characteristics to include. This process has been documented previously but is not always documented. This is because software in the field is often sold commercially so there are intellectual property considerations. The process is often that the botanical data is developed by botanists, which is either kept within an organisation, or behind a paywall. Then it is handed off to programmers to develop an application which is then sold. This process inhibits innovation in the field, as the time needed to do both processes is great, and having both botanical and programming is not so common. This hinderance was also a limitation of this study. If the botanical information was already available in a digital form, the time taken for data gathering could have been much less. Additionally, matrix keys were often only applied for global databases. The smaller scale scope of this thesis tends to utilise dichotomous keys or other resources.

A method partially considered in this paper is the use of more data and machine learning for the inclusion and order for characteristic types for the matrix. Considering the number of occurrences for each characteristic type within literature, could provide a more analytical way of choosing which characteristic to include. Additionally, recording data from within the application could allow analysis of which characteristics are often used. With this data the order of the characteristic type could change to create a better program design. Machine learning would be beneficial when dealing with these large quantities of data. The conceptual aspect of it having training data and test data provides a stronger assurance the changes to the program are beneficial (and not that the data sets used are too small).

As this is a relatively small matrix key, it could provide a good educational tool on matrix keys. When learning aspects of programming small datasets are often used as not to confuse the users and also to make understanding and fixing issues easier. Also, there is a disparity between botanists and programming skills. And tools used to educate often use computer science or physics or mathematical examples, this may be unnecessary daunting for botanists, as understanding of the dataset may be lacking. This ‘afraidness’ is seen in learning plant families, (Kissi & Dreesmann, 2018) states familiarity with mobiles making using their application easier, the same logic could be applied to this scenario.

Program design

The layout of the program emulates the same standard structure of a BorderPane (Fig 1 & 3-6). However, because of the largely different requirements of each sub application, the outcome does not initially appear cohesive. More time spent mapping out the layout of each sub application, especially as additions such as a dictionary would be applied across sub applications, would help to achieve a more user friendly GUI. Cohesion in design would also be more apparent in a more developed version of this application. Artistic choices, such as logos, fonts can colour schemes would also encourage a greater sense of cohesion. Furthermore, these choices would allow users to get a sense of topic (that it is biological and educational) and these features often give a better initial impression.

The program is designed such that the Excel matrix can be interchanged with other matrixes. This ensure the production of this program can go to be utilised in many more research projects. This interchangeability both applies to the matrix and the game. So other data with names (plant families, animals, or even more unrelated things like furniture or book collections) and categoric descriptors (colours, shapes, sizes) could be made into a matrix and put into this system. This approach was so that the research could be progressed from researchers from a variety of backgrounds. As found in research, plant comprehension is low, so this approach prevents limitation on future research.

Probability modifier

To combat (false positive of global characteristics info) probability modifiers that select against uncommon families could be implemented within the database design. Additionally, if the application was to be expanded over a larger region it could be preferable to have a matrix with data applicable globally and relying on probability to act as a selective tool. Ultimately the tool could provide datasets for every location it is used in, but this could be an effect trade off during development

Merriam-webster lookup percentiles are rounded to the nearest 10%, which allows for a finer tuning of the probability matrix by using 10 values for the weighting (Dictionary, 2002). The look up percentiles may not, be a good indicator for the likelihood of students to know the terminology. One source of potential discrepancy is words with more than one meaning, such as ‘heads’ is likely to have a higher look up rate, whilst users are only interested in the non-botanical descriptions. Another consideration is the counter-intuitiveness of using the look up value, a high look up rates suggests a term that is infrequently known, whilst frequently used, however low lookup rates suggest occurrence of the terminology is so infrequent , people don’t usually come across it and therefore don’t look it up, or it could indicate knowledge of it is so universal, lookup is not necessary.

Questionnaire

The questionnaire provides a structure in which the program could be assessed in future development. Additional development is needed to allow the program to be distributed as an executable jar file without error. Currently the questionnaire can only take place in person, utilising IntelliJ, as the external files (images) are not contained within the jar file. The questionnaire is designed to be completed online, with minimal modifications making it anonymous also. This design was in hopes that more people could leave an honest and critical review, and reviews could come from undergraduate students globally.

Game

The game diverges from many of the other educational games. This is because the game is unique. The Kissi & Dreesmann (2018) application is based on a pre existing software, and Jacquemart et al. (2016) relies on webpage design (and structure). This make production of a tailored application more difficult. My game was not limited by this and was able to be an application that can serve as a reference tool as well as a bespoke game. Other bespoke game applications for biological educational topics are an underdeveloped area of research.

Improvements

This matrix was developed from data within Excel. However, using relational database software may be more beneficial during expansion. Excel was an adequate tool as multimedia sources were not used, and all data could be displayed on one table. If the matrix was expanded to detail genus or species information, creating additional tables and defining the relation between them would be more beneficial. Data entry could be sped up as the characteristics found in the family could be automatically applied to the genus within. Creating a relational database would be the best way to deal with the hierarchical structure of plant systematics. Excel is also an impractical tool for storing images or video. Using a database, or any software that can handle more diverse data types, would be beneficial.

The addition of multimedia, and especially images is frequently stated in research. Although it ended up being beyond the scope of this research, it has been seen to be a beneficial tool and should be added.

Conversion of this application to a mobile format would be a potential improvement. There is lots of research stating the benefits of mobile applications: that they are more portable, a high percentage of the population has them. However, this implementation has downsides. Compatibility between different phones, and their operating system is harder to maintain. Screen size is limited so concessions are often made on the amount of information that can be displayed, as in (Kissi & Dreesmann, 2018) the software was ultimately tested on a tablet. If it were a larger scale project, such as that of iNaturalist, this could be maintained. But, the difference in goal, with iNaturalist wanting more images and data to be uploaded, and this program being largely static and infrequent changes in information, dedicating so much man power may be unnecessary (Seltzer, 2019a).

Critical reflection

Upon reflection of my project, I think the greatest benefit and improvement to the work would have been to develop a stronger sense of the direction of my project, so that I would have not attempted so many different things. This work is a matrix key, computer program, a computer game and a questionnaire. Each of these tasks could have be made into their own independent thesis (or at least the major part of one). In taking on many tasks, it made the overall results less complete than I would have liked. Additionally, when writing up the work, I found it very hard to go from topic to topic, whilst trying to maintain a cohesive ‘story’. However, the benefit of this approach is that it generated many new ideas, as it introduced me to information from many different fields: education, global databases, game design, programming, mobile application design, botanical information. But the result feels rushed.

Ultimately, questionnaire result was not what I had hoped. This was because I felt stressed with time restraints. Given more time to dedicate to making a program I was happy with, and developing a questionnaire more, I think I would have been more dedicated into finding participants. In the Kissi & Dreesmann (2018) (and others), they apply a scientific, or documented in literature, approach to questionnaire creation. I did not feel I had time to consider researching and utilising this approach. This lack of ‘depth’ I can see in some aspects of the other areas of my work, but is not always negative, as more restriction on my project were necessary to deliver a result. Perhaps another area of improvement would be to detach my personal feeling toward the work, as others may not be so critical, especially for the questionnaire.

**Conclusion**

To conclude, this paper demonstrates a new approach combining matrix keys and novel educational game to improve learning within botany. It can act as an educational tool for learning plant identification, and a learning tool for matrix keys. Additionally, the game brings a memorisation technique, and thus a useful revision tool for students, into a modern context. Many aspects of this work can be altered for different use cases. The program is designed such that matrixes can be interchanged, with the matrixes being in any field. The game would still function with this change, making this a useful tool across educational fields. There are thus many avenues of future research, with input from programmers’, educators, biologists and more based on this paper.

**References**

Agricultural revolution. (2015). In T. E. o. E. Britannica (Ed.): Encyclopædia Britannica, inc.

Althoff, T., White, R. W., & Horvitz, E. (2016). Influence of Pokémon Go on physical activity: study and implications. *Journal of medical Internet research, 18*(12), e315.

Attigala, L., De Silva, N. I., & Clark, L. G. (2016). Simple Web‐based interactive key development software (WEBiKEY) and an example key for Kuruna (Poaceae: Bambusoideae). *Applications in plant sciences, 4*(4), 1500128.

Balmford, A., Clegg, L., Coulson, T., & Taylor, J. (2002). Why conservationists should heed Pokémon. *Science, 295*(5564), 2367-2367.

Baskauf, S. J., & Kirchoff, B. K. (2008). Digital plant images as specimens: Toward standards for photographing living plants. *Vulpia, 7*, 16-30.

Biodiversity Information Standards (TDWG). (2019). Retrieved from https://www.tdwg.org/

Burgun, K. (2012). What Makes a Game? Retrieved from https://www.gamasutra.com/view/feature/167418/what\_makes\_a\_game.php

Christenhusz, M. J., Fay, M. F., & Chase, M. W. (2017). *Plants of the World: an illustrated encyclopedia of vascular plants*: University of Chicago Press.

Class BorderPane. Retrieved from <https://docs.oracle.com/javase/8/javafx/api/javafx/scene/layout/BorderPane.html>

Dictionary, M.-W. (2002). Merriam-Webster. On-line at http://www. mw. com/home. htm.

Erkens, R. (2019). *A guide to 50 often encountered Dutch plant families during Field skills in biology (PRA2009)*. Maastricht Science Programme.

Extra Credits: Games in Education. (2013).

Fassbender, E., & Heiden, W. (2006). The virtual memory palace. *Journal of Computational Information Systems, 2*(1), 457-464.

Hagedorn, G. (2007). *Structuring descriptive data of organisms—requirement analysis and information models.*

Hebert, S. (2018). The Power of Gamification in Education | Scott Hebert | TEDxUAlberta: TED.

International Year of Biodiversity. (2010). Retrieved from https://www.iucn.org/content/international-year-biodiversity

Jacquemart, A.-L., Lhoir, P., Binard, F., & Descamps, C. (2016). An interactive multimedia dichotomous key for teaching plant identification. *Journal of Biological Education, 50*(4), 442-451.

Juul, J. (2018). The game, the player, the world: Looking for a heart of gameness. *PLURAIS-Revista Multidisciplinar, 1*(2).

Khan, A., Ahmad, F. H., & Malik, M. M. (2017). Use of digital game based learning and gamification in secondary school science: The effect on student engagement, learning and gender difference. *Education and Information Technologies, 22*(6), 2767-2804.

Kiesler, S., Kraut, R. E., Koedinger, K. R., Aleven, V., & Mclaren, B. M. (2011). Gamification in education: What, how, why bother? *Academic exchange quarterly, 15*(2), 1-5.

Kirchoff, B., Remington, D., Fu, L., & Sadri, F. (2008). *A new type of image-based key.* Paper presented at the 2008 International Conference on BioMedical Engineering and Informatics.

Kissi, L., & Dreesmann, D. (2018). Plant visibility through mobile learning? Implementation and evaluation of an interactive Flower Hunt in a botanic garden. *Journal of Biological Education, 52*(4), 344-363.

Learn a language for free. Forever. (2019). Retrieved from https://www.duolingo.com/

Leggett, R., & Kirchoff, B. K. (2011). Image use in field guides and identification keys: review and recommendations. *AoB Plants, 2011*.

Mark, J. J. (2019). Sushruta. Retrieved from https://www.ancient.eu/sushruta/

Pinola, M. (2019). The Science of Memory: Top 10 Proven Techniques to Remember More and Learn Faster. Retrieved from https://zapier.com/blog/better-memory/

Plitzner, P., Henning, T., Müller, A., Güntsch, A., Karam, N., & Kilian, N. (2017). *Bottom-up taxon characterisations with shared knowledge: describing specimens in a semantic context.* Paper presented at the S4BioDiv@ ISWC.

Pokémon GO. (2018). Pokémon Nintendo / Creatures.

Pokémon GO. (2019). Retrieved from https://play.google.com/store/apps/details?id=com.nianticlabs.pokemongo&hl=en

Seltzer, C. (2019a). About. Retrieved from https://www.inaturalist.org/pages/about

Seltzer, C. (2019b). Teacher's Guide. Retrieved from https://www.inaturalist.org/pages/teacher's+guide

Silva, H., Pinho, R., Lopes, L., Nogueira, A. J., & Silveira, P. (2011). Illustrated plant identification keys: An interactive tool to learn botany. *Computers & Education, 56*(4), 969-973.

Stagg, B. C., & Donkin, M. (2013). Teaching botanical identification to adults: experiences of the UK participatory science project ‘Open Air Laboratories’. *Journal of Biological Education, 47*(2), 104-110.

Taking Action for Biodiversity. Retrieved from https://www.cbd.int/2011-2020/about

Theophrastus (371 - 287 BC). (2011). Retrieved from https://www.tcd.ie/Botany/tercentenary/origins/theophrastus.php

Thomas, R. L., & Fellowes, M. D. (2017). Effectiveness of mobile apps in teaching field-based identification skills. *Journal of Biological Education, 51*(2), 136-143.

Van der Meijden, R. (2005). *Heukels flora van nederland 23e druk*: Wolters-Noordhoff.

Vries, H. d., Roeke, T., & Lemmens, M. (2018). *Observation.org, Nature data from the Netherlands*. Retrieved from: https://www.gbif.org/dataset/8a863029-f435-446a-821e-275f4f641165/metrics

Watson, J. (2012). *Reality ends here: environmental game design and participatory spectacle*: University of Southern California.

Winston, J. E. (1999). *Describing species: practical taxonomic procedure for biologists*: Columbia University Press.

Yates, F. A. (1966). The Art of Memory. Chicago: U: of Chicago Press.

Zaiane, O. (1998). The Object-Oriented Data Model. Retrieved from https://www2.cs.sfu.ca/CourseCentral/354/zaiane/material/notes/Chapter8/node3.html