X Modal
X Cultural
X Lingual
X Domain
X Site
Global OER Network

Grant Agreement Number:	761758
Project Acronym:	X5GON
Project title:	Cross Modal, Cross Cultural, Cross Lingual, Cross Domain, and Cross Site Global OER Network
Project Date:	2017-09-01 to 2020-08-31
Project Duration:	36 months
Deliverable Title:	D6.2 – Report of in-the-wild studies investigating performance and usability of the initial WP6 services for virtual and real-world adaptive learning
Lead beneficiary:	UCL
Type:	Report
Dissemination level:	Public
Due Date (in months):	31 August 2019
Date:	31 August 2019
Status (Draft/Final):	Final
Contact persons:	Stefan Kreitmayer, Yvonne Rogers, Emine Yilmaz and John Shawe-Taylor

Revision

Date	Lead author(s)	Commments
22 August 2019	Stefan Kreitmayer	Initial draft
30 August 2019	Stefan Kreitmayer	Revised draft in response to feedback from Yvonne Rogers and Richard Noss
31 August 2019	Stefan Kreitmayer	Final version

Contents

List of Figures

1 Introduction

WP6 focuses on understanding and improving the learner experience with Open Educational Resources (OER). This report summarises the research aims, methods and outcomes in M12-M24 which concentrated on self-directed learners as the primary user group. All the studies reported here relate to Task 6.2: "In-the-wild studies of the use and experiences with the initial X5GON services".

The research questions addressed in this report followed logically from previous findings, including the empirically validated FRAME taxonomy which conceptualises how learners can Find, Recommend, Assess, Manage and Engage with OER (see D6.1). Future work outlined in Year 1 was partly addressed in Year 2. Three key challenges were identified in Year 1. These challenges revolved around the questions of how to design appropriate interfaces for helping learners (1) preview content, (2) reflect on their progress and (3) involve peers in their learning journey [1]. These three challenges, namely Preview, Progress and Peers (PPP) were addressed using two in-the-wild studies which entailed designing two different user interfaces for studying with OER. The design, evaluation and description of both studies were guided by the PPP and FRAME taxonomy.

The first in-the-wild study focused on peer involvement by investigating how learners can benefit from using OER in pairs. 72 pairs of lifelong learners were recruited and observed while using a given set of OER materials online and face-to-face. The study resulted in a comprehensive understanding of individual learner's needs and expectations and the variety of ways in which pairs can collaborate, using a mix of novel and traditional interfaces. The study found that both face-to-face and online collaboration can be effective and rewarding. However, the findings also highlight the importance of careful matchmaking, particularly taking into account individuals' background knowledge and study interests.

The second in-the-wild study focused on helping individual learners with previewing content and reflecting on their progress. This iterative design study led to the development of the X5Learn dashboard which has played important roles in multiple WPs (see below and D1.3, D5.2, D7.3 and D8.3). A key outcome of this study was a new type of document visualisation that allows learners to efficiently preview, explore and discover information in large collections of documents while maintaining focus on their primary learning task. In order to generate these visualisations automatically, a new method had to be devised to dynamically segment content and enrich each segment with annotations using a topic extraction method (see D1.3). Evaluations showed that topic annotations can enable new and engaging ways for learners to interact with OER-based knowledge representations that are directly applicable to lifelong learning. Future potential for topic annotations to enable active and social learning was also highlighted as an outcome. For instance, by tracking learners' engagement with topics over time, it is possible to infer individual learner profiles that can be used for matchmaking between peers and content. A further side effect of the topic extraction method was that the resulting annotation data can be reused for other purposes, such as the automatic detection of duplicate documents in order to improve the quality of search results.

Both studies used a similar methodology as in Year 1 [1], relying heavily on participatory design, iterative prototyping and detailed qualitative analysis of system logs, questionnaires and interviews with users and domain experts. Additional lab-based evaluations played a supporting role while investigating specific aspects of the user experience by means of predefined tasks that participants performed under controlled conditions. Moreover, additional insights were gained through interviews with educators and practitioners in various fields of education. The main part of the research approach, nevertheless, focused on data from learners in naturalistic scenarios. All design and evaluation methods were informed by recent literature in the fields of Education, Technology-Enhanced Learning (TEL), Information Retrieval (IR), Explainable Artificial Intelligence (XAI) and Human Computer Interaction (HCI).

The following sections describe the two in-the-wild studies, followed by a discussion of findings and plans for Year 3.

2 Study 1: Support for pair learning

This section presents the first in-the-wild study, focusing on pairs of self-directed learners using OER.

2.0.1 Rationale

Open Educational Resources, such as video tutorials, have become a popular way for self-motivated learners to develop skills, such as learning about machine learning. However, many individuals find it difficult to get started and stay motivated over time. For instance, dropout rates in MOOCs are notoriously high for a variety of reasons [2]. Lack of social connection has been identified as a major issue with remote learners, since building an engaging learner community through forums and video conference technologies is not straightforward [3]. A survey of learner attitudes found that learners associate face-to-face courses with more opportunities for interaction and better learning outcomes than online courses [4]. Prior research investigating blended approaches has suggested the use of MOOC material in face-to-face settings as a promising solution [5]. Moreover, empirical evidence suggests that having a (remote) "study buddy" can benefit the learning experience. Madland et al [6] showed that study buddies can support each other socially, emotionally, providing that both are equally engaged. In order to support engagement, "planning prompts" were shown to be effective in MOOCs [7] with individual learners but evidence is scarce regarding the effectiveness of planning prompts in pairs.

2.0.2 Motivation

This study explored the feasibility of using OER video tutorials in pairs in order to leverage well-known benefits of collaborative in a self-directed setting without the rigidity of large, expert-facilitated courses. Participants were self-motivated adults with little or no background knowledge in the field of machine learning who were keen to learn about machine learning with a study partner. Over the course of 10 weeks, each pair was observed as they followed a series of machine learning tutorials over 10 weeks while collaborating face-to-face or remotely. Insights about the feasibility, benefits and challenges of pair learning journeys were gained by analysing system log data, questionnaires and semi-structured interviews with participants. The findings demonstrate that pair learning with OER video tutorials can be effective and motivating and the benefits can sometimes exceed learners' expectations. Both face-to-face and remote conditions are feasible although pros and cons exist, preferences vary between learners and hybrid forms can also be attractive. New challenges are highlighted regarding the design of novel interfaces for matchmaking, content recommendation and using OER as a pair.

The study was initiated in Year 1 and continued into Year 2. The data collection and most of the data analysis are now complete and a draft paper is under internal review. The aims, method and preliminary findings of this study are detailed in D6.1. Newer findings are presented below.

NB: The series of video tutorials used in this study was commonly referred to as a "course" by participants. Whenever the word "course" is used in the following, readers should be aware that the intervention differed from today's typical online courses. Particularly, it came without any tests, tutors, cohorts, interaction with the instructor, automated emails or certificates. Care was taken by the researcher to stay in his role as observer rather than acting as an instructor.

2.1 Background and aims

Related work and aims of this study are detailed in the previous report (D6.1). To summarise, the study explored the following hypotheses:

H1: Pairs who meet face-to-face have a more satisfying learning experience (measured by the number of tutorials completed) than pairs who meet remotely using Skype.

H2: Pairs who make a pact that sets out agreed commitments (e.g. whether or not to do extra work between meetings) will complete more tutorials compared to those with no pact.

2.2 Method

To test the hypotheses, a 2x2 factorial design was chosen. 60 pairs were recruited and assigned one of 4 conditions (15 pairs in each condition):

1. f2f-pact
2. f2f-nopact
3. remote-pact
4. remote-nopact

For each pair, we measured satisfaction by counting how many tutorials (out of 10) each pair completed over the course of ten weeks. Differences between the group means were analysed using a two-way ANOVA test.

Every participant was interviewed at the beginning of the study and given a questionnaire to asses their prior domain-knowledge and attitudes using multiple-choice questions and 5-point Likert scales.

For each pair, system logs were analysed, including the time of access and user comments for each tutorial.

Semi-structured post-interviews were conducted with participants who were available, using a medium of their preference (usually phone or Skype and occasionally email).

2.3 User interface design

A dedicated user interface was designed in the form of a simple web application (see D6.1). Each pair was provided with a shared login, allowing them to access the videos, track their progress, take notes and leave feedback. Pairs were instructed to use the interface together simultaneously, e.g. sitting next to each other or by sharing the screen via Skype.

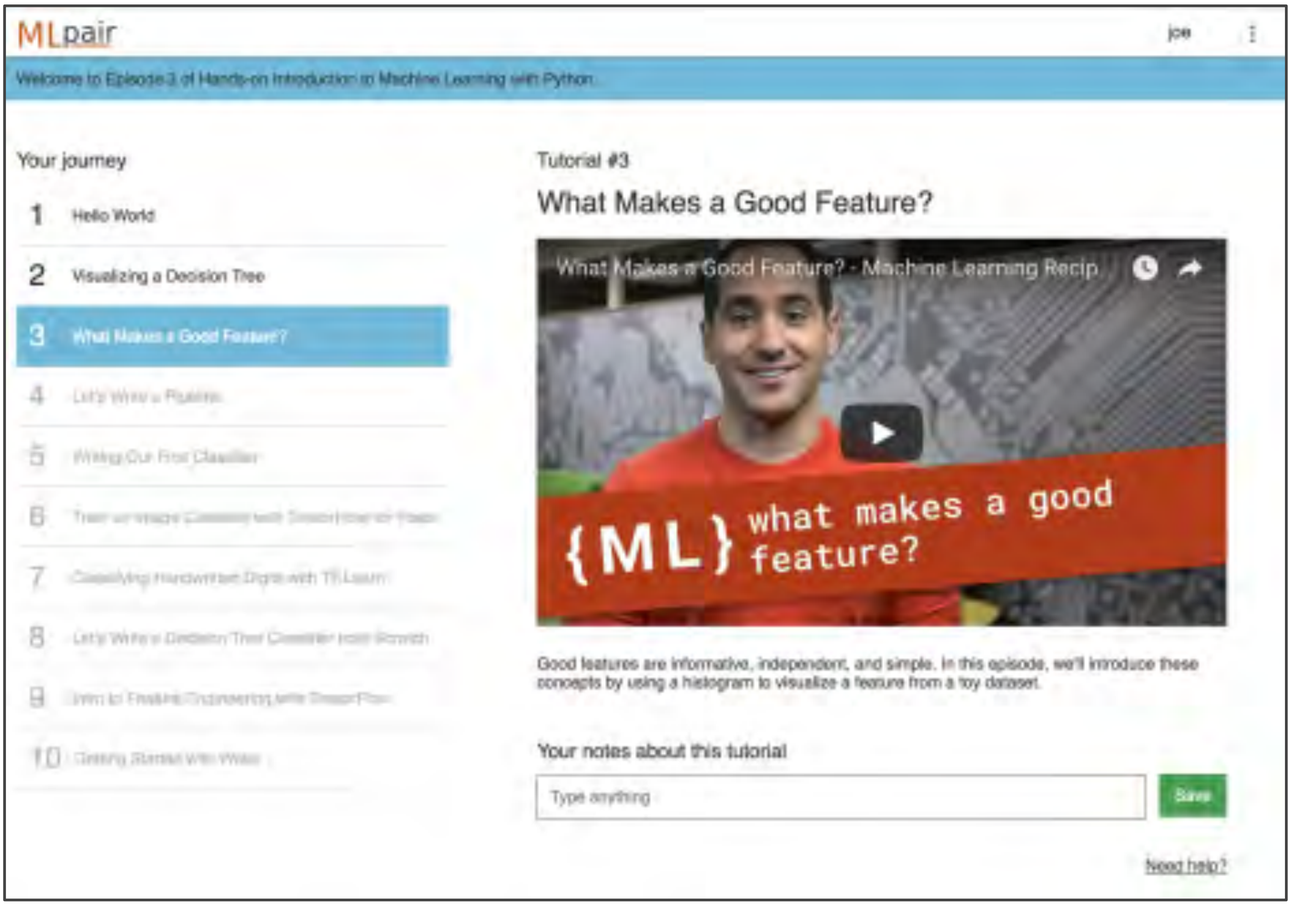

Figure 1: Screenshot of the pair learning interface.

The main design goals for the interface were as follows:

1. Shared account: Allowing both partners to login as a pair, using a memorable URL and team login, in order to share access to the material, progress and notes.
2. Easy to use: Using conventional design patterns and avoiding complicated or unnecessary features.
3. Distraction-free: Avoiding any features that could potentially distract from the content, such as recommendations or pop-ups.
4. Active learning: Encouraging users to take notes and do the exercises described in each video before moving on to the next.

Each tutorial typically contained some explanations of a new concept or technique, followed by a hands-on exercise with code examples. In order to proceed to the next tutorial, participants had to confirm that they completed the exercise by answering "yes" in a simple confirmation dialogue. This trust-based mechanism was considered appropriate, since the course offered no certificate or other reason for participants to cheat the system. Completion times and user comments were logged for analysis.

2.4 Findings

The following sections summarise the key findings from this study. First the results from the quantitative analysis of completion rates are presented, followed by qualitative accounts of how participants used the shared interface, how pairs proceeded through the course, perceived benefits and risks of pairing, trade-offs between face-to-face and remote collaboration, and participants' attitudes towards matchmaking. The FRAME taxonomy is used to describe additional aspects that contribute to a holistic picture of the intervention and preliminary findings.

2.4.1 Quantitative analysis of pair completion rates

Pairs who failed to meet at least once were excluded from the analysis. One pair who had initially been recruited for the remote condition was transferred to the face-to-face condition as the initial interview revealed that both participant lived in the same household. The final numbers of pairs in each condition are:

- 16 f2f-pact
- 14 f2f-nopact
- 15 remote-pact
- 11 remote-nopact

2.4.2 Face-to-face vs remote

Between face-to-face and remote pairs, the average completion rates were very similar, with greater variance in the remote condition.

Mean completion rate: 5.77 ($s = 2.96$) in f2f pairs, 5.83 ($s = 3.49$) in remote pairs.

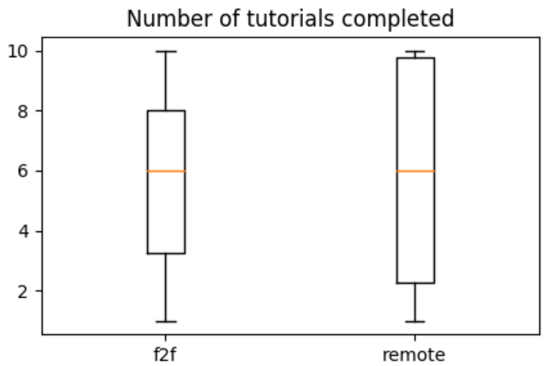

Figure 2: Comparison of face-to-face and remote pairs

2.4.3 Pact versus no-pact

Mean completion rates in each condition:

5.8 (s = 3.05) in f2f-pact pairs

5.73 (s = 2.97) in f2f-nopact pairs

5.31 (s = 3.4) in remote-pact pairs

6.43 (s = 3.63) in remote-nopact pairs

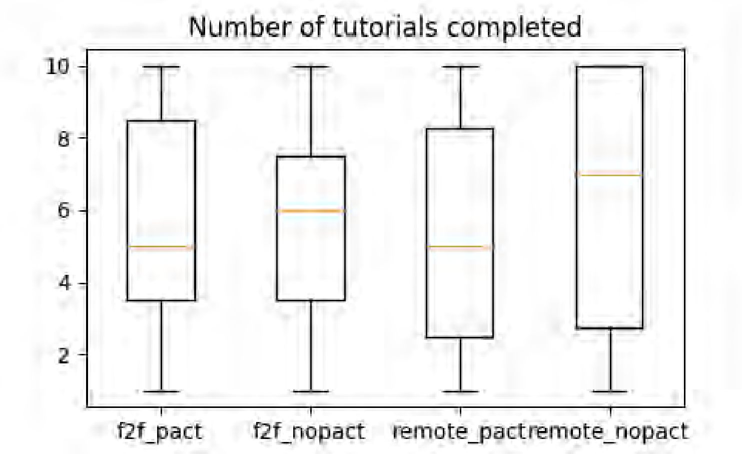

Figure 3: Comparison of different conditions

2.4.4 Analysis of Variance Test

A two-way ANOVA test did not show a significant difference between the conditions.

Univariate Analysis of Variance

Between-Subjects Factors

		N
location	f2f	26
	remote	30
intervention	nopact	25
	pact	31

Descriptive Statistics

Dependent Variable: completion

location	intervention	Mean	Std. Deviation	N
f2f	nopact	5.7273	2.96954	11
	pact	5.8000	3.05193	15
	Total	5.7692	2.95713	26
remote	nopact	6.4286	3.63137	14
	pact	5.3125	3.40037	16
	Total	5.8333	3.49466	30
Total	nopact	6.1200	3.30807	25
	pact	5.5484	3.19206	31
	Total	5.8036	3.22727	56

Tests for Heteroskedasticity

F Test for Heteroskedasticity[a,b,c]

F	df1	df2	Sig.
.298	1	54	.587

a. Dependent variable: completion

b. Tests the null hypothesis that the variance of the errors does not depend on the values of the independent variables.

c. Predicted values from design: Intercept + location + intervention + location * intervention

Figure 4: Overview of ANOVA test results

Tests of Between-Subjects Effects

Dependent Variable: completion

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	9.391[a]	3	3.130	.289	.833	.016
Intercept	1857.317	1	1857.317	171.410	.000	.767
location	.157	1	.157	.014	.905	.000
intervention	3.734	1	3.734	.345	.560	.007
location * intervention	4.848	1	4.848	.447	.507	.009
Error	563.448	52	10.836			
Total	2459.000	56				
Corrected Total	572.839	55				

a. R Squared = .016 (Adjusted R Squared = -.040)

Figure 5: ANOVA test results

2.4.5 Shared interface use

The feedback from participants regarding the usability of the interface was consistently positive overall, with several participants describing the interface as pleasant and easy to use. The note-taking feature was used frequently, resulting in a substantial amount of qualitative data.

Many of the face-to-face pairs used the interface as expected, using either one or two computers physically next to each other. However, there were also cases where the interface was appropriated in unexpected ways. Particularly, the assumption that the pairs would always use the interface simultaneously was confirmed in some, but not all cases. For instance, one pair reported that they met in a coffee shop where the high ambient noise and poor internet bandwidth prevented them from watching the videos together. The pair therefore decided to do the tutorials individually (as homework) and use the meetings to compare and discuss their solutions, sometimes involving printouts of programming code.

In the remote condition, some pairs decided to switch from screen sharing to asynchronous or semi-synchronous collaboration, using email or instant messaging (e.g. WhatsApp). In some cases, the reasons given included poor internet bandwidth affecting the user experience of screen sharing.

In both conditions, there were occasional reports of one participant disengaging from the course while the other decided to proceed on their own.

The log data showed that the tutorials were typically accessed in the intended sequence and that most pairs did the exercises before proceeding to the next tutorial. Some pairs, however, reported that they occasionally decided to "cheat" the system in order to preview upcoming videos. They did so either by falsely answering yes or by looking up the tutorials on YouTube.

2.4.6 Different ways to continue

The assumption that both participants would use the interface simultaneously was confirmed in many but not all cases. For instance, some pairs reported that their preferred way to collaborate was to meet physically in a quiet room and share one computer to do the work. Other pairs who met in coffee shops reported that the noise and poor wifi made watching the videos together unpractical so they decided to do the exercises individually as homework and use the face-to-face meetings to discuss each other's solutions. In some of these cases, the pair agreed on a shared pace of one tutorial per week in order to stay in sync, while other pairs went out of sync after a while. Overall, we identified four typical patterns, as shown in Figure 6:

1. The pair completed the course together.
2. The pair did a few tutorials together but then decided to stop.
3. One participant dropped out while the other continued.
4. Both participants continued, each at their own pace, while keeping each other updated.

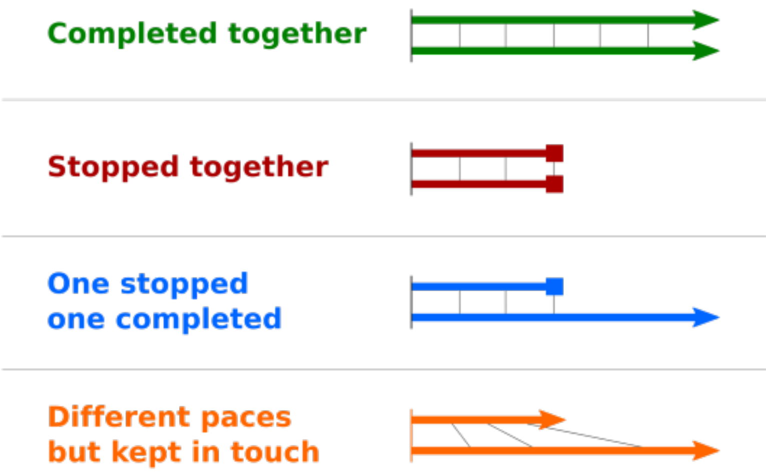

Figure 6: Four different types of pair learning journey

The following description illustrates an example of the last category where both partners completed the course at different paces. Both participants explained in separate interviews that this arrangement was mutually agreed and justified by differences in background knowledge, interest in detail and life circumstances: "We had different ideas about what to do. [...] We did a bit of learning together [...] Then we discovered that we are on different paths." / "He is in a very busy time [. . .] he hasn't programmed in a long time. I have much time [. . .]." Despite the lack of synchrony, both expressed a sense of motivational benefits of having a partner: "I went through all the videos [independently]. Having said that, committing to go through the course with somebody else does make it harder for me to quit."

In cases where an individual decided to discontinue their participation, the reported reasons ranged from domain-specific considerations (e.g. learned enough or interest shifted) to changes in external circumstances, such as pregnancy, a new job, or moving abroad. The other person then either continued the course on their own or stopped as well. Reconnecting a different partner was not possible in this study, due to experimental constraints.

2.4.7 Perceived benefits and risks of being in a pair

Most pairs overcame their learning challenges in a self-sufficient way through joint problem solving. Occasional emails asking the researcher for help were the exception. As mandated by the study design, the researcher took care to stay in his role as observer rather than acting as an instructor.

As mentioned in the summary of preliminary findings in D6.1, having a study companion was perceived as helpful and motivating by the majority of participants. Many said that having a partner was essential for them to stay motivated when facing difficult challenges, such as complex new topics and confusing error messages. Many said that without a partner to make them feel accountable and committed to a time schedule they would have given up, e.g. "I wouldn't have been so persistent through the course if I hadn't felt like: Hey, I'm meeting [name] and we have something to do together."

Other frequent themes revolved around the enjoyment of shared problem solving, socialising, getting help with installation and troubleshooting, the joy of helping the other person, not feeling alone, learning from more advanced peers and learning by teaching.

"She comes up with very good questions. She made me think. She actually challenged me and I really enjoyed that. I had to read something up and when we met next, I kind of explained it to her. It's wonderful. The experience is really great."

A sense of satisfaction was common among pairs who stayed together. However, on at least one occasion, a pair was found to stay together out of politeness, despite considerable dissatisfaction on both sides. The pair had agreed to meet weekly at work during a fixed, 30-minute lunchtime slot. One of them, an experienced programmer, said that the short time limit forced them to rush through the tutorials, leaving little time for turn-taking, questions and explanations. The other participant, a novice at programming, found themselves unable to keep up during the sessions or catch up in between: "I didn't gain so much confidence because I didn't get much practical experience. [My study partner] did all the work. It would have been helpful if I had done the exercises between the meetings on my own." Both hinted at a mild level of awkwardness on both sides which they ascribed to unfortunate circumstances. Both explained that being employed at the same large company (but not colleagues) implied certain expectations of social protocol that made it difficult to change the arrangement.

By analysing interview transcripts and participants' comments in the interface, the following themes were identified:

Benefits:

- Having someone to ask for help quickly and easily
- Joy of explaining to each other
- Joy of tackling problems together
- Sending each other interesting links
- Noticing when the other one gets stuck and helping out
- Looking forward to the next meeting
- Feeling accountable for the other

Risks:

- Different levels of commitment or different preferred pace
- Awkwardness (for both) when skills are too different
- Feeling guilty about falling behind or running on ahead
- Being alone in case the partner drops out
- Poor personal chemistry
- Continuing for the wrong reasons, e.g. to save face

2.4.8 Face-to-face versus remote collaboration

In post interviews, many participants who collaborated face-to-face said that they enjoyed the experience because of the human connection and shared focus which they could not imagine from studying in a remote setting. One participant argued: "You get a deeper human connection than on Skype" and "sitting next to each other, we're both really focused."

On the other hand, some participants said that remote collaboration can be a viable alternative: "I travel, so meeting in person isn't always an option." Moreover, some pairs chose to work individually and discuss their "homework" in regular or spontaneous meetings: "We catch up through WhatsApp constantly. It's great!"

2.4.9 Analysis of learner notes

While studying, most pairs made extensive use of the note-taking functionality of the provided user interface. This resulted in a total of 579 comments which were analysed by the researchers. The analysis consisted of three phases. The first two phases resulted in a coding scheme and list of keywords for each category. Automatically coding the comments using keywords resulted in an initial table which was then manually verified in a final phase of analysis.

The resulting coding scheme contained 7 categories as listed below, including keywords and examples. Multiple categories per comment were possible.

Category 1: Technical infrastructure Key observation: Comments related to the installation, maintenance or compatibility of software tools that are not specific to machine learning Keywords: docker graphviz colab spyder ide notepad jupyter install installing installation 2.7 deprecation deprecated conda anaconda miniconda version windows mac linux Example: “Really struggled with getting all the packages right in Anaconda on Win10”

Category 2: Learning about machine learning Key observation: Reference to the content or presentation of the tutorial, including the explanations, difficulty, didactic aspects, pace, code or use of extra material. Keywords: classifier tutorial presenter presentation fast slow example examples Example: “Got general idea. Idea of subsets is logically and fits in with programming algorithm. Presentation went very fast. Could be difficult to keep up at times.”

Category 3: Affective and emotional aspects of the learner experience Key observation: Strong expression of positive affect (e.g. enthusiasm, enjoyment, looking forward, ...) or negative affect (frustration, disappointment...) Positive keywords: great! cool! interesting! awesome fun excited exciting fantastic enjoyed Negative keywords: frustrating frustration frustrated annoying annoyed sadly pain painful disappointed Example: “This one was exciting!”

Category 4: Problem solving Key observation: Reference to problems and solutions (or absence thereof) that the pair encountered. Focus on problems that relate to the knowledge domain, as opposed to e.g. software installation. Keywords: stackoverflow overflow experimenting experiment try tried trying google googling googled search solve solving solved resolving resolve resolved solution tricky later worked trouble troubleshoot troubleshooting Example: “It worked, but again took a lot of troubleshooting.”

Category 5: Collaboration with the partner Key observation: Use of first-person plural, reference to partner, or any indication that (some of) the work was done collaboratively Keywords: we us our ourselves pair together Example: “We got it running”

Category 6: Individual work Key observation: Use of first-person singular, or any indication that the commenter did (some of) the work individually Keywords: i me my myself Example: “I wasn’t able to visualize the graph in pdf format”

Category 7: Help seeking Key observation: Addressing the researcher directly or indirectly Example: “I wasn’t able to install Docker in my machine as I am missing some of the components required for the installation. Is there a way to access it via Collab instead of installing on my machine?”

Analysing the comments in this way revealed the fact that some pairs used the interface in different ways than the researcher had expected. Some participants wrote their comments from a “we” perspective (see Category 5 above) while others used the word “I” (Category 6). This observation was further investigated in follow-up interviews, leading to a detailed understanding of the different ways in which pairs chose to collaborate. For instance, in several cases it turned out that the partners decided to study individually, each at their own pace, while using regular meetings to catch up and discuss.

At the time of writing, a more detailed analysis of the comments is in progress with a focus on differences between face-to-face and remote pairs.

2.4.10 Attitudes towards matchmaking

As part of the study design, participants were given the choice between signing up with a partner or being matched with another participant based on compatible time schedules and locations. 33 pairs were matched by the researcher and 23 pairs were self-paired. Time and location were the only criteria used by the researcher when matching participants. Other aspects, such as prior knowledge, interest, motivation, age, gender, nationality, etc were not taken into account and no assumptions were made by the researcher about which individuals would fit well together. This simple and neutral strategy was chosen deliberately for two reasons. Firstly, to avoid bias in the matching. Secondly, to better estimate the potential for automating the matchmaking process in the future by trying to anticipate the kind of matches that a simple algorithm would produce. Indeed, while the task was eventually performed by a human, using emails and spreadsheets, it could have been automated. The manual approach was chosen due to the novelty of the task and the reasonably small number of participants required for the study.

Participants were found to be surprisingly open to the idea of being matched with a stranger. Despite the option of sign up with a friend, more than half chose to be matched up with a stranger. When asked about their choice during initial interviews, participants gave various explanations. Some said that they had considered asking among friends or family but wanted to avoid a risk of being turned down or feeling that they know too little about the other person’s interests. One participant who regularly went to a machine learning meetup pointed out that the meetup would be a good place to ask but decided against this option because they felt that such a request would not be considered normal in that environment.

Self-pairing was not found to be a predictor of satisfaction, as many of the highly satisfied pairs were paired by the researcher. Some self-matched pairs completed the entire course and some dropped out early. The same was the case for researcher-matched pairs.

Reflecting on their own experience, one participant pointed out: “I would rather pair with a stranger than a friend. Because friends can easily procrastinate together.” Another participant said: “With someone you don’t really know it’s harder to let them down that than with a friend or colleague”. Many participants who were artificially matched spoke enthusiastically about how much they enjoyed studying with the other person. Several said that they found a new friend in their partner or that they embarked on another course or project together, indicating that the matching led to favourable outcomes that exceeded participants’ expectations and extended beyond the scope of the study.

However, not all matchings were considered a success. More than one participant stopped responding at some point or expressed frustration about their partner being unreliable or difficult to reach. In these negative cases, it was not always feasible to pinpoint the exact causes. However, personality may have played a role, as some of the positive comments suggested: "We are very much on the same wavelength and that made the experience very very pleasant." Another participant suggested: "As long as a person is interested, open, humble and flexible, they will keep the other one on their toes - no matter what their background and education is."

2.4.11 Finding, Recommending, Assessing, Managing and Engaging (FRAME)

The FRAME taxonomy was used to describe five relevant aspects of user interaction with OER:

1. Finding OER: In this study, the learning materials were predefined. However, many pairs reported that they used additional resources, such as Wikipedia, blog articles, forum posts (particularly www.stackoverflow.com) and GitHub issues. Typically, these resources were found by entering specific questions or error messages into a generic web search engine in order to look up detailed concepts and solutions as required by the task at hand. Google Search was a popular choice among search engines. None of the participants reported using an OER-specific search engine, such as http://oercommons.org or https://oasis.geneseo.edu. Searching together, rather than alone, was described by several participants as surprisingly enjoyable, stimulating and productive.

2. Recommending: Some of the tutorials included recommendations of further reading materials. The researchers decided against adding additional recommendations through the web interface in order to avoid distraction. The X5GON recommender engine was not fully developed at the time of the study. Peer recommendation was observed to some extent, as some pairs reported sending each other links to additional resources they found.

3. Assessing: The materials were chosen by the researcher based on various criteria, including factual accuracy, production quality, popularity and technical relevance. Many participants commented on the quality and usefulness of the materials. This feedback contributed to a better understanding of how learners assess OERs and which criteria are considered important. For instance, some learners mentioned that the presenter spoke very quickly, causing them to pause the video frequently. Another aspect was that software versions used in the tutorials did not always match the installed versions on participants' computers.

4. Managing: In order to keep the interface and the learning journey simple, the materials were presented in a linear and straightforward fashion. The purpose of this design goal was to avoid decision paralysis in pairs and keep participants focused on the task at hand. Some participants mentioned that during the course they collected useful additional resources, using their preferred tools such as Evernote, browser bookmarks or simple text files.

5. Engaging: Participants' engagement falls into four main categories: Watching the videos, doing the exercises, discussing and commenting. Therefore the learning was social and active.

2.5 Discussion

Below we discuss the findings, followed by implications for future design.

Studying as a pair was described as a highly rewarding experience by individuals who worked closely with their study partners. Although the quantitative analysis did not confirm the expected tendency for face-to-face pairs and pact pairs to complete more tutorials, there were striking individual preferences and rationales given for the approach of each pair. A common theme among highly satisfied pairs was that they developed a trusting and respectful connection with their partner, characterised by mutual awareness, empathy, courtesy and feeling accountable. This observation is in line with prior evidence suggesting that social and emotional support are important factors for satisfaction in study buddies [6]. Moreover, it was found that highly committed pairs found a variety of flexible arrangements in response to situational demands, such as noisy cafés and poor internet connection. The diversity of effective strategies far exceeded the set of options provided by the pact interface. Therefore, our findings suggest that sometimes it is better not to engineer learning practices but let them emerge from a beneficial relationship between both learners. Having a supportive interface is key to allowing flexible, individual study practices to emerge.

The finding that many participants preferred to be paired with a stranger relates to a prior study that investigated the behaviour of peer tutoring students and found that strangers gave different kinds of feedback to peers than they did to friends [8].

2.5.1 Implications for future design

Implications of the above findings for future research and design are discussed below, including aspects of interface design, offline collaboration, matchmaking, exit and re-entry strategies, previewing content, progress tracking and peer involvement.

2.5.2 Designing user interfaces for pairs

The interface was generally considered successful in allowing participants to preview content, track their progress and support peer collaboration. However, it was also found that participants appropriated the interface in unexpected ways. Particularly, the designers' initial assumption that pairs would always use the interface simultaneously was confirmed in some cases, but not with all of the pairs. The findings suggest that some learners have a need to switch between synchronous collaboration, asynchronous collaboration and individual work. Novel interfaces for studying with OER should allow this flexibility through appropriate interface mechanisms. An aspect for consideration might be to use a mix of personal devices (such as mobile phones) and shareable devices (such as laptops or tablets).

2.5.3 Offline collaboration using online material

Our findings suggest that it is timely to reconceptualise what is meant by "online learning". In this study, we described a situation where the lecture materials were provided on the web (YouTube videos) but the learner-learner interactions took place in one-on-one conversations. Unlike typical online courses, web forums and social media, these conversations were private and "off the record", providing a safe space for learners to build ideas and reflect. Many learners spoke enthusiastically about how easy, efficient and satisfying it was to (a) ask their partner when something was unclear, (b) explain newly learned concepts and processes to each other, (c) search the web for information together, (d) devise more and better solutions than each individual could have come up with on their own, (e) solve new problems independently without the direct intervention of an instructor or tutor. Similar benefits of joint problem solving have been found in earlier studies in the context of pair programming [9].

The model described here differs from the classic MOOC model where learners are encouraged to communicate through written communication in forums and social media features. Some of the potential advantages of written formats are that postings persist over time and can reach a wider audience, thereby increasing the likelihood of qualified answers. However, sometimes it is more important to talk to someone familiar who is available instantly and happy to help, rather than pausing one's task to write a forum post and waiting for a response which may take hours or days. Another issue with MOOCs is that learners can be hesitant to expose themselves to random strangers via forum or video chat, which impedes peer-to-peer interaction [3]. In our pairing model, strangers who were matched initially got to know each other over repeated meetings with the same person, thereby accumulating opportunities for social bonding and building trust. With some learners, this dynamic seemed to benefit their confidence and motivation to continue.

Whether the pairs met face-to-face or through digital channels may be less important than the fact that they had an authentic human connection, allowing them to bring their ideas to life in a familiar, private, one-on-one conversation.

2.5.4 Matchmaking outcomes and reflections

In cases where an individual discontinued their participation in the course, a variety of reasons were reported, ranging from changes of interest in the topic to changes in life circumstances, such as pregnancy, a new job, or moving abroad. The wide range of reasons is in line with prior findings with MOOC participants [2]. In contrast to MOOCs, however, disengaging was sometimes associated with negative consequences for the other person. Several participants who had lost their study partners and wanted to continue asked the researcher to be re-assigned a different study partner. Although re-assigning was not an option within the experiment, the findings indicate potential for future interfaces to support re-matching study partners as the situation demands. Ideally, such functionality might include the possibility for pairs who discover internal differences along the way (e.g. different pace, level of detail or personal differences) to gracefully disconnect and reconnect with more suitable partners. In order to provide this kind of support at scale, a system would need to "get to know" individual learners in order to estimate their mutual compatibility. Steps in this direction have been taken in the form of subsequent designs, namely X5Learn (see below) and TrueLearn (see D1.3), where part of the goal is to measure an individual's background knowledge, interests and preferences over long periods of time. Naturally, this required a shift of design focus, away from the idea of pair-based user accounts and interfaces, back to more traditional, single-user accounts and interfaces.

The findings indicate that there is demand for new tools and services to facilitate matchmaking among self-directed, lifelong learners. When asked whether they could imagine using a website to find a study partner, most interviewees were excited about the idea. Some suggested that the basic functionality could be similar to a typical dating website, with some obvious modifications. Although building and maintaining a fully-featured social network platform for matchmaking is probably outside of the scope of X5GON, it may nevertheless be possible that some of the outcomes in Year 3 could provide useful stepping stones in this direction. For instance, a future version of the X5Learn dashboard might allow users to publish their TrueLearn profile as a web URL or API endpoint for a trustworthy third-party social network site or app that specialises on matchmaking. There may be potential scope for collaborating with an education startup for this purpose.

As mentioned in D6.1, solutions for matchmaking will likely involve advanced learner modelling, data visualisation, and interface design. Progress in this direction has been made in Year 2 in the form of the TrueLearn learner model (see D1.3). Plans for Year 3 include the possibility to explore the design of dedicated user interfaces for matchmaking, both learner-to-learner and content-to-pair. The potential of TrueLearn for matchmaking lies in the idea of comparing learner profiles to identify commonalities in background knowledge, interests and preferences. Moderate accuracy will be sufficient for this purpose, presumably, since our findings suggest that minor differences can easily be worked around and even lead to stimulating experiences, such as learning by explaining.

Naturally, it is impossible for any matchmaking system (human or otherwise) to take every potentially relevant factor into account when making predictions. This needs to be considered when setting expectations and planning the design and evaluation of future systems. TrueLearn covers only a small subset of potentially relevant factors. As our findings demonstrate, the pair learning experience can depend on many unforeseeable factors, including interpersonal aspects, contextual demands regarding logistics, social and cultural norms, variable schedules and changes in professional and private circumstances. Inevitably, a pair's flexibility to work around various combinations of obstacles will vary from case to case. Therefore, the goal should be to provide the best recommendations possible within the limits of available data, while also providing appropriate fallbacks and exit strategies.

2.5.5 Exit and re-entry strategies

Participants who left the course early mentioned a variety of personal and academic reasons, similar to previous findings in the context of MOOCs [2]. However, whereas MOOCs continue unperturbed when someone leaves, a pair ceases to be a pair when someone decides to leave. Moreover, a pair may discover differences in interest or learning preferences along the way, causing them to go separate paths.

Several participants who wanted to continue after losing their study partner asked the researcher to be re-assigned a different study partner. Although re-assigning was not an option within the constraints of the experiment, the findings indicate potential for future interfaces to support re-matching study partners as the situation demands. Ideally, such functionality might include the possibility for pairs to gracefully disconnect when appropriate and reconnect with more suitable partners. While providing elegant exit strategies is primarily an interaction design challenge, the problem of finding new partners comes back to matchmaking.

2.5.6 Previewing content

Previewing the given course material was deliberately discouraged in this study, in order to keep learners focused on the tasks at hand. Nevertheless, some pairs admitted that they peeked into some of the later tutorials upfront, which required "cheating" the interface. This finding corroborates the notion that some learners care strongly about the ability to preview content. Future designs should take this into account and explore appropriate mechanisms for learners to look ahead.

In addition to the given course material, pairs occasionally used other information sources (not purely OER) which they usually retrieved through standard search engines. As reported in the interviews, typically the information needs revolved around certain error messages or domain-specific terms. Data about content queries and any previews were not collected.

2.5.7 Progress tracking

Progress tracking was integrated into the design for two reasons: Firstly, to make it easy for learners to continue where they left off. Secondly, to collect quantitative data about their progress.

Since the course consisted of a linear sequence of short tutorials, it was decided to represent them as a simple ordered list. An additional constraint was imposed which encouraged users to complete a tutorial before starting the next. This simple design worked well in most cases, i.e. when pairs used the interface simultaneously as intended. However, two unexpected behaviours were found, as mentioned above:

1. Peeking ahead: Some pairs or individuals cheated the interface in order to look at upcoming tutorials in advance.
2. Individual use: Some participants used the interface individually, rather than simultaneously with their partner.

Both types of unexpected use resulted in certain parts of the log data becoming blurred, limiting their use for quantitative analysis to some extent. Negative impacts on the user experience were unlikely.

In order to measure progress and activities more accurately, future designs for progress tracking in pairs should likely involve a combination of personal accounts with pair activity. Using personal accounts would allow individuals to keep their data beyond the lifetime of a particular pairing. By accumulating a long-term, personal learning profile, a recommender system could then use these data to make increasingly accurate suggestions of content and further opportunities for collaboration.

2.5.8 Peer involvement

In light of these findings, it became clear that future designs should anticipate the possibility that learners' study preferences and needs vary by situation and between individuals. Particularly, while some learners may find the idea of a shared user account motivating, our findings suggest that individual user accounts may be preferable in some cases, as they allow for personalisation and grant the individual more flexibility to shape their own learning journey. Ideally, future interfaces should support learner's ability to (a) study at their own pace and level of detail, (b) pair up spontaneously as the situation affords, (c) disconnect gracefully and find a new partner when appropriate. Lifelong learning should not imply lifelong pairing.

As a consequence, the question remains how to support peer motivation using individual learner accounts. This challenge has previously been addressed in the context of MOOCs [3, 7] but prior research is scarce regarding how to support peer motivation in a lifelong learning context where learners study independently of fixed course structures.

2.6 Conclusion

Our in-the-wild experiment with 72 pairs of self-directed learners showed that augmenting online resources through offline discussions can be extremely motivating and satisfying for some learners. Notable benefits include learning through explaining, joint problem solving and peer motivation. The study also investigated whether making an explicit "pact" increases completion rates and whether the effect differs between face-to-face and remote pairs. A two-way ANOVA test showed no significant effect on completion rates, suggesting that explicit pact-making adds little value in addition to the strong implicit commitments and communication that come naturally with being in a pair. Qualitative analysis of interviews and system logs resulted in recommendations for the design of future interfaces and AI-supported matchmaking.

In order to provide support for matchmaking at scale, a system would need to "get to know" learners individually in order to predict their compatibility with each other. Steps in this direction have been taken in the form of subsequent designs, namely X5Learn (see below) and TrueLearn (see D1.3), where part of the goal is to collect data about an individual's background knowledge, interests and preferences over long periods of time. This goal required shifting the focus of design away from pair-oriented user interfaces towards more traditional, single-user interfaces that are optimised for personalisation. Personalised learning, as we understand it, should not imply that learning happens in isolation. Rather, the aim of personalisation should be to help the learner connect with real-world opportunities for active and social learning that fit the person's needs.

3 X5Learn

The following sections describe the ongoing design, development and evaluation of the X5Learn website. X5Learn is an online learning application that encourages users to explore and navigate an ever-growing landscape of OER. In order to help learners find engaging content and efficient pathways among myriads of documents, we have devised a solution that combines novel approaches to document visualisation with tailored machine-learning algorithms for user modelling, content recommendation, knowledge tracking and quality assurance.

Primarily designed as a web application for self-directed lifelong learning, X5Learn aims to support active, connected, personalised learning with OER. For this purpose, it combines content-centric features, such as search and recommendation, with learner-centric features, such as note taking and support for personalised learning journeys. It differentiates itself from traditional OER tools and personal learning environment by allowing the learner to interact with meaningful *fragments* and *topics* within OERs, in addition to entire documents and collections thereof. This focus on small units aims to bring clarity to the learner at a fine-grained level of detail and allow them to *mix and match*, e.g. to find alternative examples of a particular topic in different resources, different modalities or different cultural contexts. For this purpose, X5Learn capitalises on a new approach that combines topic extraction (see D1.3) with efficient preview mechanisms, allowing the user to explore, discover, and accumulate topics and resources and build a personal learning profile.

X5Learn, developed at UCL and hosted by Pošta Slovenije, aims to be a flagship interface for X5GON, designed in response to prior findings and tested with real users in the wild. Furthermore, X5Learn aims to function as a testbed for future interfaces and algorithms. The joint design effort between HCI and AI researchers has led to a series of user interface prototypes, a production-ready web application and a better understanding of OER use cases for lifelong learning. Furthermore, demonstrating prototypes to experts at international conferences has produced valuable feedback and opportunities for X5GON to collaborate with charities, NGOs and startups.

While our previous study with pairs focused on learning with a predefined curriculum, X5Learn focuses on the challenges that learners face when constructing their own curriculum, i.e. including the additional challenge of retrieving additional OERs and judging the potential relevance of the retrieved pieces of content.

The primary user group of X5Learn is lifelong learners. The secondary user group comprises educators, charities and NGOs (for the latter see D7.3).

3.1 Objectives

Our design capitalised on prior findings from Year 1 (see D6.1), particularly, our research on supporting learning journeys in the wild [1]. Three main challenges were previously identified: (1) previewing content, (2) reflecting on progress, (3) involving peers.

In the context of X5Learn, these challenges are still considered essential. However, the requirements X5Learn added additional challenges, particularly about discovering OERs. Unlike our previous studies (which relied on small, expert-curated collections of OER) X5Learn allows learners to access the entire X5GON content catalogue, using search and recommendation features. Therefore, X5Learn introduces the need for appropriate search and recommendation algorithms, automatic quality assurance, and interfaces that allow the user to make quick and accurate relevance judgements about presented content.

Functional requirements were defined in order to scope the desired functionality of X5Learn from a user's point of view. The system should allow the user to perform the following tasks: 1. search for OERs by single topic or text phrase (Year 2) and complex topic query (Year 3) 2. explore transparent OER recommendations through iterative query refinement (Year 3) 3. preview OERs of various formats, including video and pdf (Year 2) and audio (Year 3) 4. reflect on their progress in terms of resources used (Year 2) and knowledge gained (Year 3) 5. engage with peers by sharing links to resources (Year 2) and sharing journeys (Year 3)

The above requirements informed (and were informed by) a series of literature reviews and design iterations. The aim of the literature search was (a) to ensure that the design and implementation reached the state-of-the-art in relevant domains and (b) to identify any potential for research contributions.

A set of non-functional requirements and interaction design principles was also compiled. One of the key aspects in this regard was the principle of learnability, i.e. ensuring that the system is quick and easy for beginners to become familiar with, while also ensuring that the person can quickly and easily progress and master the advanced functionality.

3.2 Related work

Below we present a summary of relevant literature, grouped by research theme. A large amount of prior research has revolved around aspects of discoverability of OERs, efficient previewing of documents, and potential for personalisation. Although real-world systems like X5Learn that combine all of these principles in one cohesive application have not been found in prior literature, many studies that individually focused on one of the aforementioned principles are relevant to the design and evaluation of X5Learn and are therefore discussed below.

A diverse range of papers was considered in the review, including technical system proposals, design studies and evaluations of live systems. Discoverability and preview are introduced first, drawing on a variety of studies that may focus on either learners or teachers/lectures. Personalisation is discussed from a purely learner-centric perspective.

3.2.1 Discoverability of OER

The challenge of discoverability concerns the question how users can learn about the availability of material [10]. Dating back to the very beginnings of the OER movement, it is still considered an open research problem. Many user-centred studies in this area have focused on the need of teachers and lecturers to find material they can use in class. Although some of the issues (such as intellectual property information) are more relevant to teachers than to learners, other issues apply equally to learners. For instance, a recent evaluation of two OER meta-search tools with eight adult users identified considerable potential for improving the user experience, including (1) the need for increased transparency in search tools, including clarity about how certain quality indicators such as star ratings were calculated, (2) the importance of intuitive facets, such as content types and clear topic hierarchies, (3) the need for consistent, detailed metadata and summaries that can help the user in selecting materials [11]. One of the reasons why these findings are particularly relevant for X5Learn is that the described issues were partly due to the fact that the search results were sourced from different repositories, as is the case with X5GON.

Another approach to discoverability comes from an interaction design perspective. Cortinovis et al [12] recently demonstrated a series of prototypical interfaces that depart from traditional search and recommendation mechanisms, suggesting a blended approach instead. One of the key features was that the user could dynamically "expand" a selected search result by requesting similar items. Moreover, the user could control the degree of similarity as they navigated through the links. Test users (educators) were reportedly intrigued by this method, as it provided an alternative way to explore and make sense of large result sets. One limitation of this research is that it relied on one-off user testing sessions with a new prototype, suggesting a possible novelty effect. Another is that the evaluation focused on semi-realistic evaluations with educators. Empirical evaluations with learners in the wild or comparisons with other OER search tools were not found. Nevertheless, the study demonstrates that there is scope for design to enable new and surprising entry-points for users to explore OER repositories. The design followed a previously observed trend whereby modern interfaces are increasingly blurring the boundary between search and recommendations as systems become smarter, more interactive, social and personalised [13].

While purely learner-centric evaluations of OER discovery tools are scarce, some generic OER tools and platforms have been found to embrace learners as a secondary user group. In a review of eight web-based OER platforms, [14] found basic search functionality to be a common feature, including the ability to filter results by metadata, such as data, type of content, author and keyword annotations. Support for personalisation and peer collaboration was found in less than half of the systems under review. None of them provided content recommendations for learners.

3.2.2 Efficient previews of (OER) documents

Efficient previews are considered instrumental in many Information Retrieval (IR) tasks where users need to make fast and accurate relevance judgements about large quantities of textual information. Early computational approaches focused on generating static abstracts of texts [15]. The potential to reduce text dynamically in response to a user's information need, received attention in the early days of modern search engines. Query-biased summaries were intended to improve both the speed and accuracy of user relevance judgements and also "help users to more easily identify the relevant pieces of information that are contained in each document" [16]. Query-biased wordclouds [17] followed a similar intention with a stronger emphasis on fast visual perception by reducing the amount of text and varying the text size and colour according to relevance.

Abstract graphical representations of text content have also been explored in order to help the user judge the potential quality of the retrieved documents [18]. The above example capitalises segmenting the content into multi-paragraph chunks, similar to the approach with X5Learn.

Different types of previews can be appropriate depending on the type of content. While resources that contain images are frequently displayed as thumbnails, research suggests that the combination of text and thumbnails enables more accurate relevance judgements than either text or thumbnails [19]. Methods exist for automatically summarising videos by extracting thumbnails [20], including approaches that respond to the user's query [21].

In addition to summarising the content itself, it can also be desirable to augment the search experience with structural preview information. Providing an interactive display of result distributions across multiple search result pages resulted in users considering more mid- to low-ranking results in their selection, indicating that well-designed previews can put the user in control and make the search experience less dependent on ranking algorithms whose predictions may not always be perfect [22].

Within the realm of OER, Zhao et al [23] demonstrated how a combination of textual and visual previews can considerably accelerate simple lookup tasks in large numbers of educational videos. The solution involved representing the videos as thumbnails, including tagclouds, and highlighting points in the timeline where the transcript matched the search query. Other video-based preview methods that might to have potential applications for OER include automatic detection of lecture slides [23], grids of animated thumbnails [24] and various types of "skimming" using the progress bar to produce previews similar to YouTube or Netflix ([25, 26], Matejka et al 2013). Skimming was found to be a common strategy among young adult self-directed learners when previewing educational videos [27].

Despite overwhelming evidence highlighting the benefits of efficient previews, it appears that sophisticated preview features have yet to become mainstream in real-world OER search tools. In cases where search results contain only basic metadata, users lamented a lack of salient information that they would need in order to make confident judgements of relevance [11]. It seems reasonable to assume that, least with standard OER formats, such as text and video, these shortcomings could be addressed with reasonable effort by adopting some of the well-researched techniques described above. Little research was found regarding how learners can preview less common OER formats (such as interactive simulations, games, courses, projects, hackathons and meetups). Some of these types of resources are notoriously difficult for learners to discover and preview [28], indicating a potential need for novel techniques. In order to generate previews of interactive simulations, Ma et al [29] explored the use of heatmaps and collage to summarise relevant activity in the graphical user interface. Focusing on desktop-based, educational mathematics simulations, this study was one of very few examples addressing interactive OER. Moreover, efficient preview solutions for users with disabilities are rare [30].

To summarise, efficient previews have been found to improve the speed and/or accuracy of user relevance judgements across a variety of content types and information retrieval tasks. However, much of this potential appears to be unused in OER. Moreover, while many of the above techniques are specific to particular content types (e.g. video), there is a scarcity of generic solutions for users to efficiently preview OERs of diverse types.

3.2.3 Personalised learning with OER

Efforts towards personalisation have received growing interest in MOOC research, whereby recent studies have considered the possibility of maintaining a MOOC-independent learner profile [31].

Outside the MOOC context, prior research is scarce regarding personalised learning with OER. One of the few relevant papers proposed the high-level design of a framework for learner-centric OER recommendation [32]. On a conceptual level, the framework suggests implications for filtering, user preferences and personalisation based on interests, knowledge levels and learning objectives. Its suggestions for design were informed by prior (non-OER specific) perspectives on learning with personalized recommender systems [33]. Work-in-progress was announced towards implementing the framework as a real system that can be tested with users.

Outside of OER, a survey of the state of the art in research on recommender systems for Technology Enhanced Learning [34] identified collaborative filtering as one of the most common techniques and also highlighted a variety of alternative approaches, including the potential for recommender systems to consider contextual information and educational constraints. While a large part of the research in this area focuses on content recommendations, recent work has also looked into the possibility of recommending other aspects, such as learning activities that foster communication and metacognition.

3.3 Research Aims

The literature review revealed a number of opportunities for WP6 to produce research contributions in the areas of discoverability of OERs, efficient previewing of documents, and learner's interactions with personalised OER recommendations.

The main research aims for Year 2 focused on efficient previews, particularly:

1. to identify appropriate types of preview for common types of OER
2. to characterise the cognitive benefits of different types of preview for different learner tasks
3. to better understand common criteria and effective strategies that learners use while making relevance judgements about OER

Secondary research goals and objectives were:

1. to generate well-rounded understandings of self-directed learners' information needs in the context of OER
2. to identify appropriate representations and interfaces for allowing learners to see, change, and reflect on their information needs
3. to explore the viability of traditional and blended approaches to search and recommendation for discovering OER
4. to identify under-explored use scenarios for OER

3.4 Iterative design and evaluation in the wild

In order to address the research aims, an iterative, participatory design approach was chosen. Starting from initial requirements (see above), repeated cycles of design, prototyping and one-on-one user testing resulted in a series of prototypes and increasingly rich understandings of users' needs, habits and expectations. Participants were convenience-sampled adults from a broad range of demographics, including first-time users and recurrent users.

At the early design stages, the evaluations were designed as open-ended sessions with the main goal to elicit broad feedback and suggestions for improving the key features and overall usability. As the design matured, the focus shifted to more specific research questions, hypothesis-driven protocols, and considering participants' background knowledge. A lab experiment investigating the cognitive benefits of different preview methods is in progress at the time of writing this report.

In addition to controlled, face-to-face user testing, an online version of X5Learn has been made available to a gradually increasing number of test participants. This method of soft-launching the web application served multiple purposes: Firstly, to elicit feedback from users in-the-wild and secondly, to test the technical robustness of the system under real conditions, including regular features updates. Key stages of the design process are presented and discussed below.

3.4.1 Early sketches and mockups

Figure 7 shows the first version of the interface design. It was implemented as a semi-interactive click dummy, with some interactive areas and some areas mocked up. The goal was to evaluate our first educated guess of elements to be included on the main screen. One of the main assumptions behind this version was that the most frequently used functions should be available directly on the main screen, while secondary functions are available through menus. The layout was split into 4 main areas..

1. Progress visualisation (upper half) using a board-game metaphor with the learner's current position indicated by the person icon. Full yellow circles represent completed activities, empty circles represent unstarted/recommended activities. Lines represent transitions from one activity to another (past or recommended).

2. Kanban-board (lower half, left and centre) for the user to arrange OER-based activities (white cards) in three categories: Done, Doing, Planned. Activities could be moved from one category to another by marking them as started or finished. The user could comment on a particular resource or a particular activity.

3. Bookmarks section, allowing the learner to temporarily store OERs that could be turned into activities (using the "Plan activity" button which causes to move the card into the column labelled "Doing")

4. Chatbot interface, allowing the learner to ask an AI agent for recommendations.

This initial design was closely based on the notion of the learning journey as a set of projected pathways (graph of documents) as suggested by [35]. Experts and test users who evaluated the

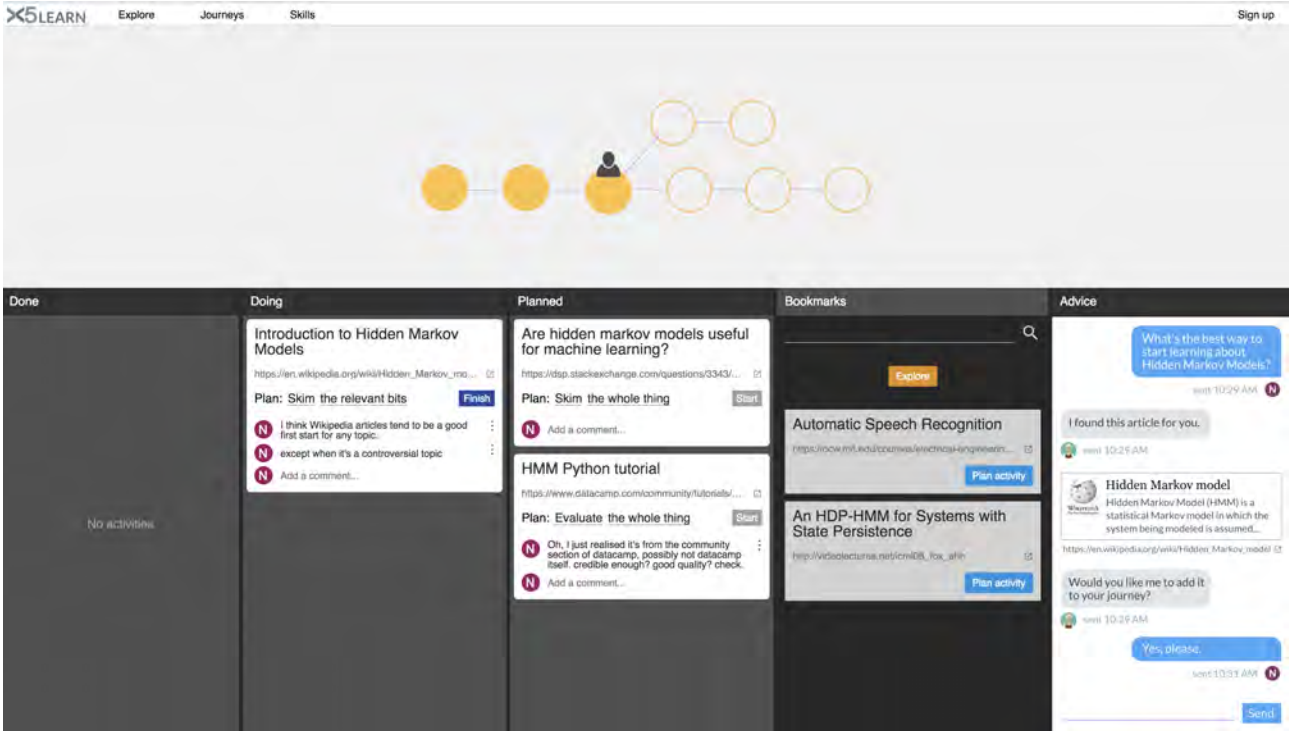

Figure 7: Early interface design

design in interviews were intrigued by the board-game metaphor. All found it intuitive to understand. However, the other interface elements were found less self-explanatory, namely the Kanban board, the Bookmarks and the Chat.

The conceptual distinction between resources and activities was considered intuitive by some, but not all users. Moreover, the fact that the interface represented a relatively advanced stage in the learning journey was confusing for many users who did not understand how the interface got to that state. In light of this valuable feedback, it was therefore decided to reset the focus on first-time users and design an easy-to-use start page, see Figure 8.

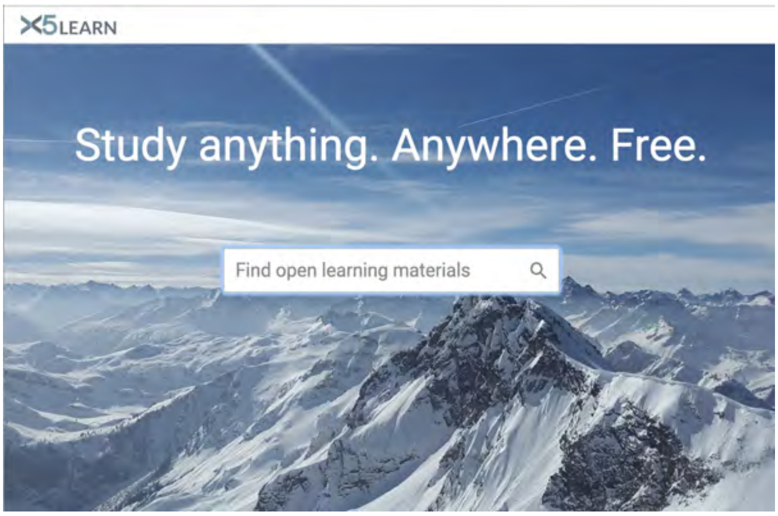

Figure 8: Minimalsitic start page design

Most test users said that they appreciated the calm and pleasant background image. One test user described the overall appearance as simple and balanced, arguing that it radiated a sense of openness and made them feel in a good mood for studying.

Two experts suggested that the start page should be populated with some attractive, high-quality, featured content. The rationale was new users should be able to:

1. see at a glance what the site offers
2. start exploring immediately with one click, rather than having to think of a search term which some might find tedious.

3.4.2 Previewing content

The prototype shown in Figure 9 explored the possibility to explore search results using thumbnail images. By hovering over a card, the user could initiate a slideshow of thumbnails, similar to hovering over a card on YouTube.

The card design was received very positively. Test users described the visual exploration approach as curiosity-inspiring and praised the ability to take a closer look at content without leaving the page, simply by hovering.

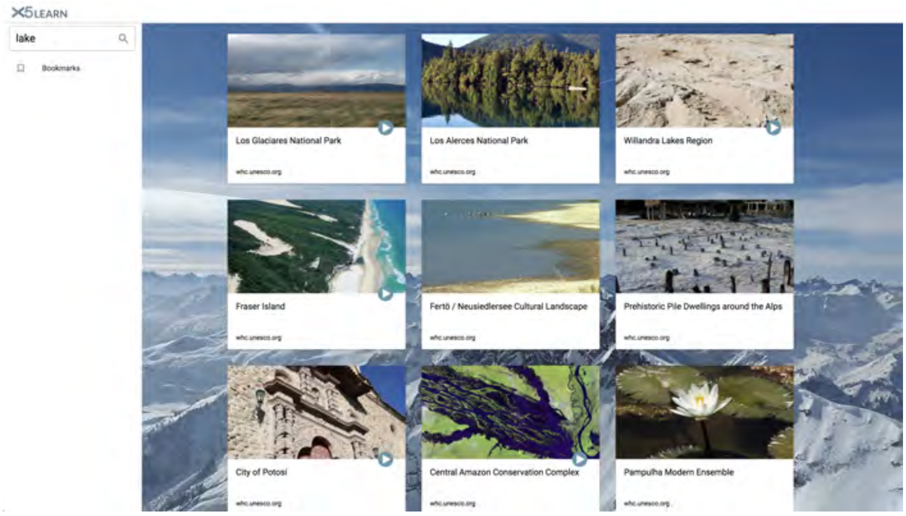

Figure 9: Early card-based design with image thumbnails showing UNESCO World Heritage Sites

In order to access video content, it was decided to embed a video player in a standard modal, as cards provide too little space, see Figure 10.

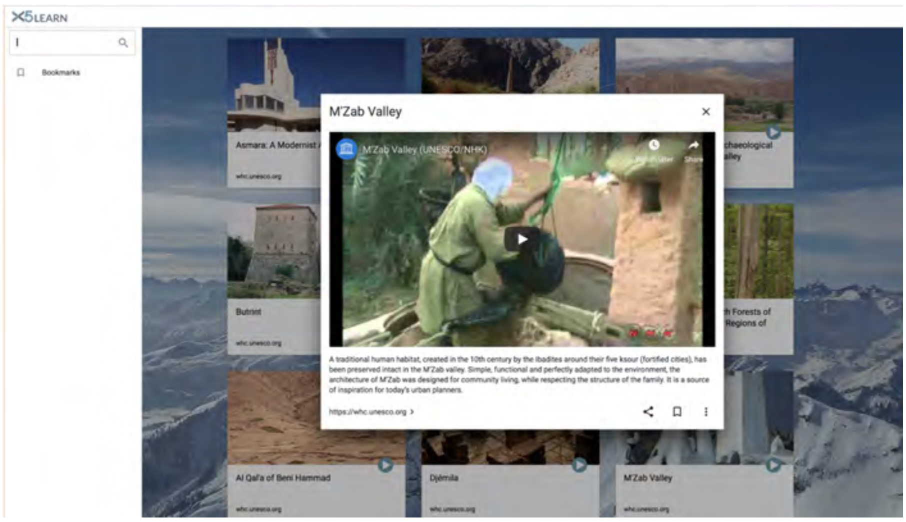

Figure 10: Clicking on a card opens the main content in a modal

The simple and familiar design was very popular among test users. The designers felt confident in the current solution with regard to discoverability and preview. At least for picture-heavy content it appeared to work well. Consequently, the design was tested with a different set of data, namely text-heavy lecture recordings sourced from www.videolectures.net. Figure 11 shows the first 9 search results for "Music" on videolectures.net, as presented through X5Learn's card layout.

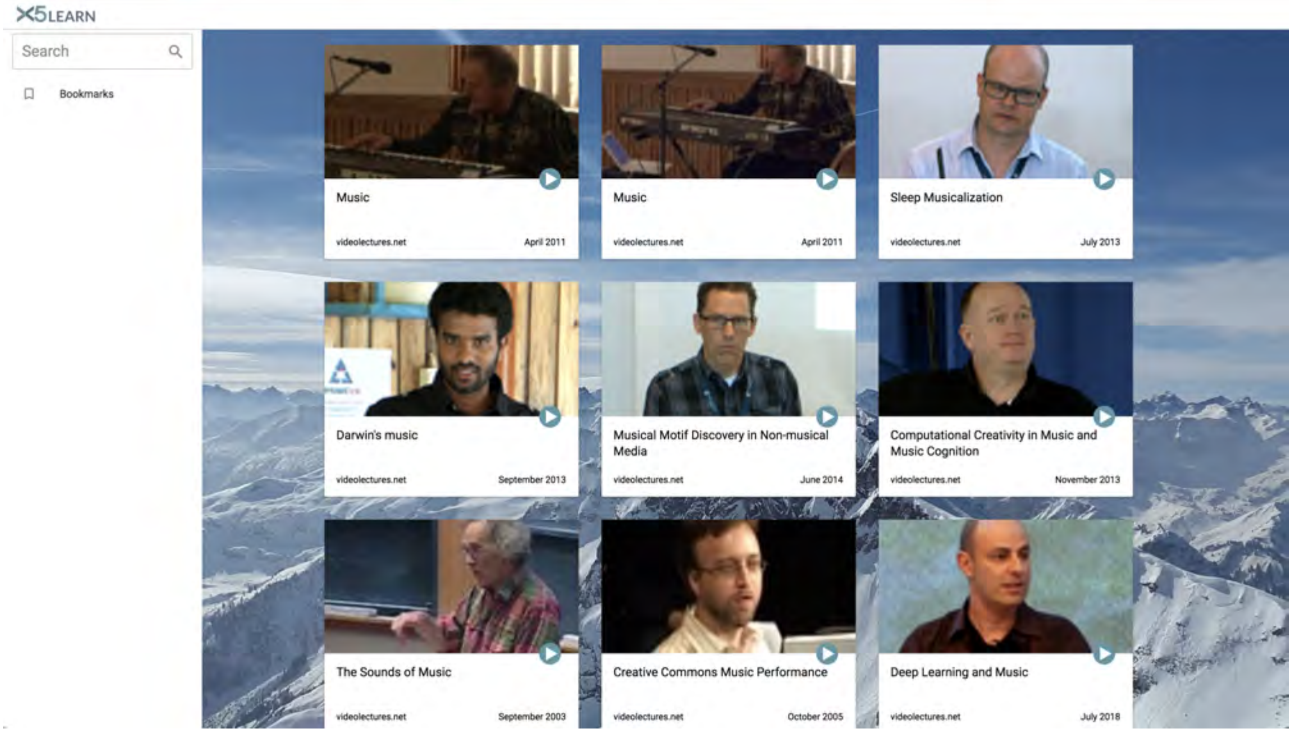

Figure 11: Many talking heads - search results for "Music" on www.videolectures.net as represented by thumbnails on cards.

When test users hovered over a speaker's photo, the interface revealed a series of other stills of the same speaker. This was unanimously described as rather uninformative, suggesting that multiple thumbnails per resource added little value with regard to users' ability to judge the potential relevance of lecture-type OERs.

Realising that text-heavy material did not benefit substantially from extended graphical summaries, the hover mechanism was redesigned. The goal was to replicate the joyfulness and fluidity of using the hover feature UNESCO version but with text data instead of image data. The resulting new version introduced a timeline widget for every resource, allowing the user to "skim" through the document by hovering over the timeline. A context menu was designed to display the main topics at the corresponding position under the mouse pointer (see Figure 12).

In order to extract the topics, the documents were pre-processed by dividing the text in to chunks of roughly equal length, similar to TileBars [18]. Each chunk was then analysed using Wikifier (as described in D1.3).

As a side effect, the timeline provided space for highlighting parts of the resource that the user had previously visited. Therefore, the timeline was used as a combined tool for preview and progress tracking. Users responded very positively to the skimming feature. One user compared the experience with speed reading. In the example shown in Figure 12, the user had read the second chapter (blue area on the left) of an e-book. Hovering over the timeline shows a summary of key topics in the corresponding section of the text (or transcript in case of audio/video).

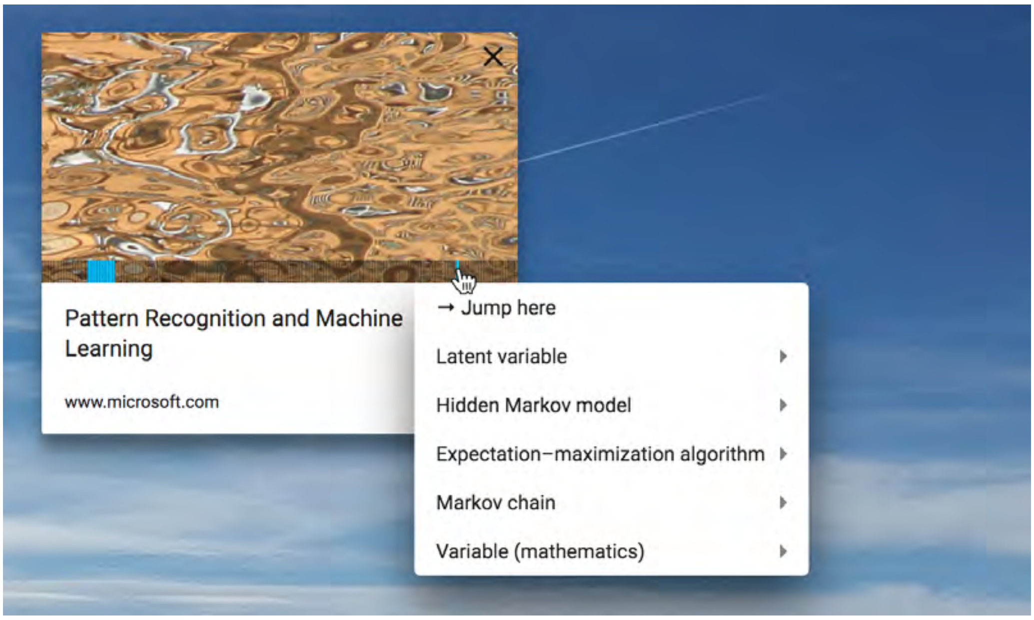

Figure 12: Context menu showing extracted topics

In order to let the user interact with the extracted topics, Wikipedia definitions were provided. An additional button allowed the user to annotate a topic with “I know this well”. This function constituted the first of a series of designs towards supporting personalisation using Wikipedia-topic based knowledge tracking (see Figure 13).

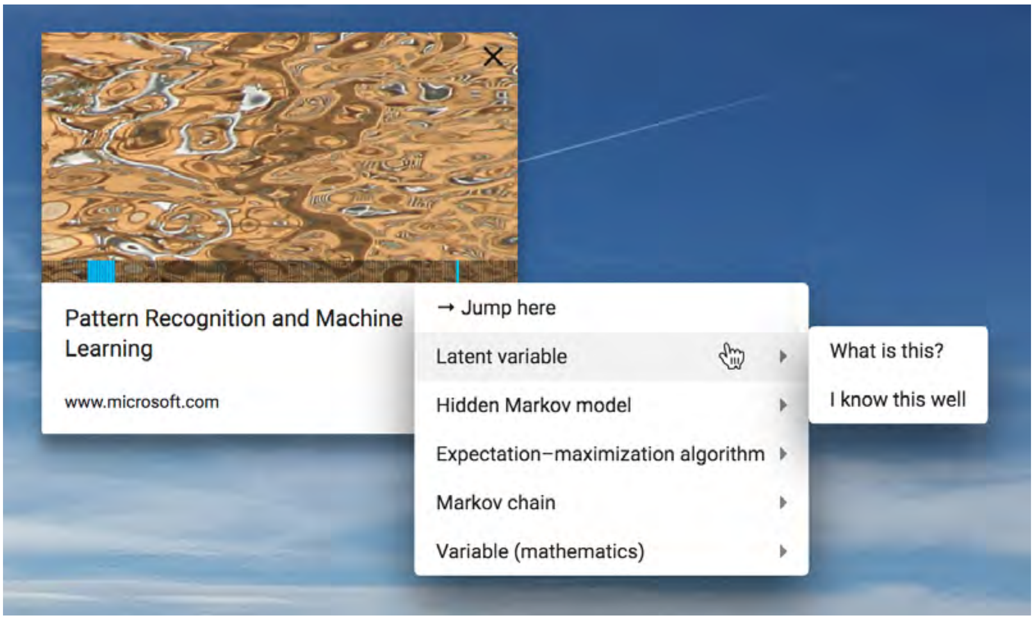

Figure 13: For each identified topic, the menu allows the user to ask for a definition or indicate prior knowledge.

3.4.3 Ranked tagclouds

The next design iteration explored the suitability of text-based previews instead of thumbnails. Our design is similar to the wordcloud design by Williams et al [17] but differs in that the items are ranked vertically and left-aligned for faster readability (see Figure 14). The items represent the five main topics of the corresponding document, with the most frequently occurring topic at the top. The name “ranked tagcloud” was chosen for this design. A small-scale, within-subjects comparison between thumbnails and ranked tagclouds led to the following conclusions:

1. Ranked tagclouds are consistently easy to read and informative, independent of the format, modality or structure of the OER.
2. While thumbnails can potentially be distracting, superficial or misleading, ranked tagclouds aim to represet the content in an objective, unemotional way, which may sometimes be preferable.
3. In cases where an OER lacks a descriptive title (e.g. “Lecture 15”), the ranked tagcloud can partly compensate by covering in the information gap. Theoretically, this might occasionally prevent good OERs from being unfairly overlooked based on unlucky titles.

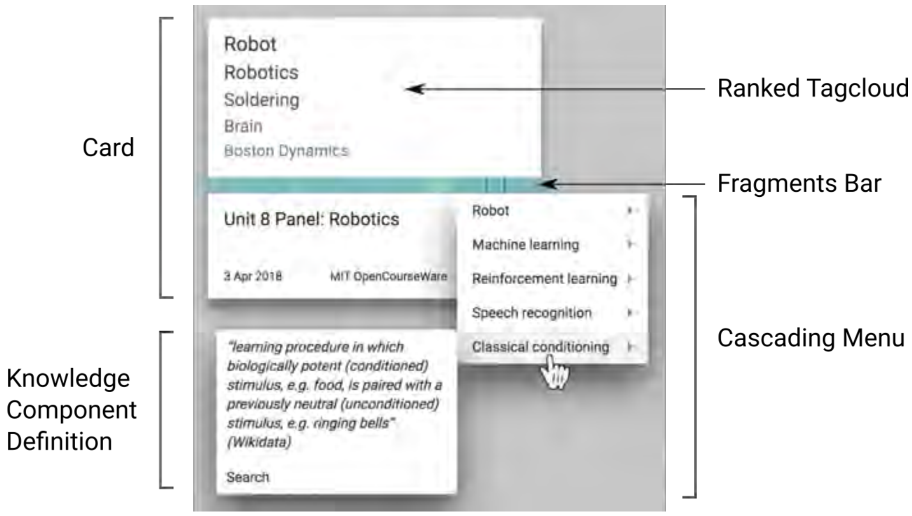

Figure 14: OER card with ranked tagcloud and menu fully expanded.

3.4.4 Different previews for different users

Ranked tagclouds are the simplest type of visual summary and were therefore chosen as the default preview mode, in order to avoid overwhelming new users with too much detail. Two further types of visual summary were devised in response to advanced users desiring more detailed information. The first type ("Compact") is similar to wordclouds except that the font size remains constant and a every topic is associated with a "bubble" whose size represents the frequency of a topic in the resource (see Figure 15).

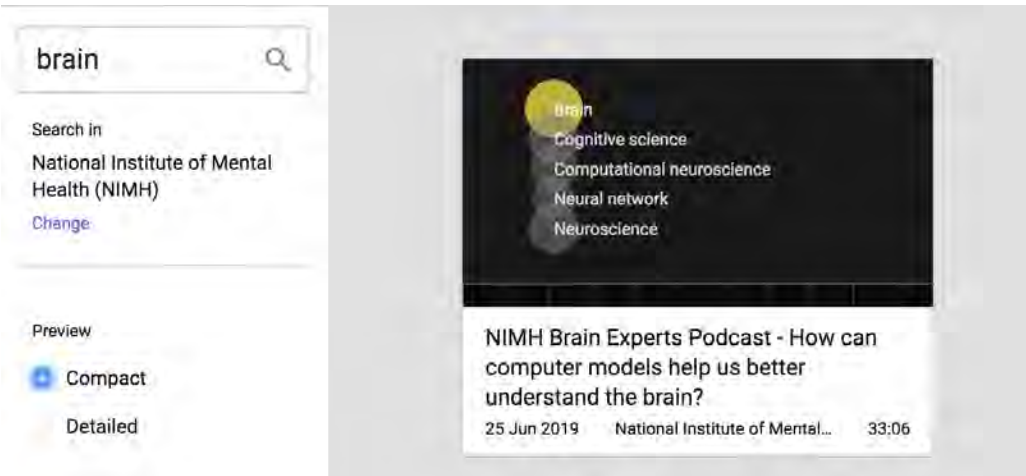

Figure 15: Compact preview mode

The second type ("Detailed") separates the bubble into small dots that indicate points in the resource where the topic is mentioned (see Figure 16). This mode allowed advanced users to quickly eyeball the distribution and co-occurrence of topics in a resource. As expected, this mode tended to be preferred by advanced users. Strong preferences were also expressed by users with experience in reading spectrograms or sheet music.

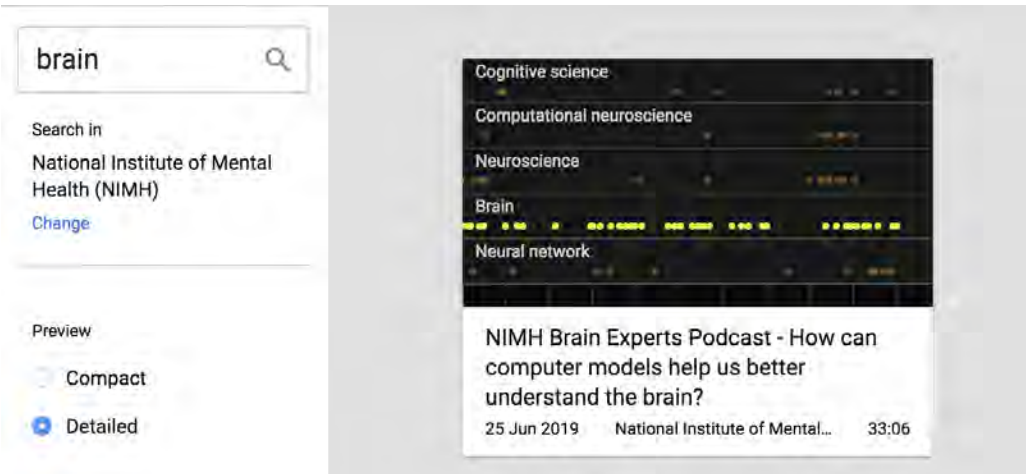

Figure 16: Detailed preview mode

3.4.5 Consistent gestures

A familiar set of mouse gestures are used consistently throughout the interface, in order to provide consistency, predictability and learnability for the user. With regard to topics, chunks and a mentions, the following principles apply:

1. Topics that match the search query are highlighted in colour
2. Hovering over a thing (e.g. a topic, a chunk or a mention) shows information about that thing.
3. Clicking on a thing (e.g. a topic, a chunk or a mention) navigates to the corresponding point in the resource.

These rules were informed by a well-established convention in web browsers that hovering over a link opens a tooltip, whereas clicking results in navigation. All preview modes and menus adhere to these principles, as shown in Figures 17, 18 and 19.

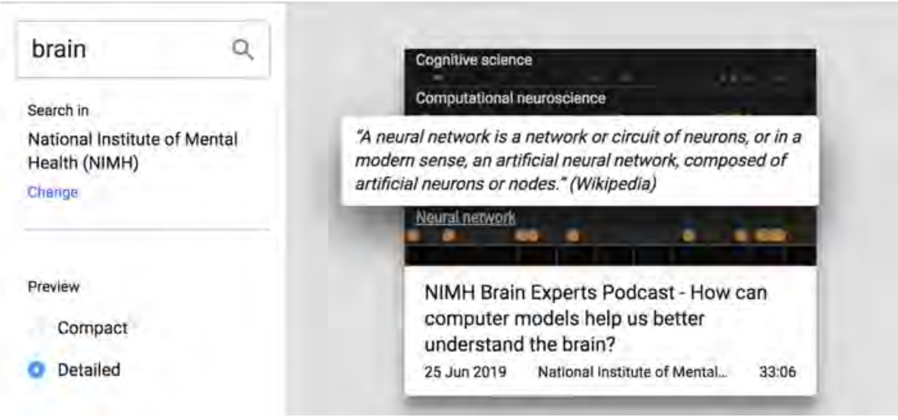

Figure 17: Hovering over a topic (e.g. "Neural network") shows its definition.

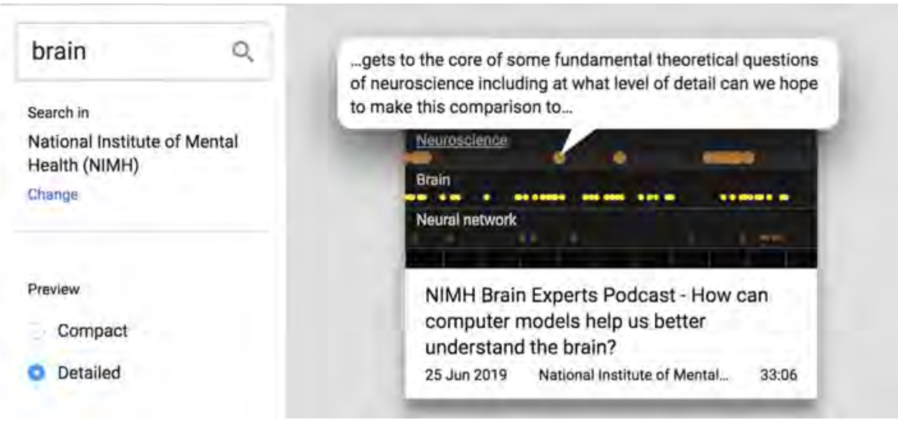

Figure 18: Hovering over a mention (e.g. of the topic "Neuroscience") shows what the text says about neuroscience at that position.

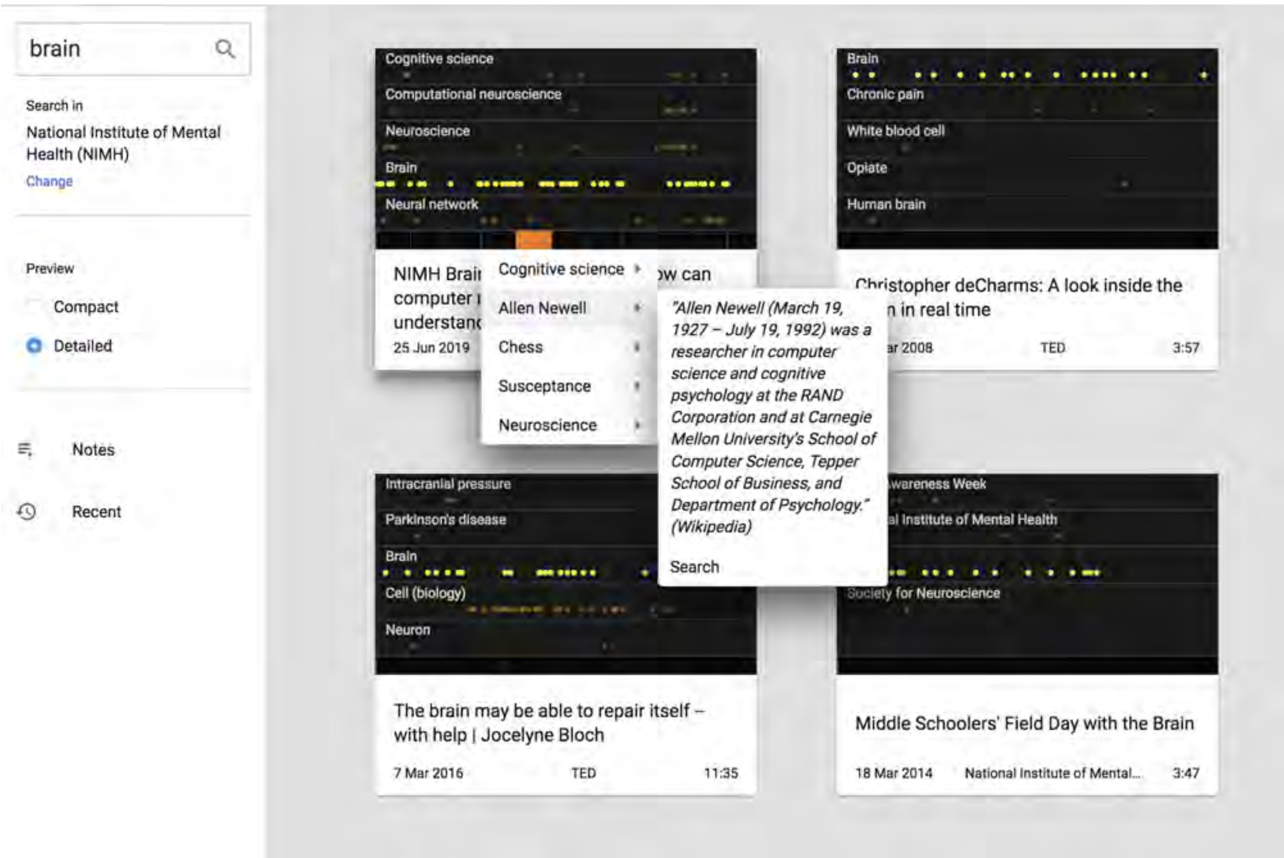

Figure 19: Hovering over a topic in a menu opens its definition.

3.4.6 Learner scenarios to simulate progress

Allowing the user to reflect on their progress was considered a key design goal for X5Learn. Below we outline the pivotal stages in the design of progress tracking. While the design of (a) previews and (b) progress tracking happened more or less in parallel, the two are reported separately for the sake of clarity.

Early on it was decided to ground the design in realistic use scenarios. Unlike previews which can easily be repeated during user testing, progress tracking leaves irreversible traces. Therefore, in order to allow rigorous testing of the interface, A hypothetical learner journey was constructed along with a few artificial ("dummy") user accounts that represented the hypothetical user at different stages in the journey. The ability to easily reset a dummy's state enabled us to repeatedly test the same feature while working on it.

Based on earlier findings, a range of representative learner scenarios were devised in the form of flowcharts (e.g. Figure 20).

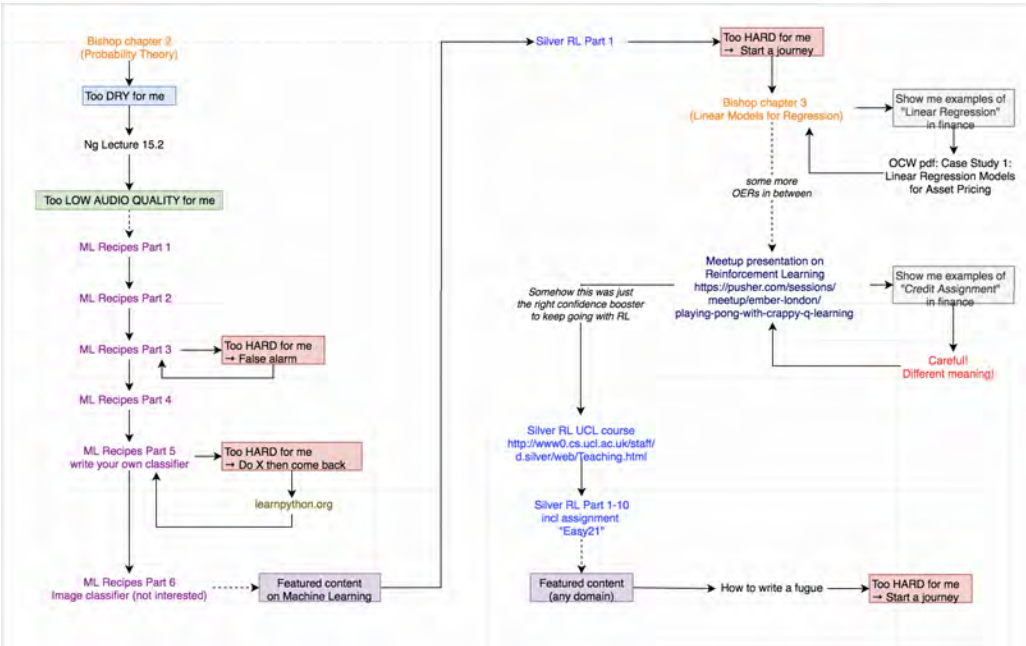

Figure 20: Example of a hypothetical learner scenario

3.4.7 Interfaces for progress tracking

The initial interface to support progress tracking was inspired by Netflix. It provided the user with a vertical "playlist" of resources that the user had started but not finished (Figure 21). Colour highlights (as shown in Figure 13) were later added in order to indicate which fragment(s) of each resource the user had visited, and at what point they had stopped in a previous session. This feature was found especially useful when dealing with large resources, such as hour-long lectures, that users did not necessarily complete in one go.

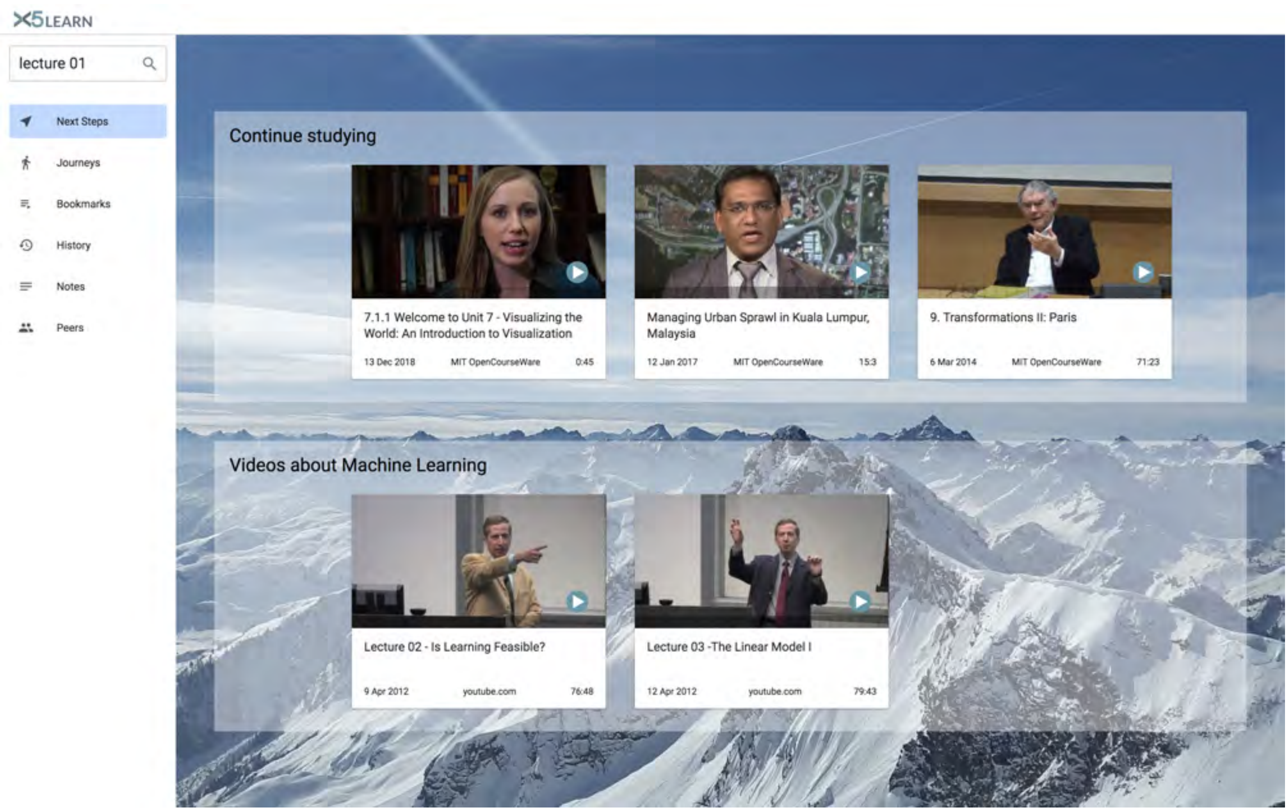

Figure 21: "Continue studying" (top) was implemented as a dynamic playlist, automatically populated by unfinished resources. Another playlist (bottom) is shown for comparison.

The question of how to select one or multiple OERs for later use raised the question of appropriate metaphors. Bookmarks, playlists, shopping trolleys, folders and tags were considered. Figure 21 illustrates a tentative implementation of a feature that allowed the user to create their own playlist for saving OERs. The design was closely modelled after YouTube's playlist metaphor. While this design had the advantage of being familiar, its downsides include a strong emphasis on collection, rather than engagement.

A later design iteration introduced a context-sensitive note-taking feature, based on the conjecture that recalling one's notes could potentially satisfy a learner's need to track and reflect on progress. A rigorous examination of this question was postponed to Year 3.

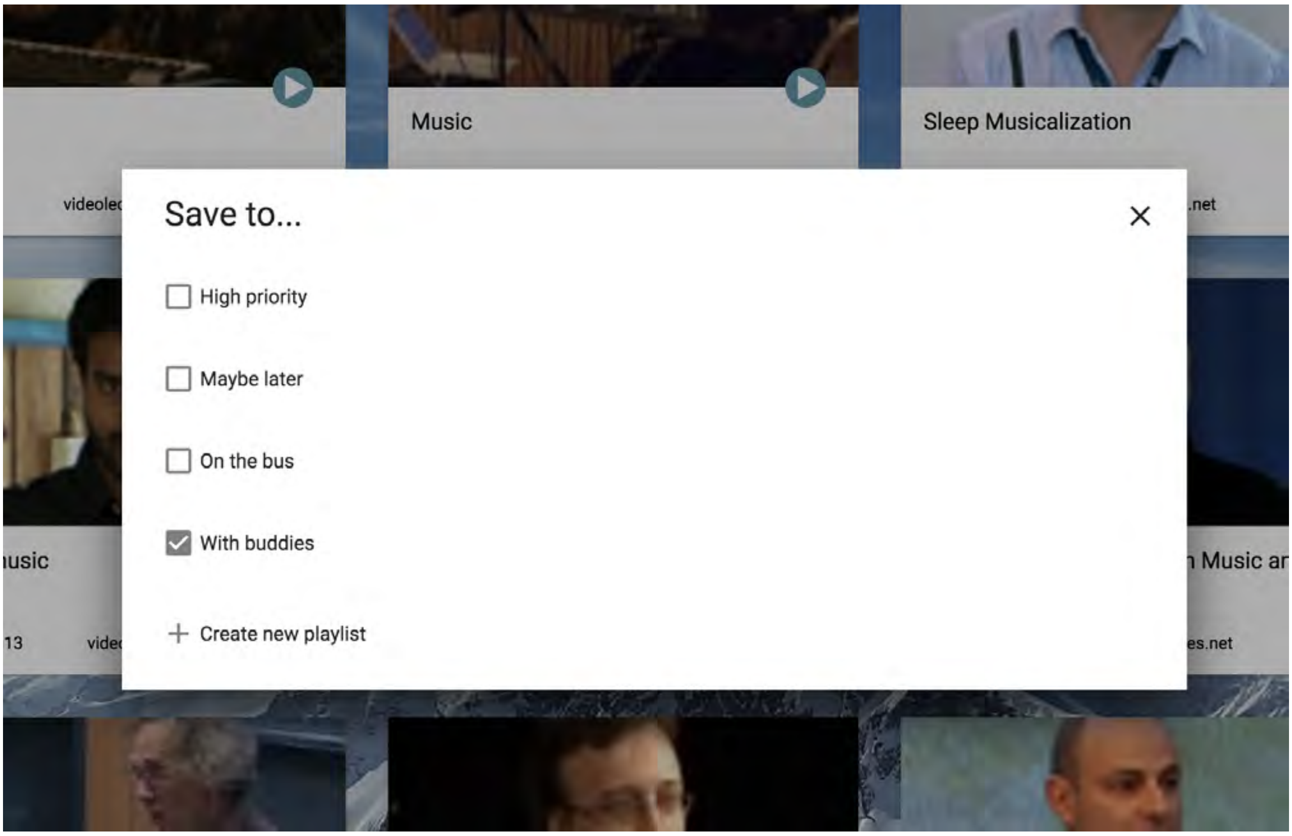

Figure 22: A progress-tracking design based on user-created playlists. In the example, a hypothetical learner has created a "with buddies" playlist, indicating that they intend to reserve certain resources for upcoming peer-learning opportunities.

3.5 Discussion

X5Learn addresses some of the key challenges with OER, predominantly the problem of discoverability. For this purpose, two methods are being explored, namely (a) efficient previews and (b) personalisation.

Efficient and informative previews are a highly desirable feature in OER interfaces, as shown in prior studies [11]. In contrast to previous systems, X5Learn aims to provide a range of preview methods for a variety of OER formats, including text, audio and video resources. Through iterative in-the-wild design and evaluation, several promising interfaces have been devised that were found to be highly engaging to use. These are already available online. More research is needed in order to characterise the cognitive benefits of different preview methods for various learning tasks. Particularly, new users should be provided with the most intuitive preview method for learning tasks that are common among this population. For this purpose, a clearer understanding of the needs of users at all stages (beginner to expert) will benefit the ease-of-use and learnability of the system.

Personalisation is a nascent area of research in the specific context of learning with OER [32]. Previous work aiming to apply personalisation to online courses [35] or lecturers [12] is of limited relevance for the aims of X5Learn where the focus is on cross-modal, cross-site learning. In this context, X5Learn is expected to break new ground by reconciling state-of-the-art recommendation technology with our comprehensive understanding of the needs of self-directed learners using OER.

3.6 Future work

In Year 3 we plan to extend the feature set of X5Learn. Particularly, it is expected that users will be able to:

1. search for OERs by complex topic query
2. explore transparent OER recommendations through iterative query refinement
3. preview OERs of various formats, including websites and and audio
4. reflect on their progress in terms of knowledge gained
5. engage with peers by and sharing journeys

Additional features are planned for matchmaking and multilingual studying.

References

[1] Stefan Kreitmayer, Yvonne Rogers, Emine Yilmaz, and John Shawe-Taylor. Design in the wild: interfacing the oer learning journey. In *Proc. of BCS HCI*, 2018.

[2] Christian Gütl, Rocael Hernández Rizzardini, Vanessa Chang, and Miguel Morales. Attrition in mooc: Lessons learned from drop-out students. In *International workshop on learning technology for education in cloud*, pages 37–48. Springer, 2014.

[3] Yasmine Kotturi, Chinmay E. Kulkarni, Michael S. Bernstein, and Scott Klemmer. Structure and messaging techniques for online peer learning systems that increase stickiness. In *Proceedings of the Second (2015) ACM Conference on Learning @ Scale - L@S '15*, pages 31–38, New York, New York, USA, 2015. ACM Press.

[4] Carrie Anne Platt and Nan Yu. Virtually the Same?: Student Perceptions of the Equivalence of Online Classes to Face-to-Face Classes. *MERLOT Journal of Online Learning and Teaching*, 10(3), 2014.

[5] Carlos Delgado Kloos, Pedro J. Muñoz-Merino, Carlos Alario-Hoyos, Iria Estévez Ayres, and Carmen Fernández-Panadero. Mixing and blending MOOC Technologies with face-to-face pedagogies. In *IEEE Global Engineering Education Conference, EDUCON*, volume 2015-April, pages 967–971. IEEE Computer Society, apr 2015.

[6] Colin Madland and Griff Richards. Enhancing student-student online interaction: Exploring the study buddy peer review activity. *International Review of Research in Open and Distance Learning*, 17(3):158–175, 2016.

[7] Michael Yeomans and Justin Reich. Planning prompts increase and forecast course completion in massive open online courses. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference on - LAK '17*, pages 464–473, New York, New York, USA, 2017. ACM Press.

[8] Michael A Madaio, Amy Ogan, and Justine Cassell. The effect of friendship and tutoring roles on reciprocal peer tutoring strategies. In *International Conference on Intelligent Tutoring Systems*, pages 423–429. Springer, 2016.

[9] Nachiappan Nagappan, Laurie Williams, Eric Wiebe, Carol Miller, Suzanne Balik, Miriam Ferzli, and Julie Petlick. Pair Learning: With an Eye Toward Future Success. pages 185–198. Springer, Berlin, Heidelberg, 2003.

[10] Jia Frydenberg, Gary Matkin, and Distance Learning Center. Open textbooks: Why? what? how? when. *William and Flora Hewlett Foundation*, pages 0–33, 2007.

[11] Talea Anderson and Chelsea Leachman. Strategies for Supporting OER Adoption through Faculty and Instructor Use of a Federated Search Tool. *Journal of Librarianship and Scholarly Communication*, 7(1), feb 2019.

[12] Renato Cortinovis, Alexander Mikroyannidis, John Domingue, Paul Mulholland, and Robert Farrow. Supporting the discoverability of open educational resources.

[13] Ed H. Chi. Blurring of the Boundary Between Interactive Search and Recommendation. pages 2–2. Association for Computing Machinery (ACM), 2015.

[14] Benneaser John, V Thavavel, Jayakumar Jayaraj, A Muthukumar, and Poornaselvan Kittu Jeevanandam. Comparative analysis of current methods in searching open education content repositories. *TOJSAT*, 6(2):21–29, 2016.

[15] J. E. Rush, R. Salvador, and A. Zamora. Automatic abstracting and indexing. II. Production of indicative abstracts by application of contextual inference and syntactic coherence criteria. *Journal of the American Society for Information Science*, 22(4):260–274, 1971.

[16] Anastasios Tombros and Mark Sanderson. Advantages of query biased summaries in information retrieval. In *SIGIR*, volume 98, pages 2–10, 1998.

[17] Jennifer Williams, Sharon Tam, and Wade Shen. Finding "Good Enough": A Task-Based Evaluation of Query Biased Summarization for Cross Language Information Retrieval. Technical report.

[18] Marti A Hearst. TileBars: Visualization of Term Distribution Information in Full Text Information Access. Technical report, 1995.

[19] Susan Dziadosz and Raman Chandrasekar. Do thumbnail previews help users make better relevance decisions about web search results? In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '02*, page 365, New York, New York, USA, 2002. ACM Press.

[20] Yitian Yuan, Tao Mei, Peng Cui, and Wenwu Zhu. Video Summarization by Learning Deep Side Semantic Embedding. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(1):226–237, jan 2019.

[21] Arun Balajee Vasudevan, Michael Gygli, Anna Volokitin, and Luc Van Gool. Query-adaptive video summarization via quality-aware relevance estimation. In *MM 2017 - Proceedings of the 2017 ACM Multimedia Conference*, pages 582–590. Association for Computing Machinery, Inc, oct 2017.

[22] Pernilla Qvarfordt, Gene Golovchinsky, Tony Dunnigan, and Elena Agapie. *Looking Ahead: Query Preview in Exploratory Search.* 2013.

[23] Baoquan Zhao, Shujin Lin, Xin Qi, Ruomei Wang, and Xiaonan Luo. A novel approach to automatic detection of presentation slides in educational videos. *Neural Computing and Applications*, 29(5):1369–1382, mar 2018.

[24] Dan Jackson, James Nicholson, Gerrit Stoeckigt, Rebecca Wrobel, Anja Thieme, and Patrick Olivier. Panopticon: A Parallel Video Overview System. 2013.

[25] Michael G. Christel, Michael A. Smith, C. Roy Taylor, and David B. Winkler. Evolving video skims into useful multimedia abstractions. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '98*, pages 171–178, New York, New York, USA, 1998. ACM Press.

[26] Justin Matejka, Tovi Grossman, and George Fitzmaurice. Swifter: improved online video scrubbing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1159–1168. ACM, 2013.

[27] Cliff Loke, Schubert Foo, and Shaheen Majid. Video seeking behavior of young adults for self directed learning. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 10647 LNCS, pages 314–324. Springer Verlag, 2017.

[28] John D. Shank. *Interactive open educational resources : a guide to finding, choosing, and using what's out there to transform college teaching.*

[29] Chao Ma, Xiangping Chen, Yongsheng Rao, and Mouguang Lin. Generating summarized preview for education resource based on exploring and comparing guis. In *SEKE*, pages 141–146, 2016.

[30] Rosa Navarrete and Sergio Luján-Mora. OER-based learning and people with disabilities. In *Proceedings of 2015 International Conference on Interactive Collaborative and Blended Learning, ICBL 2015*, pages 25–34. Institute of Electrical and Electronics Engineers Inc., jan 2016.

[31] S. Assami, N. Daoudi, and R. Ajhoun. Personalization criteria for enhancing learner engagement in mooc platforms. In *2018 IEEE Global Engineering Education Conference (EDUCON)*, pages 1265–1272, April 2018.

[32] Janneth Chicaiza, Nelson Piedra, Jorge Lopez-Vargas, and Edmundo Tovar-Caro. Recommendation of open educational resources. An approach based on linked open data. In *IEEE Global Engineering Education Conference, EDUCON*, pages 1316–1321. IEEE Computer Society, jun 2017.

[33] Jürgen Buder and Christina Schwind. Learning with personalized recommender systems: A psychological view. *Computers in Human Behavior*, 28:207–216, 01 2012.

[34] Hendrik Drachsler, Katrien Verbert, Olga C Santos, and Nikos Manouselis. Panorama of Recommender Systems to Support Learning. In *Recommender Systems Handbook*, pages 421–451. 2015.

[35] Peter A Henning, Florian Heberle, Alexander Streicher, Andrea Zielinski, Christian Swertz, Jürgen Bock, and Stefan Zander. Personalized Web Learning: Merging Open Educational Resources into Adaptive Courses for Higher Education.