Brainard the Fox (Interactive Bibliography Tool)

Contents

Executive summary

Key project outcomes

The Brainard project successfully created a functional AI-powered research companion that addresses a genuine pedagogical challenge: helping undergraduate students navigate the overwhelming scope of early medieval British history. The prototype chatbot, trained on over 100 scholarly summaries, demonstrated clear value in providing structured guidance to students approaching complex historical topics.

Testing revealed that Brainard excelled in several key areas. The tool consistently provided coherent starting points for research, with testers particularly praising its "suggested order of reading" feature that made daunting topics feel manageable. Graduate students found the opening summaries "extremely thorough" and "very clear," while librarians appreciated the tool's ability to offer "surprisingly rounded and well-expressed" answers with "engaging modesty." The chatbot successfully reduced barriers to entry by contextualizing topics within broader historical frameworks and suggesting logical reading sequences.

However, evaluation also revealed important limitations. The narrow training dataset meant recommendations could be repetitive or too broad for specific

essay questions. Some suggested sources were outdated, and the tool occasionally failed to include obvious primary texts or made errors when asked about concepts outside its training data. Most significantly, one tester demonstrated how easily the tool could generate essay frameworks, raising concerns about academic integrity and the development of independent critical thinking skills.

The project achieved its primary objective of helping students navigate difficult historical material while maintaining appropriate academic guardrails. The tool successfully consolidated recommended reading lists into an interactive format without providing complete essay solutions, though this balance proved more delicate than initially anticipated.

Key learning for the team around using AI for developing teaching and learning

The development process revealed crucial insights about implementing AI in humanities education. First, the quality and scope of training data directly determines the tool's effectiveness. Our initial dataset of 100 sources, while substantial, proved insufficient for comprehensive coverage of early medieval topics. The most time-consuming aspect was sourcing and summarizing relevant materials, highlighting the importance of adequate preparation time for AI teaching tools.

Technical collaboration between humanities faculty, the AI Competence Centre, and Digital Innovation Team proved essential. The interdisciplinary approach enabled successful integration of domain expertise with technical capabilities, though it required careful coordination and clear communication about pedagogical goals versus technical possibilities.

The testing phase revealed that AI tools in humanities education must be designed with explicit boundaries. While Brainard successfully provided research guidance, the ease with which it could generate essay content highlighted the need for clear usage guidelines and ongoing discussions about academic integrity. This suggests that AI teaching tools should be implemented alongside explicit training about appropriate use rather than as standalone solutions.

Faculty response varied significantly, from enthusiasm about the tool's potential to serious concerns about encouraging AI dependency. This highlighted the importance of broad consultation and change management when implementing AI tools in traditional disciplines. The most constructive feedback came from colleagues who engaged critically with the tool's limitations while recognizing its potential benefits.

Perhaps most importantly, the project demonstrated that successful AI teaching tools must enhance rather than replace fundamental academic skills. The most praised features were those that provided scaffolding for independent research rather than answers, suggesting that AI's role in humanities education should focus on guidance and skill development rather than content generation.

Project introduction

Background and context

The tool helps UG students who tend to struggle most with pre-modern period papers (late antiquity and the Middle Ages) due to the unprecedented scope of material, both in terms of geography and timespan. Brainard the Fox is an AI-powered research companion tool to signpost students to relevant materials from Oxford's vast available sources. The prototype focuses on HBI1 and uses materials recommended by History Faculty.

The tool allows students to assess quickly which titles within the overwhelming reading lists they should focus on for their particular take on the broad tutorial essay questions, and to test-drive their thesis with an enhanced interactive and accessible reading list. Meanwhile the bot cannot write essays through the use of AI guardrails.

Objectives and outcomes

1. Help UG students (History or JS) navigate the most difficult papers of the course.
2. Accumulate recommended reading lists and make them interactive.
3. Help drive student engagement with existing bibliography; reduce barriers to entry; efficiency through easier access to relevant texts.
4. Increase UG students' engagement and accessibility through generative AI-based interactions.

1. A novel way of engaging with less-known, vast material for UG students at the time when their subject tutors are unavailable for consultations and peers are not reliable in their knowledge, while helping subject tutors keep abreast with the recent bibliography.

2. Consolidated body of recommended information instantly accessible for signposting and essay prototyping, without doing the heavy lifting of offering structure and doing actual writing.
3. A novel way of engaging with pre-modern texts with greater levels of UG student interest, insight and critical thought.
4. Potential side benefits to faculty include creation of better reading lists and supporting the creation of new examination questions.

Scope

Due to time and funding constraints, especially at the initial data acquisition stage, the prototype only covered just over 100 summaries of scholarly works.

Tools and technologies

The database for primary sources was Zotero, allowing form collecting and sharing metadata.

The AI Competence Centre built a backend tool for importing docs from Zotero, and managing the bulk processing of docs sent to Gemini AI for summary creation. Gemini 2.5 was selected for its large context window and OCR abilities. The tool allows document grouping, prompt experimentation, and summary exporting.

The end-user facing chatbot was created using a Chatbot Framework developed by the Digital Innovation Team which enables customised chatbots to be easily created and distributed to web users. The core AI technology was Open AI gpt-4.1 selected for its large input context window and accessed via the OpenAI Responses API.

Collaboration

Members of the History Faculty, AI Competence Centre, and Digital Innovation Team worked closely throughout the project, each contributing key components of the solution.

Project outcomes and findings

Evaluation results

Eight evaluators from diverse academic backgrounds tested Brainard, including university administrators, graduate students, medieval historians, librarians, and colleagues from other historical periods. Testing involved both structured queries about specific historical topics and exploratory use to assess the tool's boundaries and capabilities.

Quantitative data

 Evaluators: 8 academic staff/students across different roles and expertise levels

 Test queries: Approximately 30+ specific historical topics tested

 Training data: 100+ scholarly source summaries

 Response time: Immediate responses to user queries

 Technical accessibility: Successfully deployed via web interface with password protection

Qualitative insights

Feedback revealed consistent strengths in three key areas. First, the tool's ability to provide contextual overviews was highly valued, with one evaluator noting responses were "surprisingly rounded and well-expressed" with "engaging modesty." Second, the structured guidance feature—particularly the "suggested order of reading" – was universally praised as helpful for students who "sometimes struggle with" knowing where to begin. Third, the tool successfully reduced anxiety about approaching complex topics by making them "less daunting."

However, significant limitations emerged. The narrow training dataset led to repetitive recommendations and occasional inclusion of outdated sources. One evaluator noted suggestions were "somewhat hit and miss" with "nothing really bad" but lacking focus and relevance. The tool struggled with methodological questions and occasionally made errors about concepts outside its training data.

Most concerning was the ease with which the tool could generate essay frameworks. One evaluator successfully created a substantial essay outline in 30 minutes, demonstrating potential for academic shortcuts rather than skill development. This highlighted the critical need for clear usage guidelines and ongoing discussions about academic integrity when implementing AI tools in humanities education.

Lessons learned

Challenges

The primary challenge was balancing comprehensive coverage with limited resources. Data acquisition proved more time-consuming than anticipated, requiring manual sourcing and summarization of over 100 scholarly works.

The interdisciplinary collaboration, while ultimately successful, required careful coordination between teams with different expertise and priorities.

Faculty resistance presented an unexpected challenge. While some colleagues embraced the tool's potential, others expressed fundamental concerns about encouraging AI dependency in humanities education. One colleague refused to test the tool, citing concerns about making teachers redundant and environmental impact of AI systems. This highlighted the importance of addressing both pedagogical and ethical concerns when implementing AI tools.

Technical limitations also emerged during testing. The tool's responses were sometimes too broad for specific essay questions, and the narrow training dataset led to repetitive recommendations. Integration with library systems proved more complex than expected, preventing the desired direct linking to Oxford's resources.

Key takeaways

Successful AI teaching tools in humanities require extensive preparation and clear pedagogical frameworks. The quality of training data directly determines effectiveness, suggesting future projects should allocate more time and resources to data preparation. Clear usage guidelines are essential from the outset, not as an afterthought.

Faculty engagement and change management are crucial for successful implementation. Early and ongoing consultation with colleagues helps address concerns and build support. The most constructive feedback came from colleagues who engaged critically with the tool while recognizing its potential benefits.

AI tools work best when they enhance rather than replace traditional academic skills. The most successful features provided scaffolding for independent research rather than ready-made answers, suggesting AI's role should focus on guidance and skill development.

Advice for teams

Begin with extensive consultation across your faculty to understand both enthusiasm and concerns. Allocate significant time for data preparation – this will likely take longer than expected. Establish clear usage guidelines from the project's inception, not as an afterthought. Focus on features that enhance student skills rather than provide shortcuts. Plan for ongoing evaluation and iteration rather than treating the initial version as final. Consider environmental and ethical implications alongside pedagogical benefits to address colleague concerns constructively.

Appendices

Digital artefacts: for example, prompts you used, or the chatbot

- `brainard-chatbot.zip` contains the chatbot config files and summaries
- Brainard chatbot is hosted at: https://ox-dig-innov.pages.dev/brainard/v2/ (password is required to access, available from The Digital Innovation Team innovation@it.ox.ac.uk)
- The Chatbot Framework is hosted at: https://ox-dig-innov.pages.dev/chatbots/ (password is required to access) and the code is available on GitHub on request from The Digital Innovation Team innovation@it.ox.ac.uk

Raw data: Include or reference datasets, feedback forms, or other raw data collected

Templates used: Provide any templates (eg for digital artefacts or evaluation questionnaires) utilised during the project

Guidance for ‘potential’ users (eg student guidance/comms you had to develop)

Additional materials: List or attach supplementary documentation

SSO-protected information: Indicate sensitive content that requires secure sharing (this will not be published on the website).

Teams are encouraged to provide clear instructions and methodologies to enhance shareability and reproducibility.