

Generating quiz questions using Bing Copilot Chat

Context

Steve Bentley is the Strategic Learning Technology Advisor in the Strategic Teaching and Learning Team. He realised that the introduction of a facility to upload a file as part of a Bing Copilot Chat (Enterprise) prompt gave an opportunity to build on established techniques of using Generative AI as a tutor, enabling the chatbot ask questions that are focussed on the material covered in a teaching session rather than a broad topic, and for staff to be able to generate question sets that could be imported into the VLE or a classroom response solution.

The University of Huddersfield uses D2L Brightspace as its VLE and Vevox for classroom response, but it is anticipated that these techniques could be adapted for other platforms.

A series of base prompts have been provided which can be amended to suit individual circumstances and use cases, or the output can be refined with follow-up prompts.



Strategy

From a Bing Copilot Chat prompt, it is possible to upload a file of up to 1MB and reference this in a prompt. This gives the opportunity to upload the outline of a lecture and then engineer prompts which will generate questions about the content of the session. Users who have a full Copilot license are not subject to a file size limit, so can upload the full PowerPoint file, but other users can open larger PowerPoint files and use the Export option in the File menu to send the Outline to a Word file which can be uploaded instead, although the results may not be quite as good.

The Enterprise Protection feature of the corporate version of Copilot provides reassurance that the material uploaded will remain secure.

A straightforward use case might be for a learner to download the PowerPoint file for a lecture from the VLE, export the outline and upload it to Bing Copilot and launch a prompt such as:

The attached file is the PowerPoint file, or an outline of it, which was used to deliver a university level lecture. The lecturer expanded on the points covered on the slides, providing further explanation, examples and context. Ask me multiple choice questions about using the material that somebody who attended the lecture should have learned and should require application rather than simple recall. The incorrect options should be plausible but definitely wrong. Pause after each question for my answer and give me feedback on it.

This will begin an interactive question and answer quiz, with the chatbot providing feedback on the student's response to each question.

An important aspect of this prompt is the stipulation that the lecturer expanded on the material provided, rather than merely reading out the slides, and that questions should be based on what somebody should have learned from attending. This instructs the chatbot to infer and predict what the lecturer might have said rather than relying solely on the headings in the file. We also specify that the questions should be pitched to require higher level skills rather than simple recall.

VLE Quiz building

Steve was aware that it is possible to import questions to the Brightspace quiz engine using a CSV format. This is not a regularly used facility at Huddersfield because the CSV format required is very specific and it is usually easier to enter questions directly rather than craft a spreadsheet in the correct format. By including a simplified version of the template as part of the prompt, Steve was able to have Bing Copilot output a set of questions in the required format, ready to paste into Notepad or a similar text editor in order to save it with a .txt extension and import into the VLE.

The standard prompt can create Multiple Choice, Multiple Select, True/False, Matching and Ordering questions. The base prompt instructs the chatbot to select appropriate question types to suit the material, but this can be adapted as required. It is possible to create Written Response and Short Answer questions, but it was felt that it was more likely that colleagues would want to create questions capable of being automatically marked, and it was not possible to create a prompt which could reliably create adequate alternative answers for short answer questions to anticipate synonyms and alternative spellings. The other types of question which Brightspace supports, such as numeric and fill in the blanks, can not be imported through the CSV template so are not suitable for this technique.



There are limits on the length of an AI response which restrict the number of questions which can be generated from a single prompt to about 15, but follow up prompts such as **Can you give me 15 more different questions using the same template** can be used to generate further questions. Further manipulation of the prompt can refine the requirements for the questions, for example questions focussing on a particular element of the material covered or at a given level of Bloom's taxonomy.

Happily, the import function in Brightspace offers a review feature, giving an opportunity to reject poor or inaccurate questions, which has synergies with best practice around checking AI outputs before using them.

A further workflow for importing Word based model answers from previous multiple choice assessments has been developed using a modified prompt, allowing past papers to be repurposed for revision.

Vevox poll building

Vevox has an import/export feature which is based on zipped JSON files. These are not intended for human editing, but a similar process can be used to generate Multiple Choice, Multiple Select and Ordering questions ready for import. Additionally, a Word Cloud question, which could be used to provoke discussion, can be generated, because the live polling scenario means that we are not limited to questions which can be automatically marked, or even have a correct answer.

The limits on the size of output and complexity of the template mean that only around 5 questions can currently be generated, but follow up prompts can produce more questions.

Again, the import mechanism prompts for checking the questions and gives an opportunity to reject questions or edit them after import.

Benefits

Through the techniques described in this case study, a number of benefits can be realised.

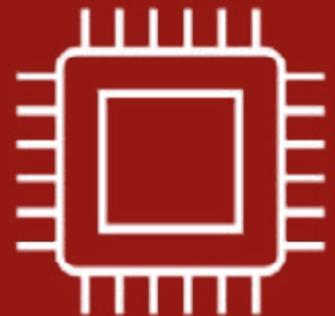
For academic staff, the opportunity to efficiently generate high quality, focussed questions based on the specific material they have been teaching. These could be used in a variety of ways as part of established pedagogical methods:

- VLE quizzes for private study, perhaps using Just in Time teaching principles, so that areas of common misunderstanding are identified and can be revisited in the next session before moving on.
- Polling questions for use with classroom response technology, to consolidate or recap on key points, and check understanding before continuing.
- As low stakes assessment (through the VLE or a polling platform) to promote engagement and attendance.
- Create banks of questions themed around each lecture, which can be consolidated to create revision quizzes.

These techniques produce questions using a range of question types, and which can require higher order skills rather than simple recall. Writing high quality questions that go beyond recall is a time consuming activity, so assistance from AI can make this much less a burden.

Practice questions are regularly requested in feedback surveys, so these techniques make it possible to provide them with minimal additional work for staff.

Additionally, this is a very tangible way that Generative AI could be useful to academic colleagues, so serves as a good means of introducing a positive, ethical use case to staff whose existing exposure may be in the context of academic misconduct, and an opportunity to explore simple prompt engineering by issuing follow-up prompts to refine the output.



Both Brightspace and Vevox offer native facilities to create questions using AI (in the case of Brightspace, this is through a paid-for add-on service which many customers will not have access to) but this approach potentially offers more customisation through prompt engineering and scalability by creating a larger number of questions using a variety of question types in a single operation. In the case of Vevox, this approach broadens the types of question which can be generated to include Ordering and Multiple Select, and allows for the content to be more targeted to the content of the teaching session rather than a broad topic.

Caveats

Like all Generative AI techniques, there is a need to check and verify the output. It is not anticipated that this will be particularly burdensome for tutors who are subject matter experts.

The quality of the output does depend on the quality of the input. There is a requirement that the PowerPoint file has been developed with some regard to accessibility principles, particularly following best practice for the use of slide titles, so that the extracted outline captures the key areas that were discussed. Slides where the reading order has been set incorrectly may produce unexpected results when generating ordering questions. Providing an incentive for staff to observe accessibility requirements when they prepare teaching materials may be a further positive benefit of these techniques.

Users who have a full Copilot license and can upload the full PowerPoint file are likely to obtain better results, as their input to the prompt will include the alt text descriptions of images and material inside SmartArt objects which are not included in an exported outline.

Feedback

The reception from staff who have attended demonstrations has been very positive, with several colleagues expressing their surprise at the breadth and depth of the questions provided, and their appreciation for the time saved.

"Using Copilot to create an MCQ quiz in Brightspace, based on the teaching resources that had been used in class, was a highly efficient and time-saving process. Copilot was able to generate a variety of questions quickly, which provided a strong foundation for the quiz. While the generated questions were useful, some did require refinement to ensure that they were clear and aligned with the specific learning objectives of the session. Some minor adjustments were needed to improve the difficulty of the questions and ensure that the answer choices were balanced and unambiguous."

"Overall, it proved to be a really valuable tool for streamlining quiz creation. It significantly reduced the time spent on drafting questions. However, a final review was necessary to refine the content, check for accuracy, and enhance the overall engagement and fairness of the quiz."

Charlotte Kenyon, Course Leader, BSc Midwifery, University of Huddersfield

As a new development which was introduced mid-year we anticipate there being more take-up during the summer period as academic staff prepare for the coming session and are able to consider changes to their pedagogical approach.

Example AI prompt: Generating Brightspace quizzes

The prompt below is intended as a starting point, the requirements at the end of the prompt can be modified to suit your specific needs. The template shown here is specific for Brightspace, but any similar CSV or JSON import template could be substituted for use with other platforms.

The code produced should be pasted into a text editor such as Notepad and saved with a .csv extension. (In Notepad, change "Save As Type" to "All Files"). This process is demonstrated in the accompanying video.

This is a template for importing questions into a VLE quiz feature, for various types of question.

```
//MATCHING QUESTION TYPE
//Choices and Matches must include text in column3
"NewQuestion","M"
"ID"
"Title","This is a matching question"
"QuestionText","This is the question text for M1"
"Points","2"
"Difficulty","2"
"Image"
"Scoring","EquallyWeighted"
"Choice","1","This is choice 1 text"
"Choice","2","This is choice 2 text"
"Choice","3","This is choice 3 text"
"Match","3","This matches with choice 3"
"Match","1","This matches with choice 1"
"Match","2","This matches with choice 2"
"Hint","This is the hint text"
"Feedback","This is the feedback text"

//MULTIPLE CHOICE QUESTION TYPE
//Options must include text in column3
"NewQuestion","MC"
"ID"
"Title","This is a multiple choice question"
"QuestionText","This is the question text for MC1"
"Points","1"
"Difficulty","1"
"Image"
"Option","100","This is the correct answer"
"Option","0","This is incorrect answer 1"
"Option","0","This is incorrect answer 2"
"Option","0","This is incorrect answer 3"
"Hint","This is the hint text"
"Feedback","This is the feedback text"

//MULTISELECT QUESTION TYPE
//Options must include text in column3.
"NewQuestion","MS"
"ID"
"Title","This is a Multi-Select question"
"QuestionText","This is the question text for MS1"
"Points","10"
"Difficulty","5"
"Image"
"Scoring","RightAnswers"
"Option","1","This is option 1 text"
"Option","0","This is option 2 text"
"Option","1","This is option 3 text"
"Option","1","This is option 4 text"
"Option","1","This is option 5 text"
"Hint","This is the hint text"
"Feedback","This is the feedback text"

//ORDERING QUESTION TYPE
//Items must include text in column2
"NewQuestion","O"
"ID"
"Title","This is an ordering question"
"QuestionText","This is the question text for O1"
"Points","2"
"Difficulty","2"
"Scoring","RightMinusWrong"
"Image"
"Item","This is the text for item 1","HTML"
"Item","This is the text for item 2","HTML"
"Hint","This is the hint text"
"Feedback","This is the feedback text"
```

It is CSV based, with values enclosed in quotation marks. Keep the line breaks intact. Write 10 questions using these templates using a variety of question types based on the attached powerpoint file. The questions should check that somebody who attended this presentation is able to apply the concepts discussed and be more challenging than simply recalling the material. Assume that the presenters did not simply read the text on the screen but added additional explanation, examples and embellishment. Each question should have a single, unambiguous correct answer. Leave the ID and image fields blank.

Multiple Choice and Multi Select should have at least 4 possible answers. For Multiple choice and Multi select, place the option rows in random order so that the first answer is not always the correct one. Vary the number of correct answers on multiple select but ensure that there is 1 or more incorrect options. Incorrect options should appear plausible but definitely be wrong. Ordering questions should include between 4 and 6 Item rows. Choose the most appropriate question types, it is not necessary to use all types. Present the questions in a continuous list - do not add headers for each question type. Present the output in the chat window, not as a downloadable file.