



Answer real questions using Wikipedia content

Axel Didier, Frédéric Assmus, Pierre Goncalves



Summary

- Reminder on the subject
- What we have done
- Issues encountered
- Ongoing work
- Future work



Reminder on the subject

- Extracting the subject of the question
- Finding the most relevant wikipedia article given this subject
+ the most relevant subpart of the article
- Giving the question and the subpart to BERT so it can find a pertinent answer
- Having a website interface for our question-answering system



What we have done

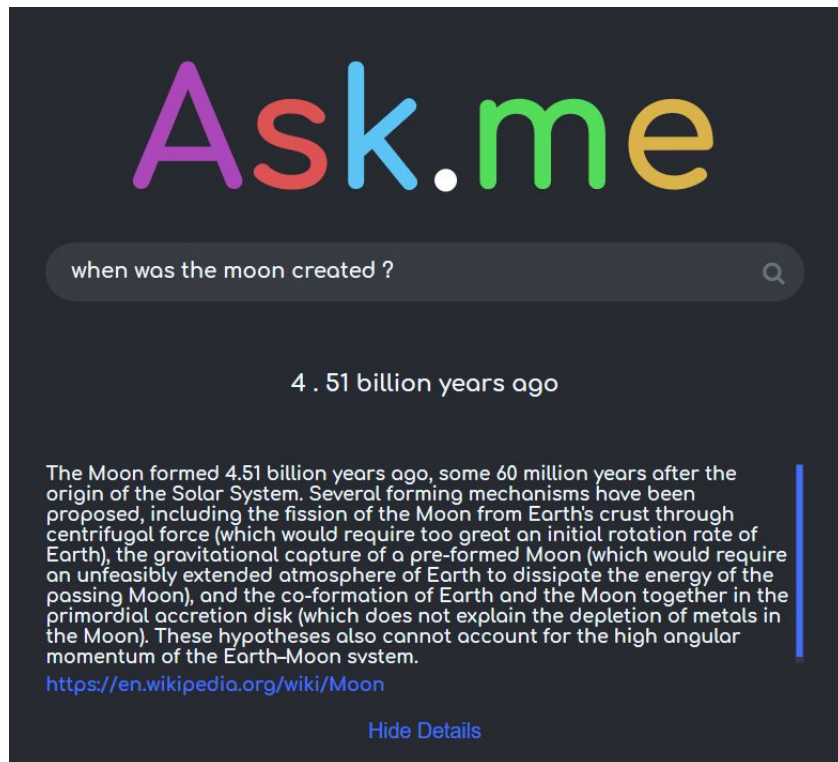
Solve previous issues and improve the system :

- **Subject Extraction** : Basic rules showed weird results for subject extraction. We extended the rule by using a combination of NER and noun-chunks extraction to solve the issue.
- **Article Retrieval** : Our way of dealing with the API produced weird results : sometimes it gave to the model an article that wasn't related to the question. We ameliorated our system by combining a new wikipedia-oriented library to the one we already used. This new article retrieval system is now most of the time able to correctly find the most relevant article related to the subject.
- **Website implementation** : Our work is now directly integrated on the website, which gives a good idea of what the final version will look like.



Implementation

- Display Answer
- Display the paragraph where the answer was found
- Link to the wikipedia page



The screenshot shows the Ask.me website with a dark background. At the top is the 'Ask.me' logo in large, colorful letters. Below it is a search bar containing the text 'when was the moon created ?' with a magnifying glass icon on the right. The answer '4 . 51 billion years ago' is displayed in a light blue font. Below the answer is a paragraph of text explaining the formation of the Moon, mentioning various hypotheses like fission, capture, and co-formation. At the bottom of the paragraph is a blue link to the Wikipedia page: <https://en.wikipedia.org/wiki/Moon>. A 'Hide Details' link is visible at the bottom right of the text area.

Ask.me

when was the moon created ?

4 . 51 billion years ago

The Moon formed 4.51 billion years ago, some 60 million years after the origin of the Solar System. Several forming mechanisms have been proposed, including the fission of the Moon from Earth's crust through centrifugal force (which would require too great an initial rotation rate of Earth), the gravitational capture of a pre-formed Moon (which would require an unfeasibly extended atmosphere of Earth to dissipate the energy of the passing Moon), and the co-formation of Earth and the Moon together in the primordial accretion disk (which does not explain the depletion of metals in the Moon). These hypotheses also cannot account for the high angular momentum of the Earth-Moon system.

<https://en.wikipedia.org/wiki/Moon>

Hide Details



Issues encountered

- **Disambiguation** : find a way to handle multi sense word (e.g “What is a python ?”)
- **Paragraph selection / time computation** : Currently looks at 15 first paragraphs of the article (to **limit computation time**), computes answer for each and retrieves answer which has the highest confidence.
Would like to find new ways to **select pertinent paragraph** without processing all articles or losing information.

Ongoing work

- Developing functions that are able to detect the type of question. Our model is only able to answer to factual general knowledge questions. Goal : Tell the user that we can't answer his question, in a natural way.
- Checking other similar projects to see if we can find new ideas to improve our current solution.

Type	Examples
Yes/No	<i>Is the earth flat ?</i>
Pseudocleft	<i>What she says is true.</i>
Comparative	<i>Is the Moon bigger than Mars ?</i>
Subjective	<i>What is Madonna's best album ?</i>
To App	<i>What's your favorite color ?</i>
Tutorial	<i>How to unlock a closed door ?</i>
Technical	<i>What is the treatment for pharyngitis ?</i>
Personal	<i>Should I talk to my ex again ?</i>
Factual	<i>When is Barack Obama born ?</i>



Future work

- Continue working on question-type detection
- Find more efficient ways to select paragraphs
- Find a way to deal with disambiguation
- Check if it is possible to extend our system to other languages
- Check if it is possible to extend our system to answer Yes/No questions too



Thank you for your attention

Questions ?