

# Solution Outline - Mid

**Study the role of Wikidata in enhancing the productivity of the Wikipedia writers and enhance Wikidata for Hindi and Telugu using gamification.**

**Name** - Himanshu Maheshwari  
**Roll No** - 20171033

**Mentors** - Tushar Abhishek and Nikhil Pattiisapu

---

In the solution outline that was presented in version 2 of 'Problem outline and defining expected outcome', there are two sub-system in the game. The first is 'Question Creator' and the second is 'The Game'. Till now I have implemented Question Creator.

Just to recap - Question creator finds out various properties of an item whose value is missing. Now if we look at all the categories that are targeted, all of them contain things that have almost similar properties. For example Chemical Elements contains Oxygen and Argon. If we look at Wikidata entries for both of them we will see that both of them have similar properties entries. For example both of them have Atomic mass, Atomic number, Discoverer, etc. Thus if some property is missing in Argon but is there in Oxygen then that property must be there in Argon too. For example - Oxygen has property "Oxidation state", however this property is missing in Argon. Thus this system identifies such missing properties and then creates questions around them.

Entire code can be found at - <https://github.com/him-mah10/social-computing-project>

## Approach

The first part of the question creator downloads all the entries that belong to a particular category. For example in the case of chemical elements, it properties of all the elements. These are then stored. The entire thing is downloaded in Hindi. We first downloaded the list of all the entries that belong to a particular category. In case of chemical elements, that would be list of all chemical elements. Then we download list of all properties of each chemical elements.

अमेरिशियम	अबियम	आइस्टीनियम	आयोडिन	आर्गन	आर्सेनिक	इट्रियम	इट्रियम	इण्डियम	इरीडियम
उनउनदिरियम	एल्युमिनियम	ऐक्टिनियम	ऐन्टिमोनी	ऐस्टाटीन	ऑक्सीजन	ओगेनोसोन	ओरियम	कार्बन	कैडमियम
कैलीफोर्नियम	कैलियम	कोबाल्टनिसियम	कोबाल्ट	क्यूबियम	क्रिटन	क्रोमियम	क्लेवीन	कुदरातु	मैडोलिनियम
गैलियम	गंधक	छांदी	जर्कोनियम	जर्मेनियम	जस्ता	ज़ेनान	टर्बियम	टाइटेनियम	टिन
टेक्नैशियम	टेनेसीन	टेलुरियम	टैटेलम	टंगस्टन	डबियम	डाम्स्टंडशियम	डिस्प्रोसियम	ताम्र	तेजातु
थुलियम	थैलियम	थोरियम	नाइट्रोजन	नायोबियम	निकल	नियोडाइमियम	नियोन	नेप्टूनियम	नोबेलियम
पलेडियम	पारा	पोटैशियम	पोलोनियम	प्रांसियोडाइमियम	प्रोटेक्टिनियम	प्रोमीथियम	प्लूटोनियम	प्लैटिनम	फर्मीयम
फास्फोरस	फ्लोरोवियम	फ्लोरीन	बर्कलियम	बिस्मथ	बेरियम	बेरिलियम	बोरियम	बोरॉन	ब्रोमीन
मेइटरनैरियम	मेण्डेलीवियम	मैन्गेशियम तत्त्व	मैंगनीज़	मोलिब्डेनम	मोस्कोवियम	यूरोपियम	यूरेनियम	रदरफोर्डियम	रीनियम
रुथेनियम	रुबिडियम	रेडॉन	रैन्टोनियम	रोडियम	लिथियम	लिवर्मोरियम	लुटीशियम	लेन्थनम	लोरेंसियम
लोहा	वनेडियम	सिलिकॉन	सीज़ियम	सीबोर्गियम	सीरियम	सीसा	सेलेनियम	सेमरियम	सेडियम
सोना	स्कण्डियम	स्ट्रॉन्टियम	हाइड्रोजन	हाफ्नियम	होलियम	हैसियम	होल्मियम		

द्वारा नामांकित  
 चित्र  
 खोज की जगह  
 GND अभिज्ञापक  
 LCCN अभिज्ञापक  
 उद्धरण है  
 खोजकर्ता या आविष्कारक  
 स्रोत द्वारा वर्णित  
 द्रव्यमान  
 का उपवर्ग  
 का उपवर्ग  
 NDL पहचानकर्ता  
 कॉमन्स श्रेणी  
 BNCf कोश  
 फ्रीबेस पहचानकर्ता  
 विषय की मुख्य श्रेणी  
 कॉमन्स गैलरी

After that is done, we randomly select two elements and find the list of all the properties that are there in first and not in second. After this list is found we randomly select one property and create question on that property. Since the property is two or three words, we can use template method to create question. The template in case of chemical elements is: 'तत्व'+2nd element name in Hindi+ ' का '+missing property +' क्या है?'. Though this method might result in

grammatically incorrect sentences, it is still comprehensible. Some questions that are generated by this methods are:

तत्व रीनियम का खोज की जगह क्या है?  
तत्व इण्डियम का BNCF कोश क्या है?  
तत्व हैसियम का द्रव्यमान क्या है?  
तत्व पोटैशियम का फ्रीबेस पहचानकर्ता क्या है?  
तत्व अर्बियम का ध्वनि का वेग क्या है?

Due to mids the meeting and finalization of 'Problem outline and defining expected outcome' became delayed and finalization was done on 2nd March(3 days before) and thus I was not able to do much.

