GSoD 2020 Project Proposal
The Open Multilingual Wordnet (OMW)

Rohitesh Kumar Jain   thekumarjain@gmail.com

Resumé:

I am a Bachelor of Engineering sophomore, pursuing Computer Engineering from Indraprastha University of Delhi, India. I have a working proficiency in English. I am a technical writing enthusiast and I always used to do the documentation part of any project my group was assigned (apart from the project of-course!) I am comfortable with python and GIT.

My LinkedIn profile

SAMPLE DOCUMENTATION: CONCEPT PAGE

My development blog

Project request:

If given a chance I can spend 40 hours (or more) a week on this project.

I have a good grasp of Python, Git, HTML, C++, Java, and C. I am bilingual and I can speak fluently in English and Hindi, I learned French at school and know a smattering of German also. I am very keen and enthusiastic about contributing to the Global Wordnet Association through this project if given a chance I will work with utmost dedication. Due to COVID-19 my University has been closed for months now, and I want to utilize this time for learning new things and contributing to an esteemed open-source organization. I have good problem-solving skills and always been at the top of my class in academics throughout my student life. I am comfortable with photo editing and can handle graphics which be required to be used in the documentation of this project. I am a quick learner who picks up things quickly and adapts to a new environment smoothly.

Project work plan:

Current State:

The current Open Multilingual Wordnet documentation is shallow and incomplete existing user guides lack content, graphics, examples, and explicit instructions. There is no user guide to assist someone searching the wordnet, how to read the returned pages, how to install, how to provide feedback. Some documentation is present on contributing new wordnet, but that documentation is neither complete nor clear to guide someone how to join the Open Multilingual Wordnet. There is

no documentation for how to interact with the Collaborative Interlingual Index (CILI), there is a CILI introduction page, which lacks information like why and how to deprecate and supersede concepts, judging new candidates. Overall the current documentations lacks in content, clarity, examples, and visuals for assistance.

Plan:

I have planned to execute this project in four phases. In phase A, I will document a guide for someone contributing new wordnet, in this phase will deal with how to create well-formatted wordnet, how to upload it, and how to make sense of the feedback. In the next phase B, I will document a guide for someone searching the wordnet(or family of wordnets), in that documentation, I will cover how to read the returned pages, how to download, and how to provide feedback. In phase C, I will document a guide for interacting with the Collaborative Interlingual Index, how to deprecate and supersede concepts, and judging new candidates will be documented in phase C. In phase D, I will populate the documentation of the Open Multilingual Wordnet, improve existing documentation of this project other than the documentation covered in phases A, B, and C. As there is overlap between the projects of GWA, phase D will also be dedicated to organizational requirements. I have planned to allocate weeks 1,2 & 3 for phase A, weeks 4,5 & 6 will be allocated to phase B, weeks 7,8 & 9 for phase C and weeks 10 and 11 for phase D.

Project Goals:

1)Analyzing the existing documentation, and determining which section requires improvement.

2) Reconstruction of the structure of existing documentation.

3) Documenting a guide for someone contributing new wordnet.

4)Documenting a guide for someone searching the wordnet.

5)Documenting a guide for interacting with the Collaborative Interlingual Index.

6) Adding examples and visuals to the documentation.

7) Populating the existing pages where the documentation seems shallow.

Timeline:

Note: This timeline is flexible and recognizes the fact that organizational needs can change.

Phase	Goal and Deliverables
Community bonding phase: 17 August- 13 September.	• In this period I will interact with the community and will get to know the community better. • Set up a communication channel. • Set up a communication time. • Refine goals and set expectations on both sides.
Doc Development phase (September 14- 30 November)	• I will work closely with my mentors and will complete my project in this phase according to the weekly plans. • Complete any survey from the Google program administrators.

Week 1-3 (September 14-27)	• Documentation phase-A
Week 1 (September 14-20)	• Documentation for a guide for someone contributing new wordnet will begin. • This week I will do the documentation, how to create well-formatted wordnet. • Clear instruction for converting the user's wordnet to the public Wordnet LMF document type definition (DTD) will be documented. • Conversion tools will also be covered in it.
Week 2(September 21-27	• This week I will do the documentation for how one will upload one's wordnet(after converting to Wordnet LMF document type definition) to the OMW. • The existing documentation is shallow (but accurate) I will populate it and add examples to the documentation.
Week 3(September 28-Ocotober 4)	• This week will be utilized for documentation of how to make sense of the feedback. • This documentation is required to be written from scratch, as there is no existing documentation on this topic.
Week 4-6(October 5-25)	• Documentation phase-B
Week 4(October 5-11)	• This week I will start documenting a guide for someone searching the wordnet (or family of wordnets). • This week I will write documentation for how to read the returned pages.
Week 5(October 12-18)	• This week I will document how to download. • There is no existing documentation for this it is needed to be written from scratch.
Week 6(October 19-25)	• This week I will document how to provide feedback to the organization. • As there is no previous documentation to take reference from, this is also going to be written from scratch.
Week 6-9(October 26-November 15)	• Documentation phase-C
Week 7(October 26-November 1)	• This week I will start the documentation of a guide for interacting with the Collaborative Interlingual Index (CILI) • This week I will document how and why to deprecate and supersede concepts.
Week 8(November 2-8)	• This week I will document additional information about CILI, and write instructions on how to download CILI and fulfill organizational requirements on the CILI documentation this documentation needs to be written from scratch.
Week 9-11(November 9-29)	• Documentation phase-D
Week 9(November 9-15)	• OMW 2.0 will be developed now, according to organizational needs.
Week 10(November 16-22)	• Documentation of the Open Multilingual Wordnet(OMW) home page seemed shallow, this week I will work on it, and add the additional information related to OMW, and I will look at which other OMW existing pages need to be populated. • Lacking OMW pages will be improved and whenever possible, I will try to add visuals in it.
Week 11(November 23-29)	• As there is overlap in projects of GWA, this week is reserved for fulfilling these overlapping organizational needs, and other

	miscellaneous documentation.
Project finalization and evaluation phase (November 30 – December 5)	• This week will be spent on proofreading the documentation for typos and errors. • I will write my project report. • I will request my mentors to review my project report. • I will finalize and submit my project report. • I will write and submit an evaluation of my mentors and my experience of Season of Docs.

Communication:

The frequency of contact will be dependent on a discussion with mentors. I prefer regular communication - my suggestion is twice every week to provide an update and receive feedback on work done so far. However, this is flexible depending on their preferences as well.

Why this project?

I have an interest in linguistics, and the idea of connecting wordnets for all the languages in the world intrigued me to join Global Wordnet Association and contribute to this project. The concept of Open Multilingual Wordnet is admirable where individual wordnets are linked to wordnets in other languages, which makes these wordnets more easily accessible. The Open Multilingual Wordnet and its components can be freely used, modified, and shared by anyone for any purpose. Although GWA has a solid idea and noble intention, the organization is not much renowned and is underrated. Probably one of the reasons is shallow documentation and lack of user guides to assist in the contribution to it and to use it. I perceived that the Open Multilingual Wordnet project is something in which I can contribute to assist the Global Wordnet Association and the open-source community in common.

Future Plans:

After completing GSoD 20, I will keep contributing to the Global Wordnet Association in other projects also. I will surely attend any conference of the organization if given a chance.

I am grateful that you read my proposal, thank you!