CSC3003S Capstone Project — Stage One

Goals (Scope) [21 Marks]

|  |  |
| --- | --- |
| Project Abbreviation and Name | AI-Web - A web application for the AI research community in South Africa |
| Client/Supervisor + email | Deshen Moodley, deshen@cs.uct.ac.za |
| Tutor + email | Jane Imrie, IMRJAN001@myuct.ac.za |
| Date | 2/08/2022 |
| Team Members | Student Number, First and Last Name, email |
|  | UPDRAU001, Rauseenah Upadhey, updrau001@myuct.ac.za |
|  | ORRKAU001, Kauthar Orrie, orrkau001@myuct.ac.za |
|  | KHSSHR001, Shreeya Khoosal, khsshr001@myuct.ac.za |
| Overall purpose and stakeholders [5] | ***Overall Purpose:***  The primary purpose of the web application is to create a platform through which users can engage with an up-to-date aggregation of AI related data.  Users should be able to perform intuitive queries in order to filter and display information according to a range of parameters. Given that the application will collate data from multiple sources, the platform can serve as a centralized resource that enable users/institutions to easily stay informed on the latest advances/contributions within the South African AI community.  In addition to this, end users can make use of the platform to view contemporary trends and statistics calculated from the repository of AI data, in a visually accessible manner. Various stakeholders, such as researchers, the general interested public, educational and funding institutions, as well as external government/accreditation institutions should be able to interact with the platform in order to retrieve relevant data that can be filtered according to several categories such as the institution, region, subtopic, research type, number of citations as well as NRF rating. Users will be able to engage with a range of research types as well as toggle through different display views as the data will be presented in an intuitive, dynamic format that is suited to its content. Finally, administrative users should be able to make use of the platform in order to carry out manual updates to the data to ensure that the data displayed is relevant/up to date.  ***Stakeholders****:*   1. Researchers 2. Broad AI community in South Africa 3. Educational Institutions 4. Funding Institutions 5. External Accreditation Institutions |
| SMART Goals [5] | ***Supported User Functionality***   * Access an overview of what institutions or regions researchers are based in * Visualise interesting trends and patterns over time within the South African AI research community * Access information about dominant research areas/topics, publications venues, co-authors and citations * Access and analyse metrics of the AI community * Access to updated research from MAG, NRF and other resources * users/institutions can stay informed on the latest contributions within the South African AI community   ***SMART Goals***   * Research and commit to the programming frameworks/languages that will be used to develop the application by the end of the first week, starting with the MERN stack. * Set up underlying database structure by pulling and collating data from NRF and MAG with appropriate DBMS (eg.MongoDB) within the first half of second week. * Integrate frontend (React) with backend to allow for simultaneous full-stack development within the second week of development and implementation stage. * Implement and test functionality for the user to perform basic search/update queries by first demo on the 19th of August. * Develop visual/dynamic representations of the data and layer this functionality on query results and integrate with prototype ideally before first demo but with a leeway of up to week 4. * Implement functionality that allows for manual/updates or synchronisation of the data by the final project submission. * Add functionality for backend to calculate patterns or trends in the research data after the development of the first prototype. * Document every step of the project on a weekly basis by creating progress reports, facilitating team meetings and tracking progress |
| Inputs, outputs and performance [5] | ***Inputs***:   * Filter parameters (dates, number of citations, topics, area/region) * Search parameters (article names, research subtopics/keywords, authors, links, and citations) * Updated data (such as new NRF rating)   ***Outputs:***   * filtered lists of entities such as institutions, research and researchers which will be sorted according to their relevance to the search query * Interesting/meaningful patterns or trends within the data * visual representations of the data such as graphs or charts   ***Response time requirements:***   * Response to queries need to be within the average response time of 200 milliseconds to 1 second to allow the user to not feel any delay * Due to large data sets used by the application, the ideal response time to filtering the data and responding to the user should be around 1 second |
| Resources and Constraints [3] | ***What resources apart from the project team are needed?***   * Online learning resources to learn required language * Communication channels/product management software to allow for effective communication between members * Tutor support * CSGitLab to backup project data   ***Constraints/Challenges:***  The time allocated for the completion of the project proves to be the most significant constraint, however the size of the dataset may also be challenging to accommodate. In addition to this, new software might have a steep learning curve and implementation may be hindered by a lack of proficiency within the required languages. These are therefore substantial challenges within the development process.  ***In what ways will the software be limited?***   * The software may not integrate seamlessly with the other development tools that are being used * The software may not be exhaustive/able to accommodate the desired functionality * If the software is proprietary, the source code will be inaccessible and thus a limited level of customization is feasible * The performance of the software is additionally limited by hardware constraints |
| Feasibility [3] | The core functionality stipulated by the requirements/brief is feasible to implement due to the availability of online resources and tools that facilitate the creation of a web application. The team members can therefore learn the required programming languages; however, they do face time constraints which introduce a significant feasibility challenge. In addition to this, long-term maintenance of the software which is created is infeasible given that students will not be incentivised to prioritise its maintenance and have limited capacity to do so. The project additionally requires that trends or patterns in the data are reported on and displayed, however large datasets and sophisticated analytical tools are required to produce any meaningful statistics from the data. Therefore, given the size and scope of the data that this would require, it may not be feasible to implement this functionality within the allocated duration of the project. |