Requirements Document

September 25, 2020
Colin Choi, Dalton Ronan, Erik Ligai, Harrison Peters, Micheal Nguyen, Md Ridwan Hasan

Project Overview

The *Responsible AI Design Assistant* is a tool to help ensure those who are deploying an AI system do so in a responsible way. The Design Assistant is a web-based survey that will assess the respondent's answers, score the responsibility of the AI system, and provide recommendations for improving the system, based on the 5 dimensions of the AI Global Responsible AI Trust Index:

1. Accountability
2. Explainability and Interpretability
3. Data Quality
4. Bias and Fairness
5. Robustness

The product will be a self-sustaining, self-assessment tool that can be completed by project members at any point of their AI project development. Questions will be flexible and catered to the respondent's domain and subsequent answers. Likewise, the score assessment report card and recommendations will be tailored to the respondent to provide them with clear and specific ways they can improve their system. Survey results will be stored for future analysis and research by AI Global.

Project Glossary

Accessibility: The extent to which how easily data is available and/or retrievable

Accountability: Refers to the extent to which an organization has good governance over aspects such as risk assessments, organizational structure, decision making, and independent review of the systems and inputs.

Accuracy: Correctness of the data; Reliability of the data

Bias and Fairness: Like the human rights harms in other uses of technology that leverage data, the harms related to the use of AI often disproportionately impact marginalized populations.

Completeness: The extent to which the data covers the problem domain (breadth, depth, scope) of the task at hand

Consistency: The extent to which the data's format is consistent with other data in the problem domain

Data Quality and rights: The extent to which a dataset exhibits exemplary levels in each of the Data Dimensions (see: Completeness, Accuracy, Timeliness, Consistency, Accessibility) underlying data rather than the algorithm itself are most often the main source of the issue. Models may be trained on data containing human decisions or on data that reflect second-order effects of societal or historical inequities.

Dimensions of the trust index: Accountability, Explainability, Data Quality and rights, Bias and Fairness, Robustness.

Explainability: AI system would be able to explain its own reasoning to a user in a format that is understandable to a human

Interpretability: Extent to which a cause and effect can be observed within the AI system

Privacy: Who has access to the data collected. Who can make changes to the data. How transparent is it to the user about what data is being collected

Responsible: AI Global "did the hard work of deciphering the best practices, policies, and principles and put them into a simple online survey". Used the "most cited principles, whitepapers, and policy documents published by academics, standards organizations, and companies and translated them into comprehensive questions." All this information is used to determine what a responsible AI is and what principles it should abide by.

Risk Assessment: Evaluation on the potential that an AI program may negatively impact its users.

Robustness: Ensures that an AI system continues to operate within safe limits upon perturbations. Must be robust or unforeseen events and adversarial attacks that can damage or manipulate such systems.
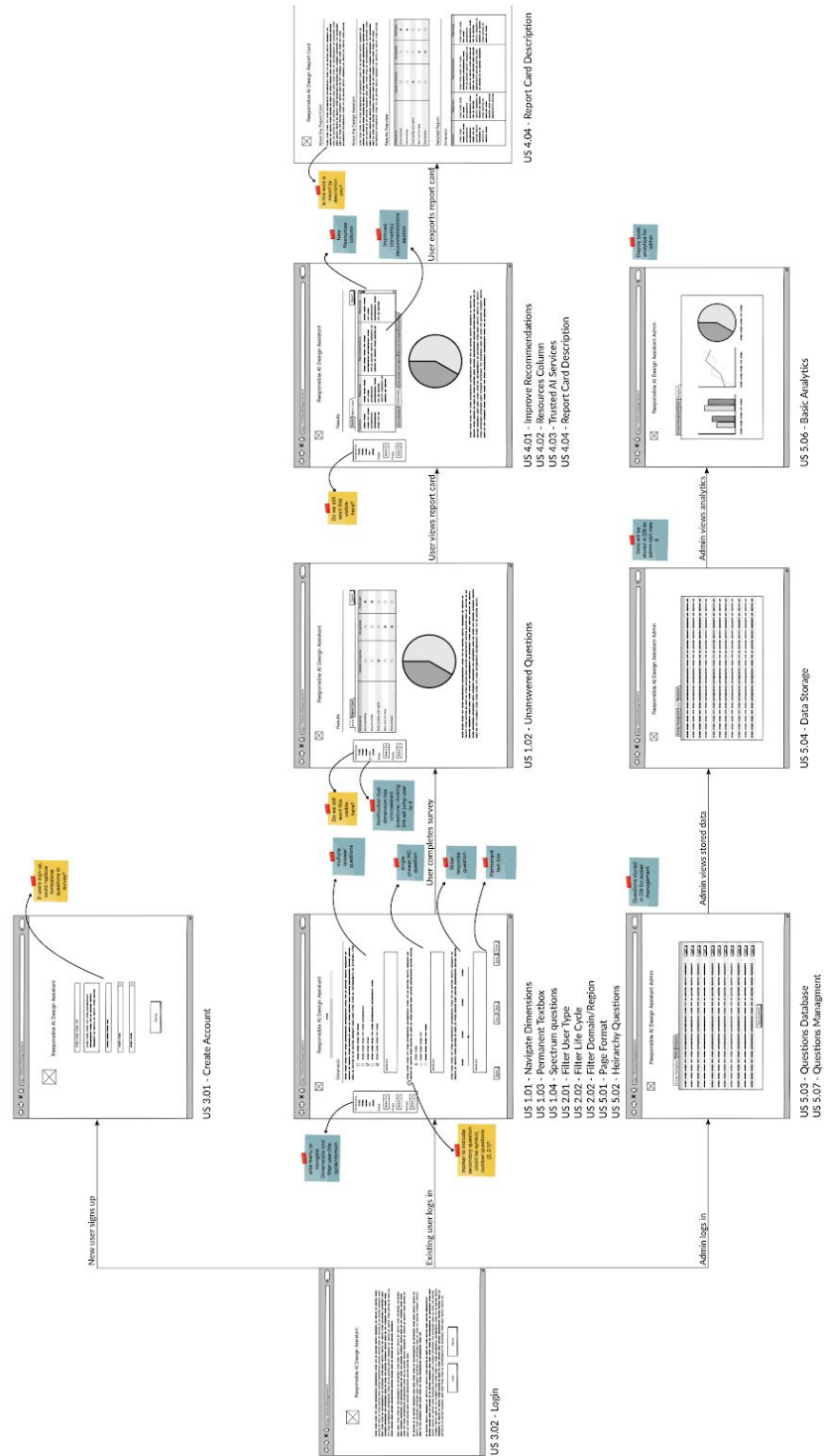
Scoring mechanism: Defined by AI Global on how each question response will contribute to the total score.

Security: How securely data is being stored. How well is it protected against unauthorized access

Timeliness: The obsolescence of the data for the task at hand

Storyboarding

High Quality pdf and individual wireframes available as separate files.

Use Cases or User Stories

US 1.XX - Survey user experience
US 2.XX - Question filters
US 3.XX - User account/login features
US 4.XX - Report card and recommendations
US 5.XX - Admin and back-end functionality

#	User Stories	Acceptance Tests
US 1.01	As a **User** I want to **have navigation controls that let me easily navigate to different dimensions** so that **I can conveniently go back and forth between questions**	● Navigation controls for switching dimensions exists when answering questions ● Clicking on controls lets the user switch between different dimensions
US 1.02	As a **User** I want to **Scroll to unanswered questions** so that **I can conveniently find questions I still need to complete**	● Unanswered questions are kept track of ● There is navigation to jump to these unanswered questions
US 1.03	As a **User** I want to **have a Permanent text box for submitting my feedback** so that **I can give suggestion on what can be improved or note any struggles I had with certain questions**	● Feedback text box for each question is permanent
US 1.04	As a **User** I want to **have a spectrum instead of check boxes for certain questions** so that **I can better represent my responses**	● Give option to add a spectrum slider as a response to questions ● Store spectrum response as a scale from 1-10 instead of yes or no
US 2.01	As a **User** I want to **filter questions by user type** so that **I am asked questions specific to my role**	● Filter exists that lets user specify their role ● Choosing a filter limits what questions are shown to the user
US 2.02	As a **User** I want to **filter on life cycle type** so that **I can answer questions specific to the current life cycle of the project**	● Filter exists that lets user specify on life cycle type ● Choosing a filter limits what questions are shown to the user
US 2.03	As a **User** I want to **domain specific extensions** so that I can **filter questions based on domain and region**	● There are different extensions for specific regions ● Questions presented to the user are filtered to only questions specific to that domain
US 3.01	As a **User** I want to **create an account** so that I can **keep track of data for monitor my progress during each cycle**	● There is a signup page ● User can create accounts
US 3.02	As a **User** I want to **login to my account** so that I can **see what changes each time I fill out the questionnaire**	● There is a login page

US 4.01	As a **User** I want to have **improved recommendations on my report card based off my answers** so that **I have recommendations tailored to my need**	- User responses are used to generate the recommendations on their report card - A system for deciding which recommendations to show based off how questions are answered is in place
US 4.02	As a **User** I want to have **a suggested resources column in my report card** so that **I can conveniently find information on how to improve**	- There is a resources section in the user's report card - Links to resources are provided based on the user's recommendations
US 4.03	As a **User** I want to **have a list of Trusted AI service providers on the report card** so that **I know where to go to get help**	There is a list of Trusted AI service providers on the report card
US 4.04	As a **User** I want to **have a title and description added to the report card**	Report card has a title and description
US 5.01	As an **Admin** I want to **have the page formatted the same regardless of how many questions are on the page** so that **everything looks consistent**	Formatting on each page of questions looks consistent
US 5.02	As an **Admin** I want to **have hierarchy questions** so that **I can present the User with secondary questions depending on their responses**	- Specific questions can be marked as primary or secondary questions - Secondary questions are given to the user depending on their response to the primary question
US 5.03	As an **Admin** I want to **store all the questions in a database** so that **all the questions are neatly organized in one place**	- All the questions from the excel sheet are moved to a database - Table in database is set up for questions with an appropriate schema
US 5.04	As an **Admin** I want to **store the data from the Users** so that **it can be accessed later for other features**	User responses are stored in a database
US 5.05	As an **Admin** I want to **have appropriate security features for the data stored** so that **unauthorized access of the data is prevented**	Security features implemented
US 5.06	As an **Admin** I want to **have access to basic analytics** so that **I can make informed business decisions**	Analytics for how many views, region, how they linked to the site, how long they are staying on the pages for are tracked and stored
US 5.07	As an **Admin** I want to **have an interface to easily remove, edit, and add questions** so that **it is easier to modify the questionnaire**	- Admin page is set up - Page allows Admin to edit questionnaire questions

Technical Resources

- DigitalOcean
  - Existing cloud service that the client is willing to provide a license for
- SurveyJS
  - Existing application leverages the SuveyJS library to generate the primary source of the content
- Bootstrap
  - Open-source CSS framework that existing application is built upon
- MYSQL
  - Open-source relational database management system that will be used to replace current spreadsheets to interface with the application

Similar Products

- PwC Responsible AI Toolkit
  - Similar type of assessment survey
  - Breaks assessment down into three categories **Ethics & Legal**, **Performance**, and **Governance**
- Socionics Personality Tests
  - Similar type of assessment survey
  - Contains Questions that allow spectrum responses
  - Provides a report card assessing the user.

Open-Source Products

- Algorithmic Impact Assessment
  - Similar type of assessment survey designed to help assist and mitigate impacts associated with deploying an automated decision system
  - Existing application is built off this open-source project
- COVID Self-Assessment
  - Self assessment survey designed to recommend services based on answers given by the user.
  - Utilizes JavaScript for their backend and CSS for their frontend