# Proposal for challenge “THE ART IN OUR WORLDS”

➔ **Solution explanation**

Art is one of the reasons humans live to see, it sparks our imagination, helps improve our creativity skills, and gives us joy, with NASA moving its data to the cloud forming a massive archive of free and open data, this art became less accessible, not only that, it is more difficult for visually impaired people to seek that art of knowledge especially for those who are into astronomy and need these resources in their research.

Our solution focuses on being creative and accessible to all. That is why we would employ technologies such as VR and text-to-speech (TTS). To describe the solution more precisely, it would be a web application with a smooth interface. The interface would offer various options for the user to interact and enter their input. The user can choose to either type the input phrase as text or say it in an oral manner using artificial intelligence techniques. Then, the user would have the option to choose the way they experience the search results. They could read the document or hear it or see the images of the search result as normal 2D images or virtual reality world.

The motto is: The art in our worlds is for everyone

Our project consists of 6 main modules:

1. Deep learning speech-to-text model using transfer learning.
2. The application date is a mixture between images and text documents, therefore, a model that can create textual representation for images is needed to unify the data format we are using to search.
3. A machine learning model that retrieves documents based on document-to-document similarity using cosine similarity.
4. Using deep style transfer between most similar image to distort images creatively.
5. Deep learning text-to-speech model describing retrieved results.
6. Web Application integrating all those Machine learning models with interface.

**➔ Problem Definition**

The main issue is to make the solution accessible to all people and offer a creative, entertaining, and informative experience, more accurate and faster search.

**➔ Background and literature review**

The objective of this review is to try to improve the efficiency of the search over a large database and enhance the relevancy of the retrieved information to the input text It’s clear from the research reviewed that the document-to-document similarity is widely used in many applications in various fields. Many approaches were applied to solve this problem, some of them were not satisfying for this project due to lack of efficiency and latency.

➔   **Methodology (the science behind the idea)**

Most of the methodologiesused to solve this challenge are domain-related to computer science, specifically the field of artificial intelligence, in the next table you can find each technique used to solve each sub-challenge:

|  |  |
| --- | --- |
| ***Sub- challenge*** | ***technique*** |
| *Speech to text Transcription* | *Streaming transcription and batch streaming* |
| *Create textual representation for each picture* | *CNN* |
| *Checking Documents similarities* | *NLP models* |
| *Manipulating images* | *GANS* |
| *Text to Speech Transcription* | *TTS models* |
| *Web application & VR image representation* | *HTML, CSS ,JavaScript*  *Unity app* |

➔  **Value Proposition**

According to design thinking methodologies there are many types of audience like lovers, haters, and outcasts, so the application is designed to suit all people, at low cost and high speed.

➔   **Resources**

Data:

1. [NASA Open APIs](https://api.nasa.gov/)
2. [Browse datasets | NASA Earth Observations (NEO)](https://neo.gsfc.nasa.gov/dataset_index.php)
3. <https://data.world/datasets/space>
4. https://api.nasa.gov/planetary/apod?api\_key=1Yrcatua18pB3TDIz703wt5URDAFJd4NwtgdLZjY&count=50

Resources:

1. <https://www.researchgate.net/profile/Faisal-Rahutomo/publication/262525676_Semantic_Cosine_Similarity/links/0a85e537ee3b675c1e000000/Semantic-Cosine-Similarity.pdf>
2. <https://www.section.io/engineering-education/speech-to-text-transcription-model-using-deep-speech/#:~:text=A%20speech%20to%20text%20model%20is%20applied%20in,audio%20and%20producing%20word%20documents%20instead%20of%20typing>.
3. <https://machinelearningmastery.com/how-to-caption-photos-with-deep-learning/>
4. <https://towardsdatascience.com/calculating-document-similarities-using-bert-and-other-models-b2c1a29c9630>

➔   **What is the role of every member of your team in the solution?**

|  |  |
| --- | --- |
| **dule** | **Developer** |
| *Speech to text Transcription* | *Nora Ekramy* |
| *Create textual representation for each picture* | *Nour Kamaly* |
| *Checking Documents similarities* | *Nourhan Abdelkarim* |
| *Manipulating images* | *Abdelrahman Sayed* |
| *Text to Speech Transcription* | *Ayman Hasan* |
| *Web application & VR image representation* | *All* |
| *Presentation* | *All* |

➔   **What is the strategy of the workflow you plan during the hackathon, and the steps you would take to finish your solution?**

We would work in parallel

Graphical user interface

Description automatically generated

➔   **Participants: Name, Mail, Phone number**

|  |  |  |
| --- | --- | --- |
| *Name* | *Mail* | *Phone number* |
| *Nora Ekramy* | *nora.ekramy11@gmail.com* | *+20 10 9903 5029* |
| *Nour Kamaly* | *nourkamaly26@gmail.com* | *+201142346682* |
| *Nourhan Abdelkarim* | *norhanabdelkarim1@gmail.com* | *+201069786425* |
| *Ayman Hassan* |  |  |
| *Abdelrahman Sayed* | *abdelrhmansayed01@gmail.com* | *+201275847533* |