ConnectNT

Bridging the gap between latest scientific advancements and the NT community

August 20, 2023

For Govhack Challenge

Using AI to better connect Northern Territory Science with the Community

Team Name

PiByte_warriors

Table Of Contents

Problem Statement

The Northern Territory aims to establish itself as an innovative society by effective communication of science within the local population. With more than 980 research publications and projects across 14 different research institutions in NT it has the fundamental resources to achieve this [1]. However, the general public is unaware of it. But with the current advancement of AI tools there is an opportunity to explore their use while compiling and reporting to the community NT's excellent track record in science and technology.

Challenges

The current measures fall short because at its nucleus it is a marketing problem which requires compiling advanced research from different institutions while providing engaging content to the layman population. Even with skilled professionals the solution is bound to be time taking and costly. However, the new generation of AI tools can expedite the compilation process and generate content on the go. This would open the doors to young adults, curious mums, and individuals of all ages eager to explore new horizons.

ConnectNT Platform

Introducing the ConnectNT platform that combines the power of AI tools into one platform especially focused on NT based research publications with the ability to translate scientific literature into local languages.

Compilation + AI Generated Short Summary

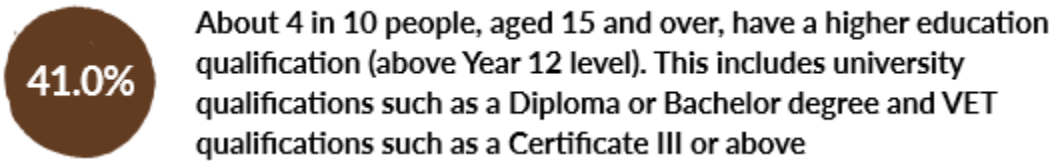


Figure 1: Story of our children and Young People 2021

Only 41% of people in NT, aged 15 and over have a qualification above Year 12 [2]. Furthermore, understanding research literature requires dense domain knowledge, proficient reading, language skills and demands time. With such a high barrier of entry, it's no wonder that curious individuals are discouraged from deep diving into the latest advancements in science.

Imagine if all of the latest literature from Northern Territories' top research institutes was available on one platform with a short AI generated summary in plain english(or other local languages). The reader can shortlist the papers they are interested in without having to read

through the entire publication.

Underlying Technology

Several open source machine learning models are available that can convert scientific research terms within a research paper into plain english. Such tools often employ Convolution Neural Networks to perform these tasks.
One such tool is the Open Accessible Summaries in Language Studies (OASIS) API. It's designed to bridge the gap between academic research and the general public by providing easy-to-understand explanations of scientific studies [3].

Virtual Assistant for Every Paper/Article

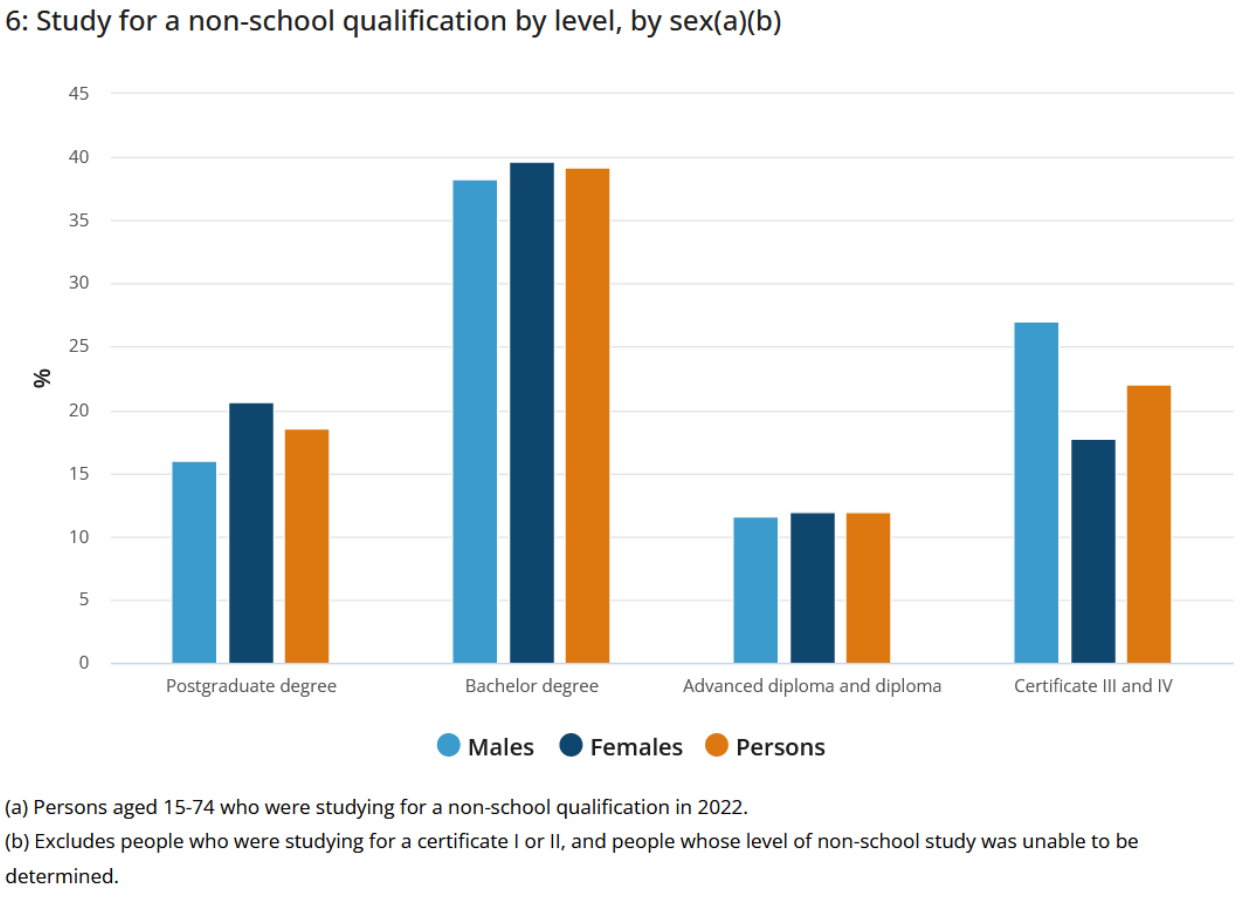


Figure 2: Education and Work, Australia, 2022

It is common knowledge that research skills need to be developed. With only 19% of 15-74 year old people studying postgraduate degrees where such skills are taught, It becomes difficult to

analyse the latest literature without external help [4]. Adding to this, reading literature from different domains like Environment, IT, Business etc. requires domain knowledge.

What if you could ask questions and get responses in your preferred language while reading the paper? This can be made possible with an AI virtual assistant that can answer questions regarding the selected paper on the platform. Once the paper has been selected through search it displays an option for a Virtual Assistant (Figure 7). Once clicked, two windows open side-by-side, one with the paper, and other with the virtual assistant to answer questions in your preferred language (Figure 8).

Underlying Technology

Advanced Natural Language Processing(NLP) techniques can be used to summarise the publication into a short and simplified summary. LSTM models like ChatGPT expose APIs that can be used by individuals or businesses to achieve this functionality. However, more accurate models also exist.

Smart Alerts

Institute
Batchelor Institute for Indigenous Tertiary Education
Center for Creative Futures
Charles Darwin University
Energy and Resources Institute
Menzies School of Health Research
Molly Wardaguga Research Centre
Northern Institute
North Australia Centre for Autonomous Systems
NT Academic Centre for Cyber Security and Innovation
Research Institute for the Environment and Livelihoods
The Australian Institute of Marine Science
The Cooperative Research Centre for Developing Northern Australia
The Drought Resilience Adoption
Innovation Hub

Figure 3: List of research institutes across Northern Territory

The ConnectNT platform compiles resources from all the research institutes throughout the Northern Territory (Figure 3). What's more is that the users have the ability to subscribe to alerts based on organisations, domains, keywords, tags and recommendations by researchers(Figure 5).

Underlying Technology

This system at its core is an AI based recommender system which uses big data to anaylse the most relevant keywords for NT and push the alerts based on those keywords for selected users. Furthermore, the matrix factorization algorithm is a popular choice to implement recommender systems.

Translation into NT Specific Languages

Figure 4: Story of our children and Young People 2021

The NT region is unique for its culture and languages. According to the snapshot of the Northern Territory from Australian Bureau of Statistics, only 57.3% of NT dwellers speak English at home as of 2021 [5]. Figure 4 shows a similar trend. After English, a lot of languages like Greek, Nepali, Hindi and other aboriginal languages share similar percentages of speakers. Such a situation requires remedial measures be taken to counteract the problem.

The ConnectNT platform will have the option to translate the text in the languages specific to the NT region. The text on the page including the name of the User Interface components will be translatable as evident in the figma prototype images number 5, 6, 7, 8 and 9.

Underlying Technology

A user preference database will be employed, and a recommender system algorithm will be utilized to suggest the most pertinent research articles or topics to users. At its core, this system functions as a recommender system, employing big data to analyze the most fitting keywords for the Northern Territory (NT). Subsequently, alerts based on these keywords will be delivered to selected users.

Poster Generation

A great way to improve community engagement is through social media platforms. In light of this, the ConnectNT platform will have a page that can automatically generate simple posters using generative AI (Figure 9). Users can then select, download and post it on their social media platforms. Doing so will have multifold benefits.

1. Improved community engagement on social media posts can drive traffic to the website.
2. NT establishes itself as a community that celebrates research culture.
3. The posters will have logos for the associated institutes, in turn promoting these institutes as centers of scientific advancements.

Other Advantages

1. Once NT establishes itself as a research driven society - solution to interstate migration - job employment [6].

User Stories

User stories can be a great way to understand the different use cases that the ConnectNT platform supports.

User Stories
As a young adult **I want to** browse through the latest articles **so that** I can get relevant material for my essays/homework.
As a professional **I want to** get the latest updates for the health and technology sector **so that** I can keep up to date with the latest market trends and identify potential business opportunities.
As a curious mum **I want to** understand the relevant literature regarding prenatal care **so that** I provide my family with the best possibilities.
As a university graduate **I want to** show on my LinkedIn that I understand the relevant literature in my field **so that** I can be more employable or be discovered by employers on

LinkedIn.

Limitations

1. Currently research papers are not translatable because they are complex without losing context.
2. Relevant Licenses and registrations for the research papers from institutes required.
3. AI tools can use the data for their own to train the models which might not be permissible by the research institutes.

Future Extensions/Releases

1. **Gamification:** Build a streaks type system so that the users can be motivated to come to the website.
2. **Voiceover**: There is a way to use generative AI to convert text-to-speech with the voice that the user wants. For example, the voice can be of an aboriginal elder. The model can be trained using the territory ANZACS dataset [7].
3. Social media APIs integration to post content on click of a button.

Data Story

The ConnectNT platform addresses the challenge of communicating Northern Territory's rich scientific achievements to the public. Despite over 980 research projects, limited awareness prevails due to the complexity of research literature and barriers to entry. AI-powered solutions can overcome these challenges.

ConnectNT combines AI tools to compile and simplify research, making it accessible. An AI-generated summary accompanies each paper, aiding comprehension. The platform's Virtual Assistant answers questions in real-time, enhancing understanding. Smart Alerts personalize content, enabling users to subscribe based on interests.

Language diversity is accommodated through translation features, catering to NT's unique linguistic landscape. Additionally, users can create shareable posters with generative AI, fostering engagement on social media. These efforts foster community engagement, positioning NT as a research-driven society.

Limitations include the complexity of translating research without context loss and compliance with research licenses. Future extensions encompass gamification, customizable text-to-speech, and social media integration. Through ConnectNT, the NT transforms into a hub of science accessibility and engagement

Appendix

Figure 5: Alerts Page

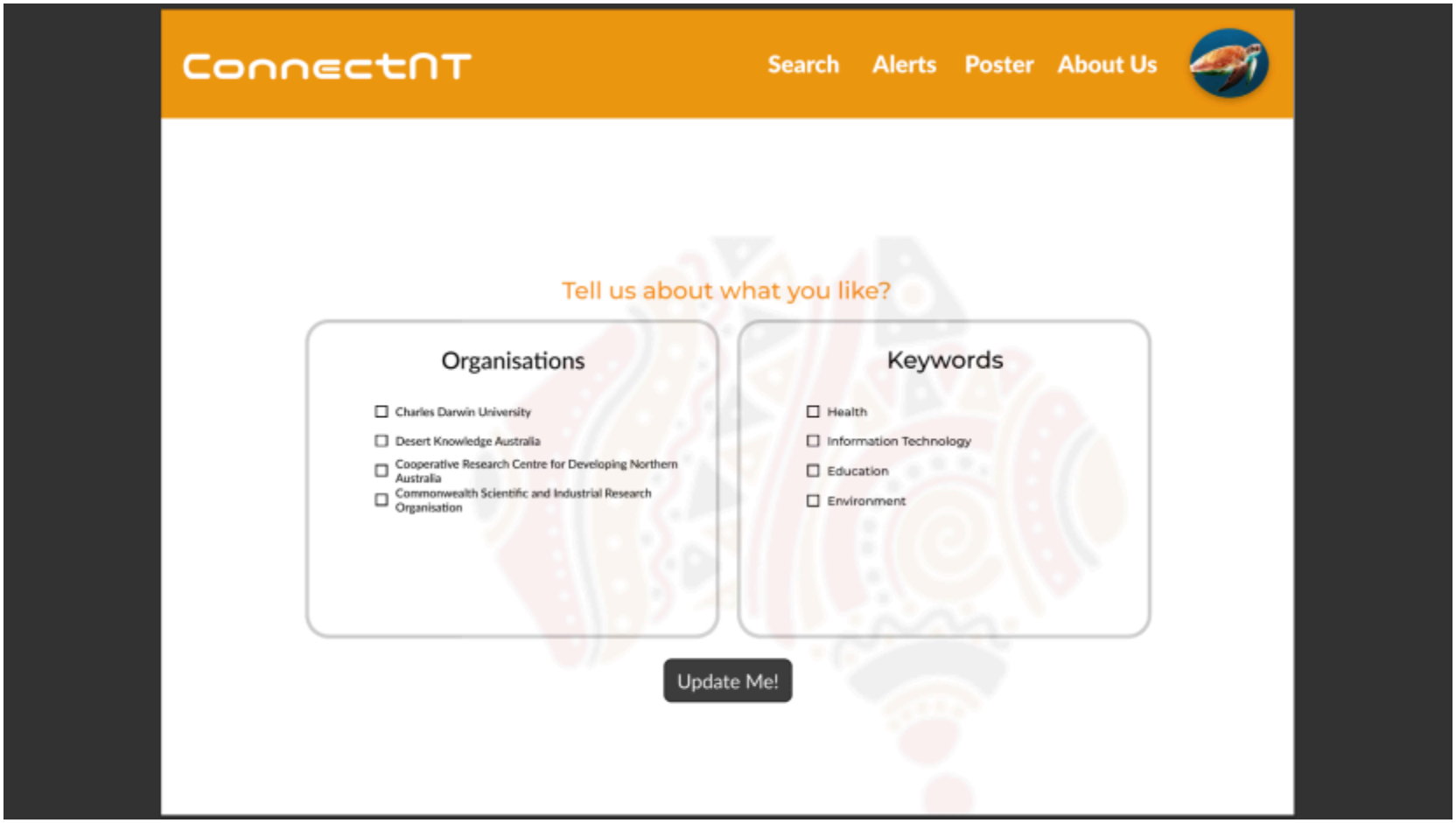


Figure 6: Search Page

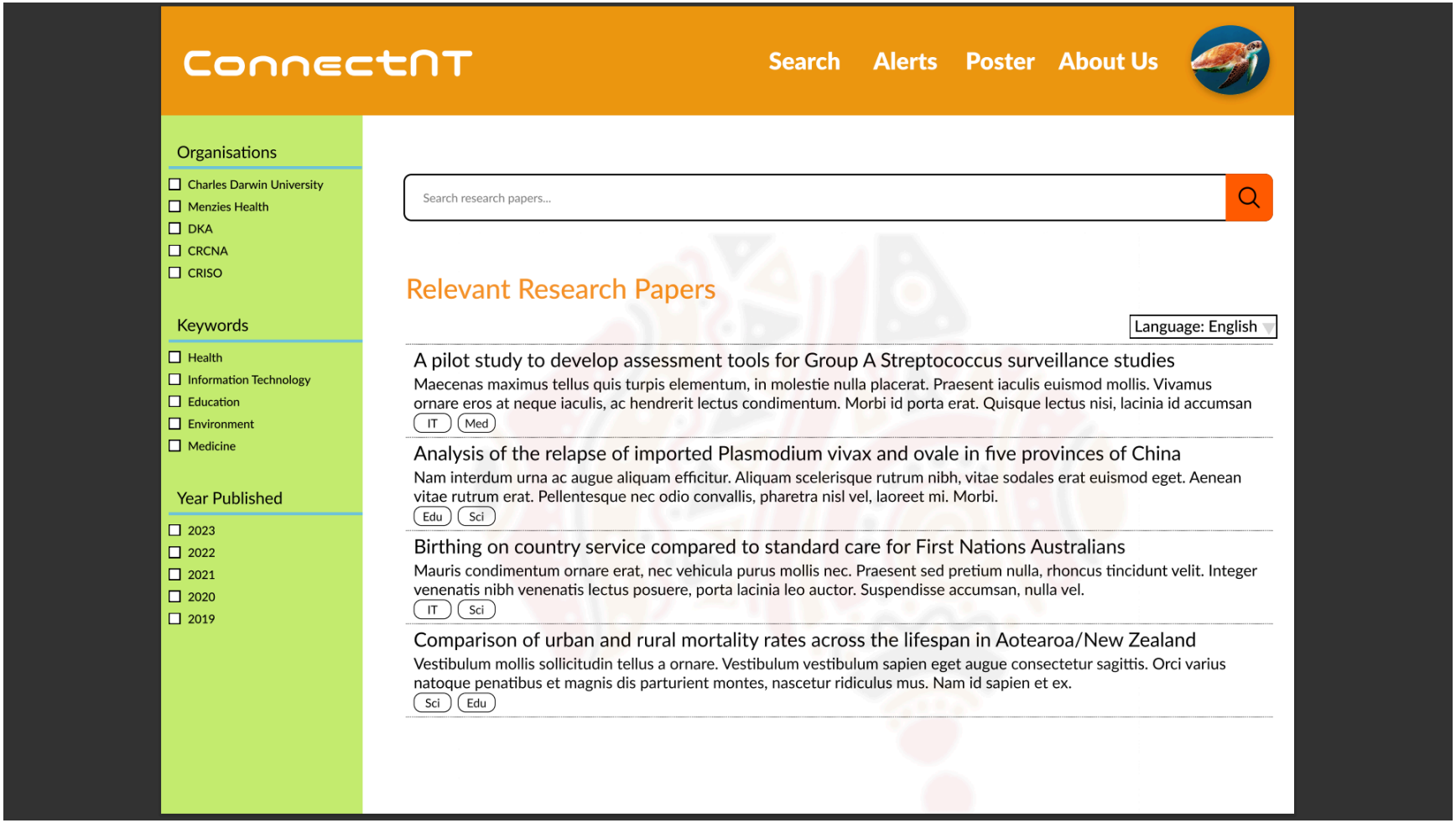


Figure 7: Detailed Research Paper Page

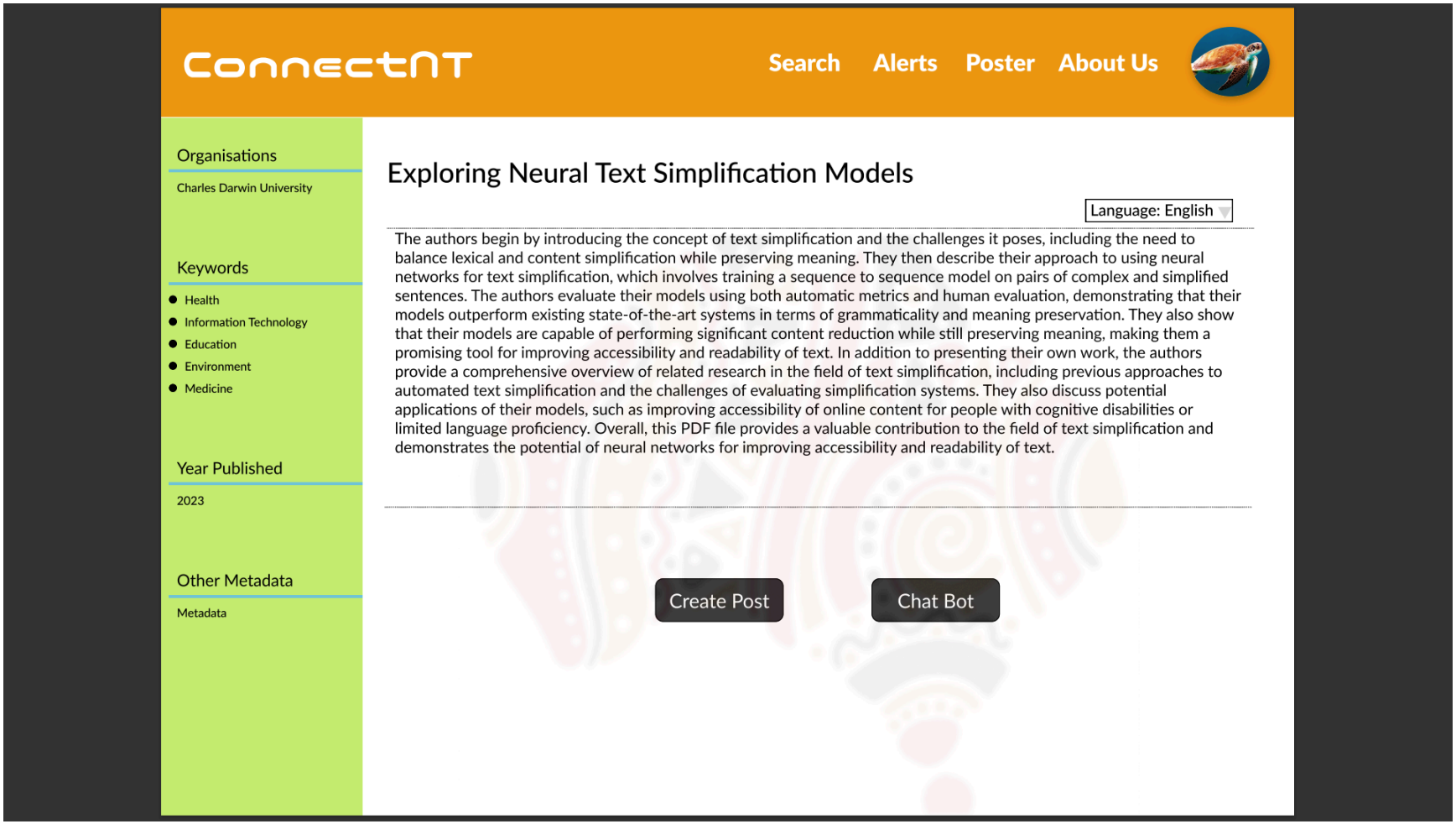


Figure 8: Virtual Assistant Page

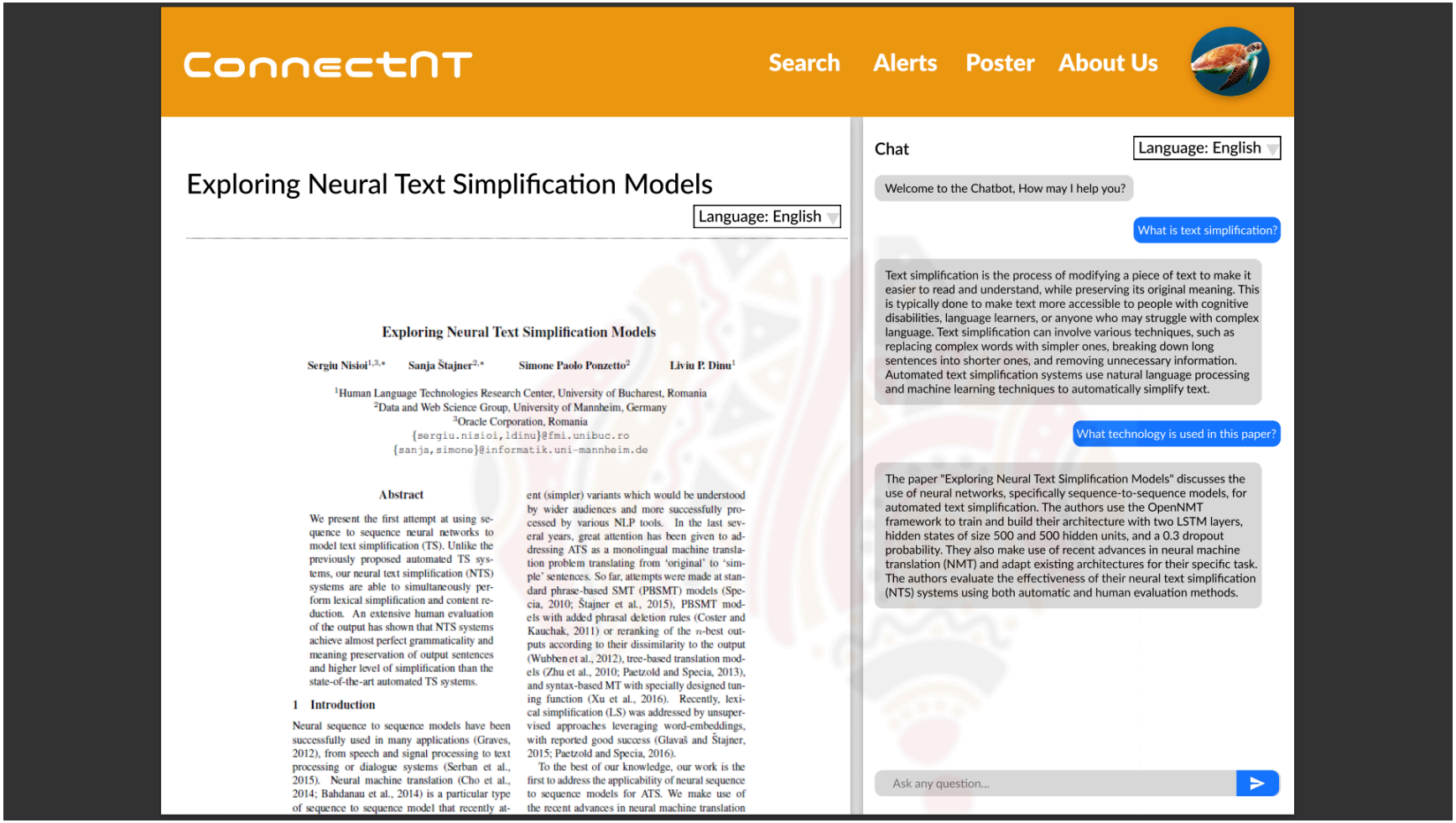
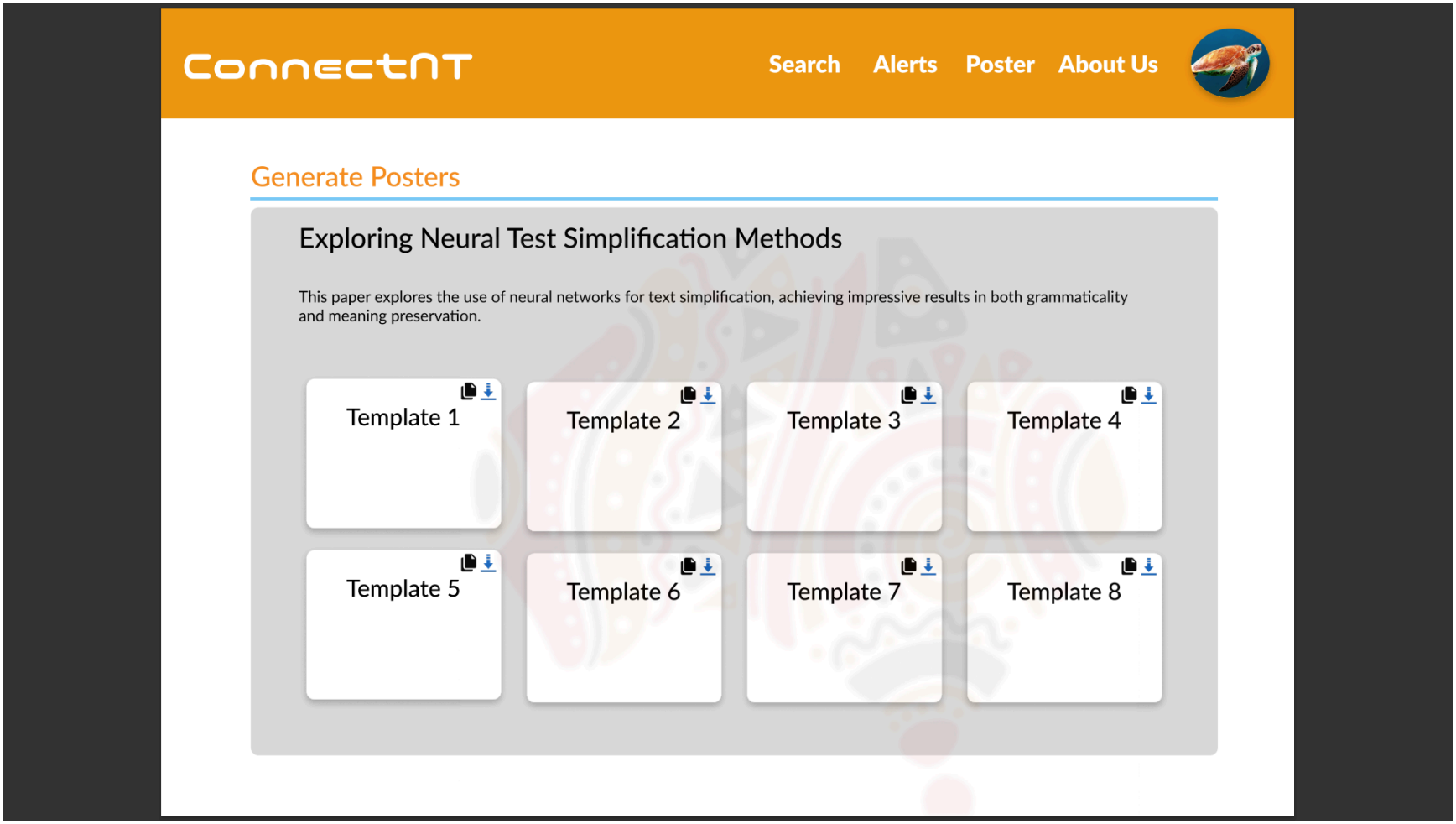


Figure 9: Poster Generation Page



References

[1] Manually going through every institute's research websites and gathering the data.

[2] Story of our young children and people 2021, https://cmc.nt.gov.au/__data/assets/pdf_file/0020/1061372/story-of-our-children-and-young-people-2021.pdf

[3] Open Accessible Summaries in Language Studies (OASIS), **https://oasis-database.org/about**

[4] Education and Work, Australia 2022. Australian Bureau of Statistics, https://www.abs.gov.au/statistics/people/education/education-and-work-australia/latest-release

[5] Snapshot of Northern Territory. Australian Bureau of Statistics, https://www.abs.gov.au/articles/snapshot-nt-2021#:~:text=Zealand%20(1.7%25)-,Language,%25

[6] Regional internal migration estimates, provisional. Australian Bureau of Statistics, https://www.abs.gov.au/statistics/people/population/regional-internal-migration-estimates-provisional/latest-release

[7] Territory ANZACS, https://data.nt.gov.au/dataset/territory-anzacs