**Conceptual EERD Generator Using NLP**

**Problem**

**Introduction to problem**

User stories serve as the backbone of communication between the stakeholders and the database developers and are used to determine the system requirements and its behaviour

One of the challenges in database design involves converting user stories into a conceptual entity relationship diagram (EERD) since this is a point in the design phase that is prone to human error and often results in discrepancies between the user’s requirements and the database design.

Furthermore, this conversion has always been a manual process an this can lead to an inefficient database design (e.g.: poor relationship definitions, multiple tables, etc.).

Finally, the manual method may not always be consistant because different team members can interpret the user story differently leading to variations in naming conventions and terminologies in the EERD.

**Problem background and statistic**

**Examples of the problem**

Consider an inventory management system for a supermarket. A possible user story in this case would involve tracking the stock levels of products, the information of the relevant suppliers, and the cost of the products.

The development team would have to go throught the manual process of translating the user story into an EERD in order to to design the database structure but this may result in inefficient database structures and poor relationship forming between entities (this can cause difficulties in retrieving and processing data). Moreover, this can be very time consuming, especially when it comes to large-scale projects.

**Attempted solution of the competitors**

Currently there are various database design tools available, but they have several limitations. Firstly, most existing tools aim to speed up the process of creating EERDs by using templates, drag-and-drop features, error checking, and more. However, these tools still rely on manual input, meaning that the developer still has to interpret the user stories themselves. As mentioned earlier, this is time consuming and inefficient.

Of the few tools that implement AI, only one was found to aid in generating ERDs rather than EERDs. The difference is that ERDs are better suited for for projects with simple data models, whereas EERDs are much more powerful at representing complex data models.

The proposed solution aims to solve these limitations by offering AI-driven user story understanding, and adaptability to diverse user story contexts; this can lead to more efficient and accurate EERD generation.

**Proposed solution**

User stories can be converted in conceptual EERDs using an AI-driven application that seamlessly. This solution would involve using NLP and machine learning to understand user stories, identify important elements such as entities na their relationships, generalisations, specialisations and multiplicities. The Conceptual EERDs can be presented in a user-friendly interface, allowing the user to seamlessly modify the diagram.

**Target audience**

**Database designers/architects software engineers**

It is essential for database designers and software engineers to efficiently develop a database’s design with reduced error. They are also responsible for making sure that the database aligns with the expectations of the user. The conceptual EERD generator application would help in this aspect by providing more accurate and consistent translations between user stories and EERDs.

**Project managers**

Project managers must esure that the software meets the business goals of the client; this may involve making sure that the project is completed within the given time frame and budget. In this case, the conceptual EERD generator would allow the database design phase to conclude much quicker.

**Resource requirements**

* Knowledge in natural language processing and machine learning
* Knowledge in front-end design to create an intuitive user interface and back-end design to handle the application’s functionality
* A dataset of user stories and Conceptual EERDs to use as training data
* Cloud infrastructure for model training and deployment

**Features of the solution**

* AI-based user story analysis
* Extraction of entities, attributes, and relationships
* Integration with common database design tools and their features
* Scalability to handle projects of various sizes