

# Log-based Data Analysis Unveiling the Past: User log-based Recommendation System for NLF Historical Newspapers

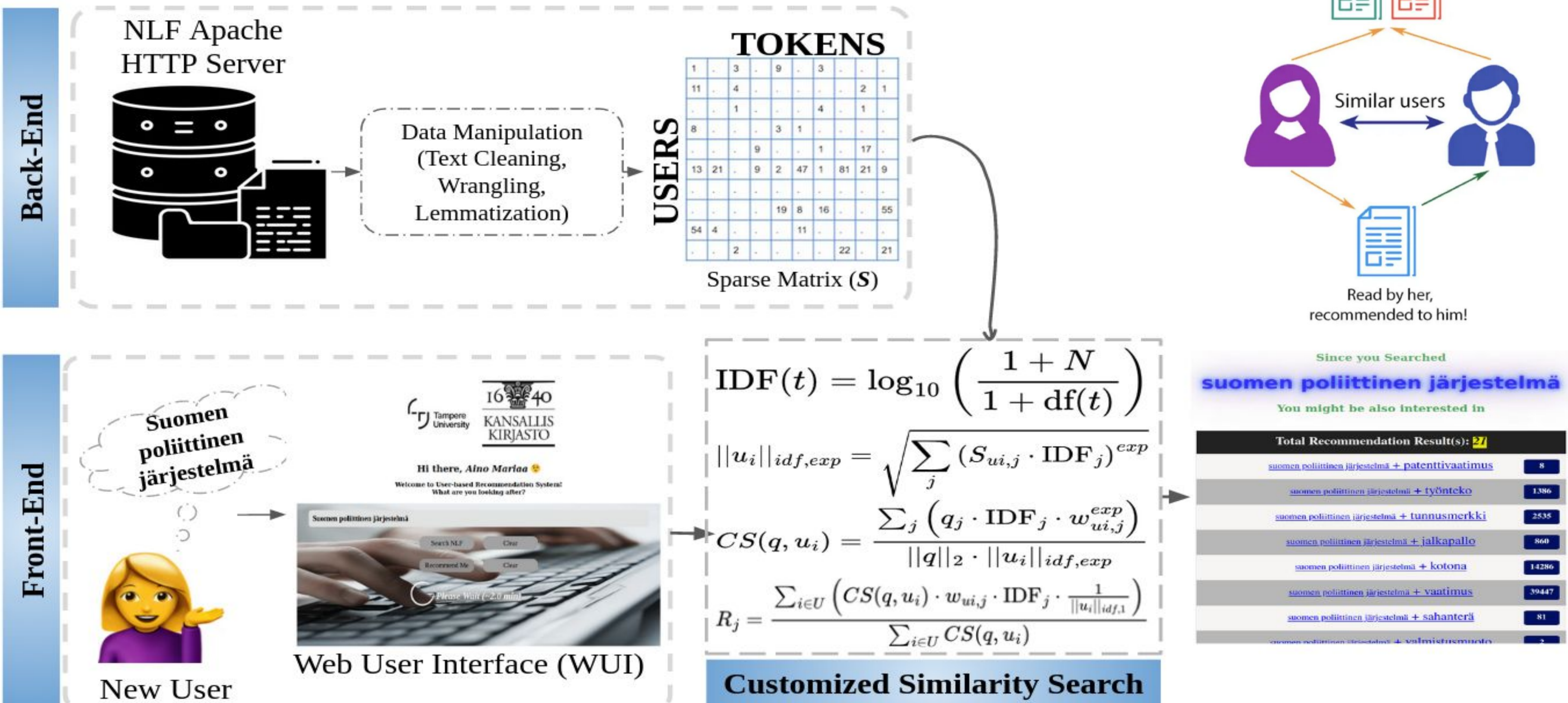
## Collaborative Recommendation in Retrieval

**Information Retrieval (IR)** aims to find relevant information from large datasets. In digital libraries, this involves retrieving precise documents from extensive archives. The National Library of Finland's (NLF) historical newspaper archives provide rich sources of information, crucial for researchers studying historical events and societal changes. However, the variability in historical language and the sheer volume of data pose significant challenges for traditional IR methods.

**Collaborative Recommendation** uses collective user preferences to enhance the relevance of search results. In historical newspaper collections, users with similar interests benefit from shared search patterns. This collaborative filtering improves the discoverability of relevant documents, enhancing overall search accuracy and user satisfaction. Our approach address this by user profile analysis, contextual relevance, and weight scoring.

**User logs** capture valuable information about user interests and interactions. By analyzing these logs, the system can tailor search results to individual needs and preferences. Session analysis provides insights into user behavior across sessions, helping to enhance the precision and *user-centric* nature of collaborative recommendations. This ensures that new users receive relevant and accurate suggestions based on collective user interactions.

## Recommendation System Architecture



## User in the loop: Satisfaction Measurement

### User Interaction:

Incorporating user Interaction and feedback in the search process is imperative for enhancing the precision and relevance of recommendation systems, along with optimizing user experience and system performance.

- **Iterative Refinement:** Our recommendation system keeps users informed, empowered, and invested by continuously refining them throughout the search experimentation.
- **User-Centric Approach:** This leads to better satisfaction by prioritizing user needs and preferences.

### User Feedback Integration:

- **Predefined Guided Tasks:** Users start interaction by entering search query prompts for specific given tasks in relevant fields, including, politics, history, or religion.
- **Post-Task Questionnaire:** Captures satisfaction levels regarding the relevance, clarity, and usefulness of the recommendations.
- **Behavioral Insights:** Feedback analysis is used to identify user interests, satisfactions and preferences, guiding towards more desirable system.

### Baseline Comparison:

A baseline comparison is crucial in developing a recommendation system for historical newspaper collections since it provides a reference point to evaluate the system performance and relevance of the obtained results to ensure the new designed system offers improvements over existing solutions.

- **Benchmarking:** Our Recommendation system is compared against the NLF DIGI website, accessible via the designed WUI.
- **User Reference:** Users familiar with DIGI can evaluate our system capabilities and recommendation results, providing a clear benchmark for satisfaction assessment.