BETTER Life

Project acronym: BETTER Life

Title: Bringing Excellence to Transformative Socially Engaged Research in Life Sciences through Integrated Digital Centers

Project number: 101071314

Deliverable 5.2

SER Self-Assessment Tool

Description: The SER Self-Assessment Tool helps researchers evaluate and improve societal impact using diagnostic surveys, CV analysis with advanced NLP models, and tailored resources.

Lead party for deliverable: EDUCONS

Document type: DEC

Due date of deliverable: 30. 11. 2024

Dissemination level: PU

Authors: Danica Mulić

Funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

BETTER Life PARTNERS:

CESKA ZEMEDELSKA UNIVERZITA V PRAZE (CZU)
PIC 999912570, established in KAMYCKA 129 SUCHDOL, PRAHA 165 00, Czechia

MARTIN-LUTHER-UNIVERSITÄT HALLE-WITTENBERG
PIC 999871539, established in UNIVERSITÄTSPLATZ 10, HALLE 06108, Germany

UNIVERSITA DEGLI STUDI DI CAMERINO (UNICAM)
PIC 999845737, established in PIAZZA CAVOUR 19F, CAMERINO 62032, Italy

ACEEU GMBH (ACEEU)
PIC 896865008, established in WILHELM-SCHICKARD-STR 14, 48149, MÜNSTER, Germany

EESTI MAAULIKOOL (EMU)
PIC 999857280, established in KREUTZWALDI 1, TARTU 51014, Estonia

DAUGAVPILS UNIVERSITATE (DU)
PIC 999830702, established in VIENIBAS STREET 13, DAUGAVPILS 5401, Latvia

UNIWERSYTET PRZYRODNICZY W POZNANIU (PULS)
PIC 999880463, established in ULICA WOJSKA POLSKIEGO 28, POZNAN 60 637, Poland

HELIXCONNECT EUROPE SRL (HELIXCONNECT)
PIC 893409868, established in 282A CAMERA 1, BUCOVAT 307352, Romania

UNIVERZITET EDUCONS U SREMSKOJ KAMENICI PRIVATNE USTANOVE (EDUCONS)
PIC 973147263, established in VOJVODE PUTNIKA 87, SREMSKA KAMENICA 21208, Serbia

Deliverable Information	
Title	SER Self-Assessment Tool
Deliverable number	5.2
WP number	5
Lead beneficiary	EDUCONS
Author(s)	Danica Mulić
Type	DEC
Dissemination Level	PU
Delivery date	30. 11. 2024

Project information	
Title	Bringing Excellence to Transformative Socially Engaged Research in Life Sciences through Integrated Digital Centers
Acronym	BETTER Life
Project no.	101071314
Type of Action	HORIZON Coordination and Support Actions
Website (Project)	www.betterlifehorizon.eu
Website (DCoE)	www.better-life-digital.eu

History of Changes			
Version	Date	Comment	Revised by
Version 0.1	27. 10. 2024	First version (ToC)	Andrea Andrejević Panić

Version 0.2	27. 11. 2024	Second version	Danica Mulić
Version 0.3	28. 11. 2024	Third version – revision and addition of images	Aleksandra Rankov, Vladan Rankov
Version 1.0	30. 11. 2024	Final version	Aleksandra Rankov, Vladan Rankov

Table of Content

List of Abbreviations

Abbreviation	Definition
ECR	Early Career Researcher
EU	European Union
HEI	Higher Education Institution
NGO	Non-governmental organisation
SER	Socially Engaged Research
SME	Small and medium-sized enterprises

Executive Summary

The Socially Engaged Research (SER) Self-Assessment Tool, developed within the EU BETTER Life project, empowers researchers to assess and enhance the societal impact of their work. Combining diagnostic surveys, comprehensive CV analysis powered by advanced NLP models and tailored resource recommendations, the tool provides a user-friendly platform for improving engagement practices.

Core features include a structured SER survey, a secure CV analysis module, and a personalised resource dashboard for tracking progress and accessing curated materials. Designed with advanced technologies and user-centric principles, the tool offers robust functionality, intuitive interfaces, and actionable insights to support researchers in refining their engagement efforts.

Built through an iterative, feedback-driven process, the tool addresses key gaps in engagement diagnostics while ensuring adaptability and long-term impact. It provides researchers with practical solutions to refine their practices and align with the EU BETTER Life Centre's mission to promote impactful, sustainable research.

1. Introduction

In the context of fostering more effective research engagement and dissemination, the development of a digital self-assessment tool has emerged as a pivotal initiative within the EU BETTER Life project. This tool aims to empower individual researchers by providing them with a tailored, interactive resource capable of diagnosing their current level of social engagement. By addressing this need, the tool aspires to enhance the overall quality and impact of research outputs across disciplines.

The envisioned tool serves as a guided, user-centric platform that not only evaluates the user's social engagement of their research processes but also offers actionable insights and practical recommendations. These recommendations encompass a variety of resources, strategies, and expert advice tailored to their specific engagement profile. Through this dual diagnostic-and-advisory approach, the tool seeks to support researchers in refining their skills, accessing appropriate resources, and aligning their practices with the overarching goals of the BETTER Life in promoting sustainable and impactful research practices.

This document outlines the methodological framework, developmental process, and projected impact of the self-assessment tool.

2. Key Components of the Tool

This chapter delves into the foundational elements that make the SER Self-Assessment Tool an innovative platform for enhancing societal engagement in research. All components below, Figure 1, form a cohesive system aimed at empowering researchers to amplify their societal impact. The tool is designed with a modular architecture to ensure functionality, adaptability and provide intuitive and personalised interaction. Each of these modules play an important role in achieving the tool's objectives of assessing and enhancing researchers' social engagement practices. Below is provided an explanation of the modules.

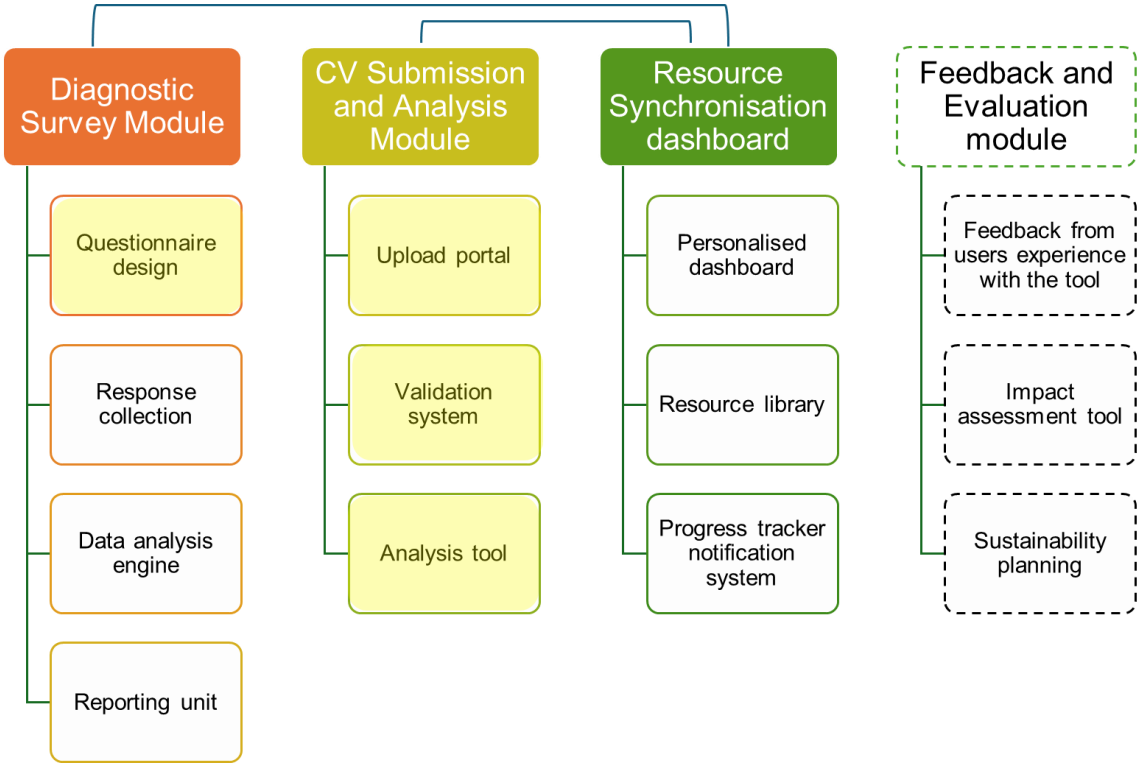

Figure 1 Self-assessment tool

2.1. SER Diagnostic Survey Module

The SER Diagnostic Survey module is central to the tool, structured to evaluate the current level of social engagement across key dimensions:

- Understanding SER: Educating users about socially engaged research.
- Planning and Relationship Building: Assessing users' strategies and collaborations.
- Co-Design and Co-Creation: Evaluating collaborative research practices.
- Implementation, Dissemination, and Assessment: Reviewing the effectiveness and outreach of research activities.
- Sustainability and Final Impressions: Highlighting long-term impact and user feedback.

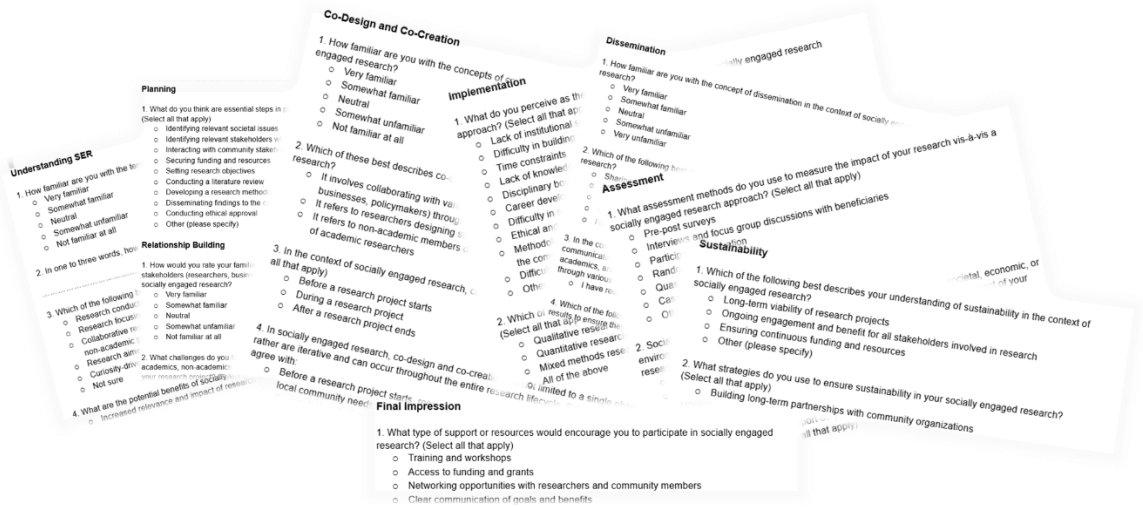

Figure 2 Illustration of questionnaire questions classified across key identified dimensions

The survey module utilises questionnaire design with tailored questions to capture dimensions of social engagement and provides real-time progress tracking and an analytics engine to process responses and provide the feedback on strengths, areas for improvement as well as the actionable recommendations.

2.2. CV Submission and Analysis Module

This module enhances the tool by allowing researchers to upload their CVs for detailed analysis supporting free-form formats. It includes:

- Upload Portal: A secure and user-friendly interface for CV submission.
- Validation System: Ensures completeness and compliance with format standards.

- Analysis Tool: Utilises advanced Natural Language Processing (NLP) techniques such as SciBERT[1] and PubMedBERT[2], to extract and analyse key data from CVs. This analysis complements the diagnostic insights by linking career achievements to engagement metrics providing in a such a way a comprehensive overview of researcher's impact.

2.3. Resource Synchronisation Dashboard

The dashboard presents a dynamic and interactive interface for researchers to track their progress, access resources, and receive notifications. Key features include:

- Personalised Dashboard: Provides a customised view of progress, resources and recommendations tailored to the user's specific needs.
- Resource Library: Offers a curated collection of materials aimed at enhancing SER practices.
- Progress Tracker: Visualises user progress through interactive metrics keeping users informed about completed milestones and pending actions.
- Notification System: Alerts users to incomplete tasks or milestones achieved ensuring timely engagement with the tool.

2.4. User Interface

The tool offered an intuitive, accessible interface with the following key pages:

- Home Page: Welcomes users with navigation, FAQs, and notifications.
- Examination Page: Allows CV uploads and survey completion.
- Progress Page: Displays progress metrics and upcoming steps.
- Diagnostic Survey Interface: Guides users through the SER survey.
- CV Submission Interface: Provides a secure environment for CV uploads.

[1] Beltagy, Iz, Kyle Lo, and Arman Cohan. "SciBERT: A pretrained language model for scientific text." *arXiv preprint arXiv:1903.10676* (2019).

[2] Gu, Yu, et al. "Domain-specific language model pretraining for biomedical natural language processing." *ACM Transactions on Computing for Healthcare (HEALTH)* 3.1 (2021): 1-23.

3. Development Process

This chapter outlines the systematic approach taken in creating the SER Self-Assessment Tool, starting with an in-depth understanding of Socially Engaged Research (SER) and careful initial planning. The process is divided into key phases from conceptualisation to implementation highlighting the modular design of the tools, the technical infrastructure and engagement strategies:

1. Initial planning and concept setup
    a. The core objectives and dimensions of SER were defined and user needs identified through stakeholder consultations
2. Tool requirements and architecture
3. Technical implementation
    a. Backend development .
    b. Frontend development
4. User engagement and integration of analytics
    a. Designed personalised engagement features, such as resource recommendations and progress tracking

3.1. Understanding SER and Initial Planning

The process began with an exploration of Socially Engaged Research (SER) focusing on identifying key metrics to evaluate engagement levels and addressing gaps in existing tools. Consultations with stakeholders helped assess user needs and define the core objectives and dimensions of SER. This phase resulted in a clear understanding of SER including its definition, key benefits and practical implementation.

SER was structure into categories representing the key dimensions of the social engagement such as understanding its concepts and benefits, research planning with stakeholder and ethical considerations, relationship building through trust and outreach, and co-creating research methods collaboratively. Additional focus areas include inclusivity, knowledge sharing and impact assessment with sustainability measures. Tools like diagnostic surveys, stakeholder mapping, and CV-based assessment were identified to support these dimensions.

During the planning phase, core modules, their interactions and the system's overall architecture were brainstormed to ensure alignment with user requirements and project goals.

3.2. SER Tool architecture

This diagram in Figure 3 illustrates the architecture of a comprehensive system for managing diagnostic surveys, CV submissions and resource synchronisation with a strong focus on user engagement and personalised support. It integrates a user-friendly interface with modular functionalities and robust database interactions to ensure data flow and management. The design utilises libraries for backend and frontend development supported by validation models and analytics to enhance usability and track engagement.

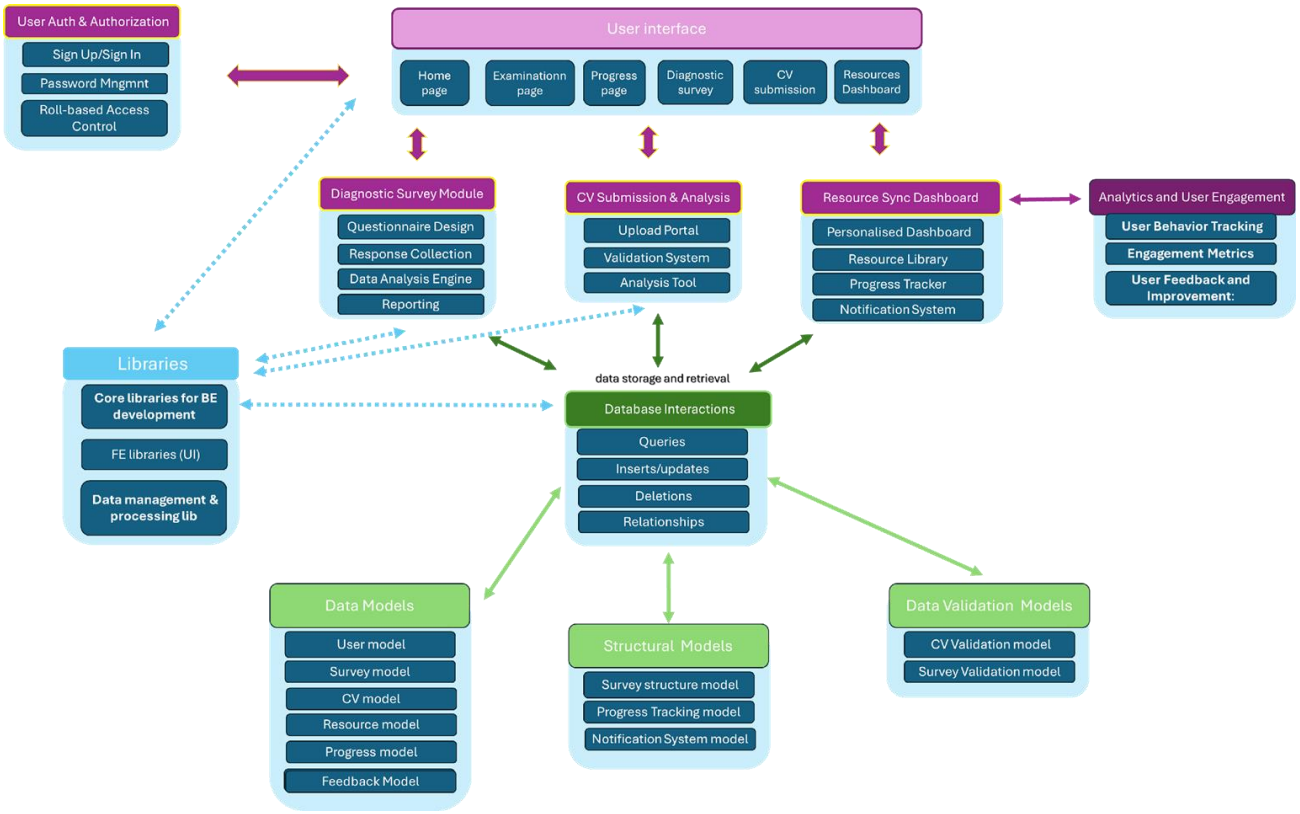

Figure 3 System (SER tool) architecture

The user interface consists of several key components designed to facilitate user interaction and improve accessibility of the diagnostic survey and resource tools. These include:

- Home Page as a central place for navigation, login, and support
- Examination Page for CV uploads, surveys, and submissions

- Progress Page that tracks user progress and next steps
- Survey Interface that guides users through surveys with tracking
- CV Submission secure platform for CV uploads and validation
- Resource Dashboard: Displays personalised resources and updates

The landing page for the SER Self-assessment tool is presented in *Figure 4*. It serves as an entry point to introduce the tool, explain its purpose and benefits and guide users toward taking action (access the tool). Here is the focus on attracting and informing potential users an encouraging them to engage with the tool. To access the tool User must register.

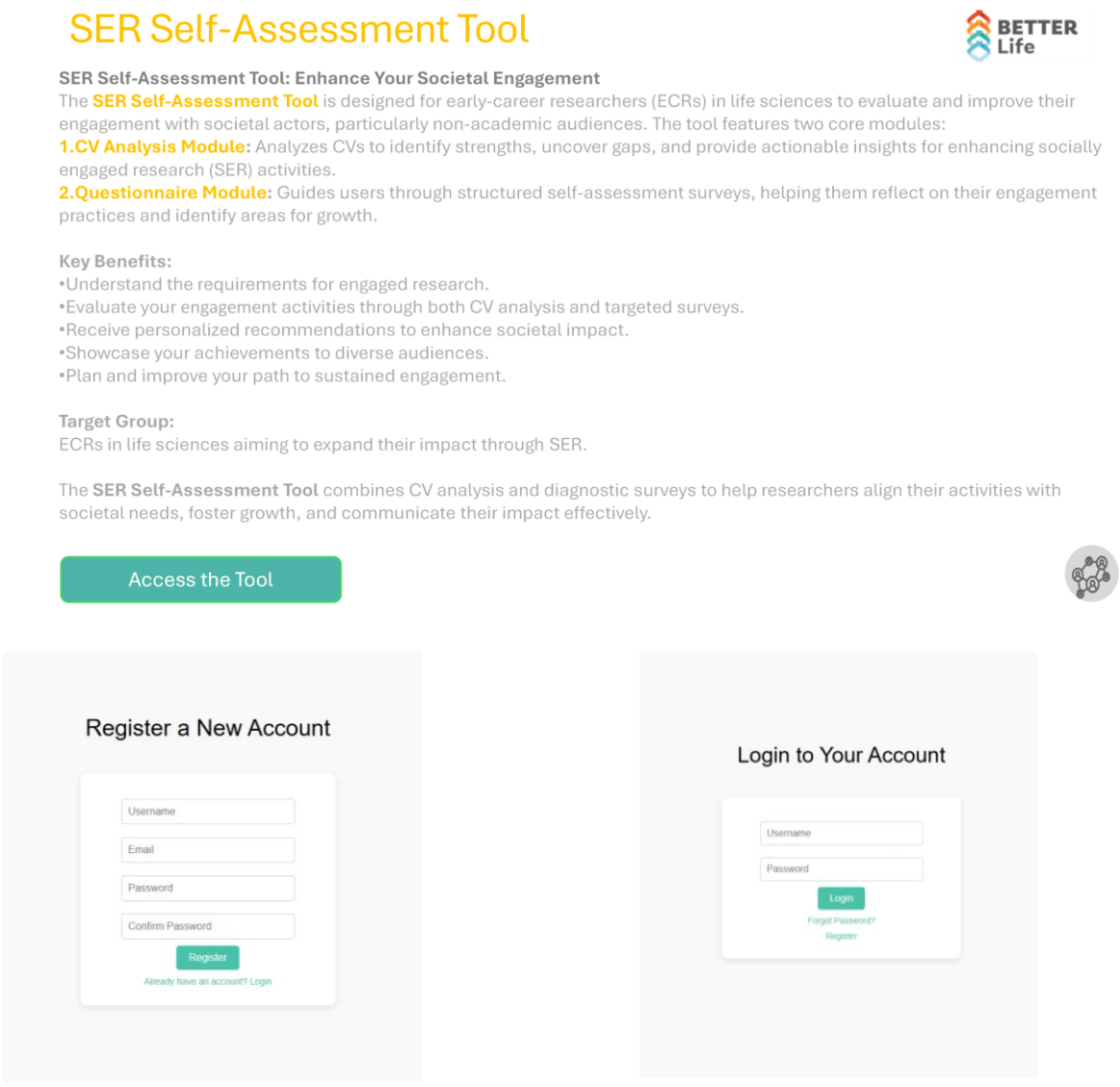

Figure 4 Landing page of a SER Self-Assessment Too and access to the tool through registration and loginl

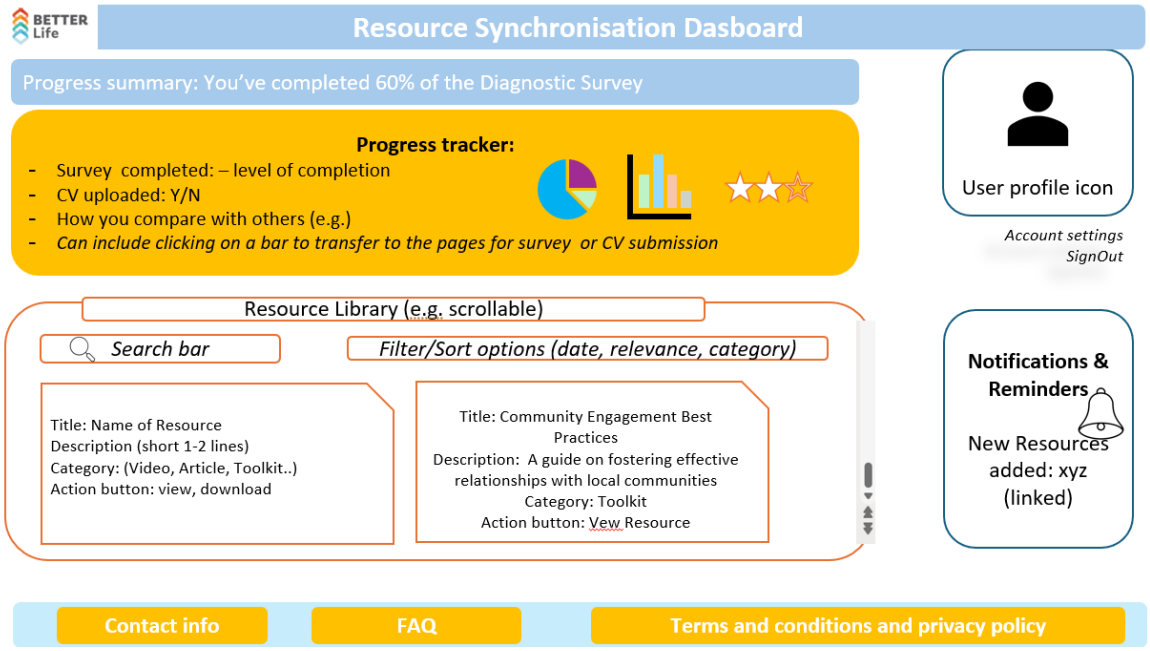

Figure 5 Conceptual prototype for the dashboard

Figure 5 presents a conceptual prototype for the dashboard designed to enhance user experience and efficiency. It is in the implementation phase, and it will manage user resources and progress in an interactive and personalised manner.

Elements, such as the Progress Tracker, Resource Library, and User Profile section are the core features essential for the dashboard's functionality and they directly impact user interaction and experience.

3.3. Technical Implementation

The technical implementation consists of two primary components – backend development and frontend development.

The backend of the SER tool was designed using state of the art technologies to process, analyse and interpret CV data and diagnostic surveys effectively. Here is an overview of the key technologies and methodologies used:

Django Framework: Django, is a high level Python web framework that simplifies building robust and scalable web applications with built-in tools for tasks like database management and user

authentication. It was chosen to provide security, scalability and modularity and for its compatibility with Python libraries.

Establishing the BETTER Life Django (BLDj) Project: The project framework was initiated and hosted on GitHub, establishing the core infrastructure for database setup, web tools, and a locally hosted server.

Integration of SciBERT and PubMedBERT: SciBERT was selected for its strength in processing scientific text, and fine-tuning was conducted to tailor it for SER assessment. PubMedBERT was later introduced to enhance domain-specific text processing, particularly for life sciences research profiles.

NLP for CV Processing: A CV processing module was developed, leveraging NLP models to analyse research profiles. Steps included tokenisation, lemmatisation, SER term mapping, and calculating TF-IDF scores for insights into SER engagement.

Clustering and Visualisation: Clustering algorithms, such as K-means and PCA, were integrated to categorise CV data, while a TF-IDF visualisation module was added to improve text analysis interpretability. Advanced techniques, including cosine similarity, were incorporated for enhanced data retrieval.

Testing and Debugging: Extensive testing validated SciBERT and PubMedBERT performance both standalone and integrated with Django. Issues such as SSL errors and image file processing errors (e.g., JPG, PNG) were resolved, ensuring stability and reliability.

Resource Synchronisation: A resource synchronisation dashboard was implemented, offering personalised progress tracking and tailored resource recommendations based on CV and survey analysis.

The frontend development focuses on creating a responsive, user-friendly interface for CV uploads, survey interactions, and resource access. Current efforts include implementing visual progress trackers, interactive dashboards, and real-time validation notifications, with ongoing work to optimise navigation, routing, and seamless integration with backend APIs for secure data handling and personalised recommendations.

End to end technical process is illustrated in Figure 6.

Figure 6 End-to-end technical process within the SER tool

3.3.1.How CV module works

The process of analysing researchers' CVs in life sciences and evaluating their engagement in SER relies on NLP techniques. These methods quantify and visualise SER related activities providing actionable insights into societal engagement. The CV module involves several key steps:

1. **Data collection and corpus creation**

- CVs are collected in PDF format (typically) and treated as individual documents within a larger corpus
- A predefined dictionary of SER keywords such as, for example, "public health," "community engagement," and "policy impact", "sustainability" is used to map relevant terms in the CVs.

2. **Preprocessing**

- Tokenisation: CVs are converted into plain text and broken down into individual words (tokens)
- Removal of common words: Common words like 'the', 'is', 'and' etc are removed to focus on meaningful content
- Lemmatisation: words are reduced to their base form (e.g. 'engaging' becomes 'engage' to ensure consistancy in analysis)
- Embeddings: Embeddings were generated using a fine-tuned version of PubMedBERT to capture and represent the meaning within the text. This approach leverages pre-trained contextual representations, enhancing the tool's ability to analyze and interpret domain-specific content effectively.

3. **SER term mapping**
- The dictionary of SER-related terms is applied to track occurrences and relevance within each CV

4. **Analysis and scoring**
- TF-IDF calculation: Term Frequency-Inverse Document Frequency (TF-IDF) scores are calculated to measure the significance of SER terms within a CV compared to the entire corpus
- Engagement analysis: High TF-IDF scores indicate strong engagement in SER-related activities. Researchers are compared based on their scores to assess individual and group engagement levels.
- Clustering: Algorithms such as K-means and PCA group CVs based on shared engagement patterns, offering insights into common areas of focus and expertise (Figure 8)

5. **Visualisation and interpretation**

Data visualisation: Bar charts and other visual tools illustrate the standing and distribution of SER-related terms across CVs (
- Figure 7)
- Resource interpretation: Patterns in SER engagement are identified, highlighting researchers with higher engagement and revealing gaps or areas for improvement.

6. **Resource recommendations**
- Based on the performed analysis, the tailored resources such as training materials, toolkits or other are suggested to help researchers address gaps and enhance their societal engagement.

7. **Progress tracking**
- The dashboard tracks user progress and provides real time updates on CV analysis, resource interaction and personalised recommendations to improve their societal engagement.

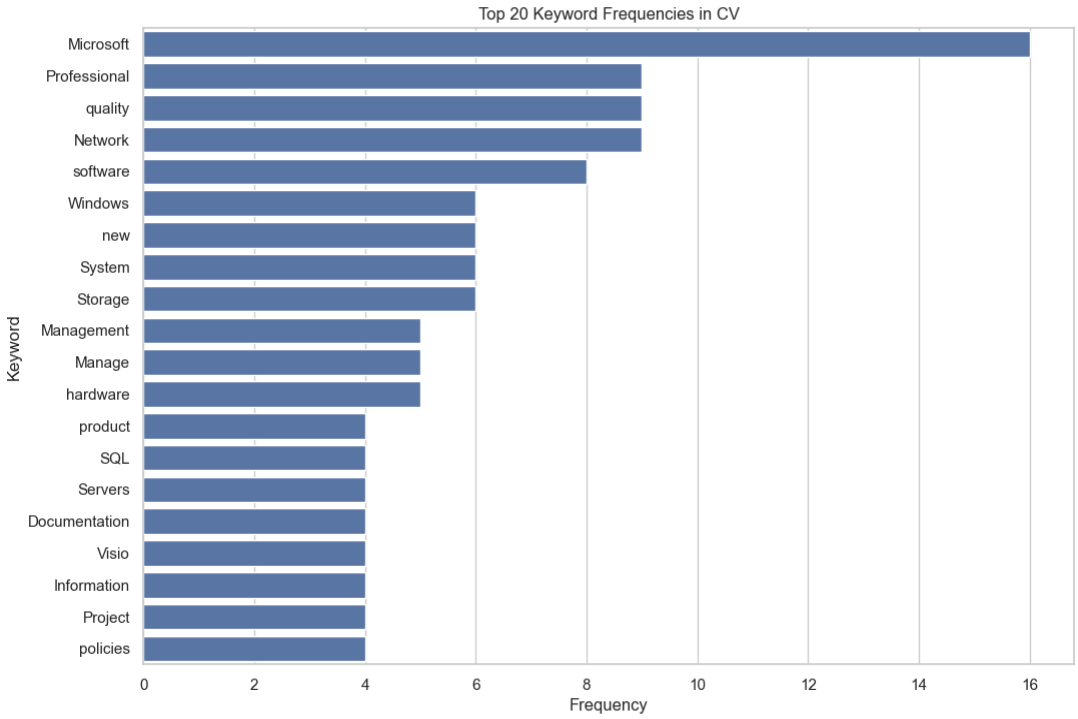

Figure 7 Presentation of keyword frequencies in analysed CV (left); Distribution

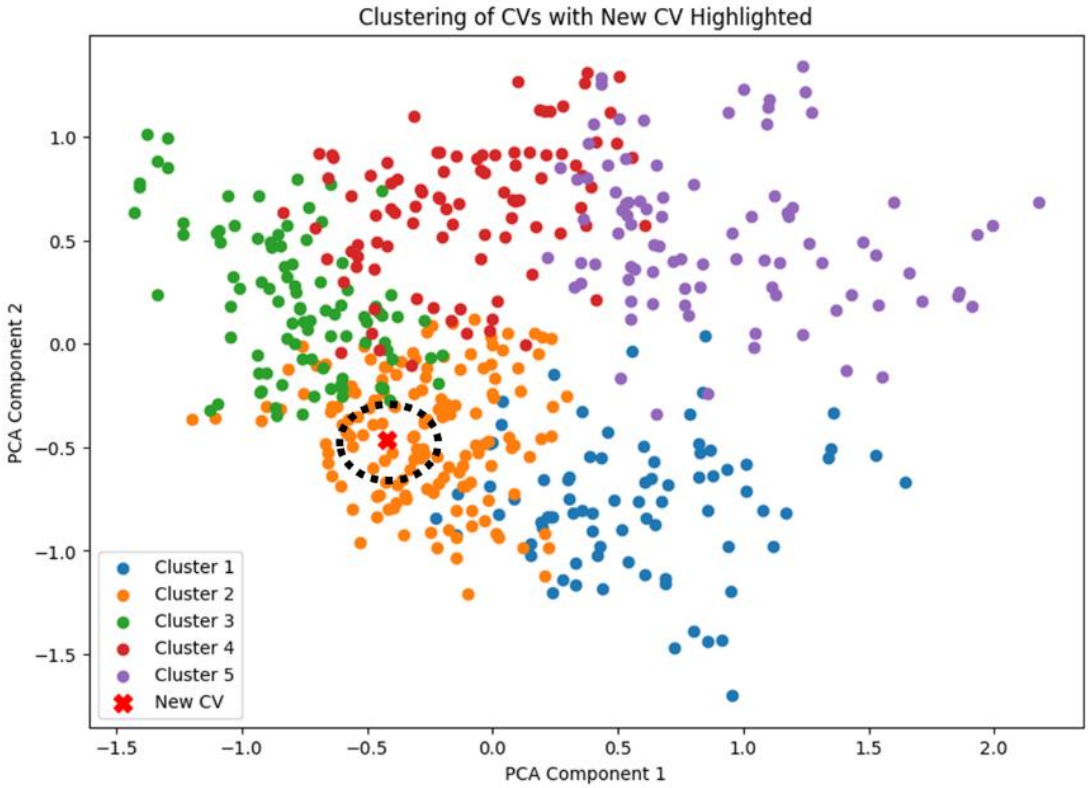

Figure 8 Illustration of assigning the new CV to the cluster

The whole process structured in this way evaluates individual engagement levels, but, in addition it supports researches in aligning their activities with SER goals. This promotes greater societal impact and collaboration.

3.4. User Engagement and Analytics

Usage metrics such as time spent on tasks, completion rates, and interaction patterns can be collected using front-end tracking tools like Google Analytics. This simple integration will provide insights into user behavior, helping identify areas for improvement and enhancing overall performance with minimal implementation effort.

3.5. Feedback and Evaluation

A feedback system is being implemented by integrating user surveys allowing users to share insights about their experience with the tool. Responses will be analysed qualitatively to identify themes and quantitatively to measure satisfaction and performance. This ensures continuous improvement aligning the tool with the user needs for a greater impact.

3.6. Sustainability Planning

To ensure the tool's long-term relevance and usability, a sustainability strategy was established:

- Regular updates to incorporate new research metrics and NLP advancements.
- Expanding compatibility with diverse CV formats and research fields.
- Continuous engagement with users for feedback and feature enhancement.

By following this detailed and systematic approach, the project team provides a robust, scalable tool that bridges gaps in research engagement diagnostics, providing researchers with actionable insights and resources to elevate their impact.

4. CONCLUSION

The development of the Socially Engaged Research (SER) Self-Assessment Tool represents a significant step forward in empowering researchers to assess and enhance their engagement with socially impactful research. By integrating advanced technologies such as SciBERT, PubMedBERT, and clustering algorithms, the tool offers a robust and scalable solution tailored to the unique needs of the research community.

The iterative development process has been grounded in collaboration, careful planning, and problem-solving, resulting in significant progress toward creating a user-friendly platform. Key components including diagnostic surveys, CV analysis, and resource synchronisation are nearing completion with the majority of the foundational work accomplished. While final refinements and testing remain, the tool already demonstrates its potential to address the gaps in existing diagnostics and provide researchers with practical personalised recommendations.

Looking ahead, the sustainability framework and modular architecture ensure the tool's adaptability to evolving research landscapes and technological advancements. Through ongoing feedback and iterative updates, the SER Diagnostic Tool will continue to support researchers in understanding, measuring, and amplifying their societal contributions.

This tool and progress made reflect the potential of technology and commitment to driving impactful research, aligning with the EU BETTER Life's mission to foster innovation and societal benefit through socially engaged research.