Further development of a Expert Finder System platform for the West German Genome Center and the Next Generation Sequencing Competence Network using the open source Profiles Research Networking Software

Profiles RNS is a research networking and expertise mining software tool (http://profiles.catalyst.harvard.edu/). Developed at Harvard, it is a open source ontology-based application. On June 18, 2020 Profiles RNS 3.0.0 was released and includes: (i) the Website source code (in C# for .NET 4.6.2) and precompiled binaries; (ii) the SQL Server (versions 2012-2017) database installation files; (iii) the documentation. By default, Profiles RNS uses the VIVO ontology which allows the addition of new sections as well as customization of the layout of the website. The original developers and the community working on Profiles RNS provides support through their google group (https://groups.google.com/g/profilesrns).

The project consists in the further development of a Profiles platform with one database and two websites, one for the West German Genome Center (WGGC) and one the Next Generation Sequencing Competence Network (NGS-CN).

The working language is English, knowledge of German is a plus. A translation of this text is available in German, the English version is the one of reference in case of doubts.

The project will start as soon as possible and should be delivered in 4 to 6 weeks.

In the following we will refer to:

* “the platform” (the whole Profiles application)
* “the software” (Profiles RNS 3.0.0)
* “the website” (as provided by the software, using Internet Information Services IIS with ASP.NET)
* “the database” (using Microsoft SQL Server as instructed by the software)

### Required skills

* Microsoft SQL: database installation and management; setup queries and scheduled database jobs; write and use scripts.
* IIS and ASP.NET: Website installation, configuration, and customization.
* Familiarity with: Resource Description Framework (RDF); Semantic Web; Ontologies.
* Communication (in English): clearly communicate progresses and delays with the product owner; clear written documentation for handover.

Project outline

Starting from a current prototype implementation and following the software documentation, the company will:

* Install the software on the designated server (virtual machine of the Hochschulrechenzentrum of the University of Bonn):
  + Either fix the current installation or re-install from scratch.
* Populate the database with the input profiles.
* Differentiate between experts from the different institutes:
  + Either using the “Groups” feature of the software or introducing a new ontology term.
* Optimize the website appearance and:
  + For the WGGC website: show only WGGC experts;
  + For the NGS-CN website: show all experts and search within the different centers.
* Optimize and debug the publication search algorithm of the software:
  + Implement the software optional feature for the ORCID identifier;
  + Fix current issues with names with special characters;
  + Fix current issues with publications not retrieved for experts.
* Implement missing features:
  + Geocoding, optional feature in the software.
* Make sure the websites are GDPR compliant (for example, cookie settings).
* Provide documentation, scripts and step-by-step guides for the most common tasks (including backups):
  + A suitable platform and/or VCS (e.g. gitlab and git) will be agreed upon.

### Project management

The company will work closely with the product owner (Dr. Antonella Succurro, WGGC Officer in Bonn) and the principal stakeholders (Prof. Joachim Schultze, WGGC spokesperson in Bonn and speaker for the NGS-CN; Dr. Mariam Sharaf, NGS-CN Officer).

The project can be performed fully remote. The servers are located in Bonn. The data of the Experts are to be treated confidentially and have to stay on the HRZ Uni Bonn servers.