SpectraWEB: A WEB Application for Storing and Retrieving Bacterial MALDI-TOF-MS Spectra

ID - 207

Dacierno Antonio[1], Facchiano Angelo[1], Mazzeo Maria Fiorella[1], Siciliano Rosa Anna[1]
[1]ISA-CNR, via Roma 52, 83100, Avellino, Italy

Motivation

MALDI-TOF mass spectrometric analysis of intact bacterial cells has the potential to be a new rapid method of identifying bacteria to species and also to strain level.

Several techniques can be used to classify mass spectra, ranging from expert based identification to statistical PCA based techniques, to Artificial Intelligence (AI) and Pattern Recognition techniques. We are interested in realizing a web based suite of tools for storing, retrieving, processing and classifying bacterial MALDI-TOF-MS spectra, by integrating and developing state of art algorithms for each processing step. In this paper we describe SpectraWEB, a WEB application for storing and retrieving spectra.

Methods

We started from the Entity Relationship (ER) model for the database (see the figure (a)). Bacteria are classified using the Genus, the Species and the Strain. For each strain, one or more Cultures are then considered. The User entity is used to model the users of the system; each user belongs to a Research Team. The Spectrum entity is used to represent actual spectra to be stored. Each spectrum, identified by means of a Spectrum ID, is related to exactly one culture and to exactly one user.

The Visibility mandatory attribute is used to model data privacy; a spectra spectrum can thus be private (only its owner can browse/download it), protected (only users belonging to the owners research team can browse/download it) or public (each registered user can browse/download it). The RawData attribute is clearly used to store the actual data.

Having in mind to use a Relational DBMS (RDBMS), the logical model has been obtained (figure (b)). The whole system has been realized using Java as coding language while the web application has been developed using Struts, a well known framework that implements the Model 2 approach, a widely adopted variant of the Model-View- Controller design paradigm. The on-the-fly generation on PDF documents containing spectra has been implemented using the well known free PDF-Java library iText. Eclipse has been used as developing platform while Exadel Studio has been used as building tool.

Results

Figure (c) shows some (simplified) use cases of the implemented system; here it is emphasized the differences between a registered user and an administrator of the system. We assume, in fact, that adding genera, species and strains is a task of the administrators while a registered standard user can just add new cultures and mass spectra. Moreover, administrator can add new users and research teams. A first use case worth to be analyzed is the one labelled Insert a Spectrum. If a user wants to insert a spectrum, he/she has to choose the related culture, so defining also strain, species and genus of the analyzed bacterial species; then he/she transmits data by uploading the data file. The system analyzes the file (several possible formats are defined) and translates data into a well defined internal structure that is finally stored as an array of bytes. Moreover, the system evaluates and stores some useful additional characteristics of the data (Xmin,Xmax, Ymin, Ymax, Points).

Another interesting use case is the one labelled Browse Data used by a registered user to retrieve and download mass spectra. The user browses data typically starting from a genus, then he/she browses species, strains and available cultures. Once the culture has been identified, the user is asked to choose among visible mass spectra, a spectrum being visible for the user if such an user is allowed to see/download it. The user also chooses if he/she is interested in downloading the data (a CSV file with raw data is then generated) or if he/she is interested in an image of the mass spectra (in that case a PDF document containing the mass spectra is generated on the fly).

Availability: http://medeaserver.isa.cnr.it:8080/SpectraWEB/

Image: http://medeaserver.isa.cnr.it/dacierno/BITS07_2.htm

Email: dacierno.a@isa.cnr.it