



EMBL-EBI Services Review 2015

On 26 to 27 March 2015 the review of the EMBL-EBI services took place. Twenty five experts, including three members of EMBL's Scientific Advisory Committee (SAC), formed the Review Panel, which was chaired by Michael Snyder from Stanford University (US). The Chair of EMBL Council attended as an observer.

Evaluation Summary

The EMBL-EBI Service Teams provide a wide array of valuable informatics services to the world. The EMBL-EBI services have had enormous impact on both medicine and industry and EMBL-EBI has reached a status of international pre-eminence. The quality and utility of the services provided as well as EMBL-EBI's training programmes and contribution to integrated activities in European Life Science research were all rated as outstanding.

Both the breadth and quality of the service activities are impressive, and the data are accessed by millions of people around the world. The services bring international recognition to both the EMBL and EBI. Whereas a decade ago these services might for the most part have been viewed as useful "add-ons" to international efforts (with some exceptions) it is now clear that the EMBL-EBI services are often leading these areas in both the information provided and the innovative ways employed to manage, access, visualise, and utilise the data. There are a number of services/activities that are unique to EMBL-EBI (e.g. Ensembl, Pfam, ChEBI, Metabolights, EMDB) and they build valuable software tools that are used worldwide (e.g. ChEMBL). They also help establish standards used by the rest of the international community. Although the services are broad, there is an impressive amount of collaboration and integration across the different activities. In 2011, the major preoccupation seemed to be how to handle the deluge of data while now (although that issue remains present) scaling plans are in place and the challenges are very different, demonstrating the consolidation and maturity of the Unit.

The service division will continue its mission of archiving, managing, and integrating the vast amounts and diverse types of data. It will also continue its outstanding training. The Review Panel endorsed the new movement of the EMBL-EBI services towards providing linkage to the clinical arena by establishing standards/platforms and reaching out to Biomedical Institutes located in EMBL member states to link their medical informatics with EBI resources and to establish at EMBL-EBI reference data sets of relevance to clinicians. EMBL-EBI has the opportunity to make a huge impact in this area. Similarly, the move into imaging by preparing reference samples and interfacing with the broader community was enthusiastically endorsed as a potentially high impact area serving a large community and one of great strength in Europe.

Much of the success of EMBL-EBI is credited to its leader Janet Thornton, Director of the EBI, its two Associate Directors Rolf Apweiler and Ewan Birney, and their predecessor Graham Cameron. Together, they are to be commended for transforming EMBL-EBI into a premier global bioinformatics centre. The Panel fully endorsed the recent designation of Ewan Birney and Rolf Apweiler as the new joint Directors of the EBI.

Response to the Panel's Recommendations

For the second time, the EMBL-EBI services were reviewed as "clusters" of related data resources rather than being classified for review by the individual Team Leader responsible. The Review Panel members were satisfied that the format of the review, which also included evaluation of the technical support teams and the training activities, permitted evaluation of the diverse areas in sufficient depth for them to form firm conclusions. Given that this required a Review Panel of exceptional size (25), I would very much like to thank the Panel members and in particular the



Panel Chair, SAC member Michael Snyder, for their excellent and efficient work and discussions and for the numerous helpful and insightful recommendations they provided. I note that the Panel recommended the generation of additional performance metrics and comparisons with other service providers for future reviews and I will discuss with my colleagues how these can best be generated. I also note here that this response does not include comment on the many valuable suggestions provided in the review sections on individual clusters. These will be brought to the attention of the cluster leaders and will be addressed by them together with EMBL-EBI's senior management. In my response I concentrate on more general issues.

The Review Panel noted that EMBL-EBI data services have attained a level of maturity in the provision of services and infrastructure that is of "industry standard". The data resources are all available 24/7 and major technical challenges like migration of all the resources from one commercial IT provider to another were accomplished with no unplanned downtime and a minimum of disruption. This has to be seen against a background of exponential data growth. The Panel notes, in contrast to the situation at the last review in 2011, that a plan to deal with this growth in a cost- and personnel-effective way is in place provided sufficient funding is available. I agree with the Panel that this represents a major step forward in the professionalisation of EMBL-EBI service provision during the last four years. In my view, this speaks to a remarkable, broadly-based, team effort and to the high quality of the overall leadership provided by Janet Thornton, Graham Cameron, Rolf Apweiler and Ewan Birney. I congratulate them and their many colleagues on this achievement.

Against this background I also find it gratifying that the Review Panel endorsed the recent decision to appoint Ewan Birney and Rolf Apweiler as joint EMBL-EBI Directors when Janet Thornton steps down later this year. I note and agree with the Panel's recommendation that it is wise to use this opportunity to look closely at the overall leadership structure of EMBL-EBI and to make any necessary changes to accommodate past growth and prepare for predictable future developments. A plan to accomplish this will be developed over the course of this year and will involve close prior consultation of all EMBL-EBI faculty by Drs Apweiler and Birney.

While dealing with the above-mentioned growth in data will remain a major task for the EBI in the upcoming period, it is by no means the only planned activity. The Panel endorsed the proposal for EMBL-EBI to work with Biomedical institutes to provide the open access reference biomolecular datasets against which medical professionals can do comparisons of confidential patient-derived data. I believe that this will be a crucial aspect of efficient precision medicine provision and an excellent opportunity for EMBL-EBI to work together with all of our member states in the individual way that will be required when dealing with medical data. The Panel also endorsed the plan for EMBL-EBI to work together with other experts to develop standards and methods for the construction and operation of imaging data resources. I find this to be a critical current challenge of considerable importance for the biomedical sciences and concur that EMBL-EBI's expertise means that it has a crucial role to play in this area.

The Panel noted that one of the two senior Team Leaders in the area of chemoinformatics had announced his imminent departure shortly before the review. I agree with their warning that it is urgent to provide stability in this important and rapidly growing area of EMBL-EBI's activity and have already, after receiving advice and input from the EBI leadership, endorsed a plan to deal with the consequences of this recent decision in both the short and the longer term.

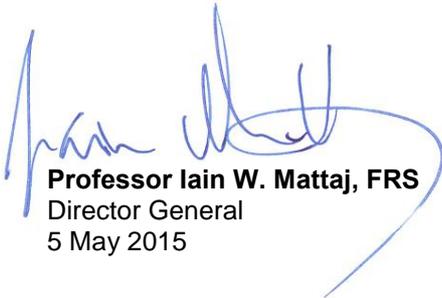
The Panel made some very valuable recommendations regarding the organisation of data in EMBL-EBI resources. For example, in the future it will be of value to find additional ways of integrating condition-dependent and condition-independent data resources. A second suggestion is to find ways in which stable and transient protein complexes can be presented as such in the protein structures databases. A final excellent recommendation is to see whether cell types, one of the fundamental units of biological organisation, can be used as an additional way of classifying datasets across EMBL-EBI.



EMBL-EBI engages in a considerable amount of training, which forms part of the overall EMBL International Centre for Advanced Training (EICAT). This was the first time we organised an external review of training activity, using EMBL-EBI training as a test case. The Panel was impressed by the range and quality of the training efforts aimed at teaching users of different levels of expertise how they can make better use of the data resources. The Panel praised those in charge of the training and noted that many individuals who are part of all the data resources participate in training. They recommend that we consider ways the training can be even better integrated into the thinking and organisation of the data service teams and I will work with my colleagues to follow up on this suggestion.

Finally, the Panel asked me to transmit two of their specific concerns to EMBL Council. The first is that they feel it is important that EMBL Council realise that, while ELIXIR is a tremendously valuable project whose realisation will be of great future benefit, at the present stage it is an activity that increases, rather than decreases, the costs, workload and responsibilities of the EMBL-EBI service staff and leadership. The Panel considered that it is crucial that EMBL-EBI funders are aware of this.

The second is a more generic worry. EMBL-EBI depends very significantly on external funds. Close to half of the total current EMBL-EBI funding is external. This level of funding reflects the tremendous effort put into raising resources by senior staff and also reflects the high regard in which the data resources are held by both funders and reviewers of funding proposals. The Panel points out however that external funding is never predictable and seldom stable. They see no alternative than to continue making use of it where possible to support EMBL-EBI activities, but feel that the member states have to be made aware of the potential danger of this source of instability for one of the world's most valuable and highly used collections of data resources.



Professor Iain W. Mattaj, FRS
Director General
5 May 2015