

EGI Virtual Machine Library: the Application Database

Tiziana Ferrari/EGI.eu Technical Director
on behalf of: Marios Chatziangelou/IASA

ELIXIR Workshop 12-13 May 2015



- A community driven central service that stores and provides:
 - ❑ **software solutions** (of *any* form, native software and/or virtual appliances), originated from almost every scientific area/discipline
 - ❑ reference of scientific **datasets** (*under development*)
 - ❑ the **programmers** and **scientists** responsible for them
 - ❑ the **publications** derived from the registered items (SW, VA & datasets)

Acts as **distribution** medium, for:

- ❑ native **software packages** (RPMs, DEBs, TARs) exposed through the UMD Community Repository to the public
- ❑ **virtual machine images** by using the HEPiX virtualization vmcatcher/vmcaster technology

- Foster the sharing of scientific codes, applications tools → application software library, communities of scientific codes
- Share and manage virtual appliances → reuse
- Link virtual appliances to datasets and publications → reproducibility of science

information retrieval

advanced searching mechanism (rated search results)
'faceted search' mechanism for refinements

quality of information

content tagging, rating, commenting per registered item contact expertise information
problem and comment abuse report centrally managed quality control

taxonomy

technical classification
scientific classification
tagging

dissemination of information

custom RSS/Atom news feeds
news e-mail subscription lists
user focused communication (messaging, requests, etc)
special dissemination tool for sending ad-hoc messages to scientists
'follow' button for receiving all the activity related to a registered item
dissemination features customizable through user preferences
sharing content with social networks



software distribution

repository for binary artifacts (lightweight release management process)
vAppliance management and distribution mechanisms

AuthN/AuthZ and security

advanced AuthN/AuthZ mechanisms ([simpleSAML](#))
integrated with [eduGAIN](#)
support for multiple accounts, [Federated Identities and/or EGI SSO and/or x509](#)
internally managed AuthZ, based on allowed actions, roles and permissions
everything is under [https](#) communication

relations...

... between all the entities listed below, are possible:

- software
- virtual appliances
- datasets
- persons
- virtual organizations
- sites / resource providers
- organizations
- projects

integration with AppDB

[RESTfull API](#), supports operations following a CRUD convention.
flexible API [stateless authentication](#) mechanism using **Personal Access Tokens** (no need for X509)
API documentation in good shape
or even, by adapting the [AppDB Gadget](#) (easy – copy & paste, one line of code – no technical skills required, you may get it [here](#))

AppDB already integrated with EGI services and beyond...

EGI GOCDB

for the list of sites and their end-points

Top-BDII

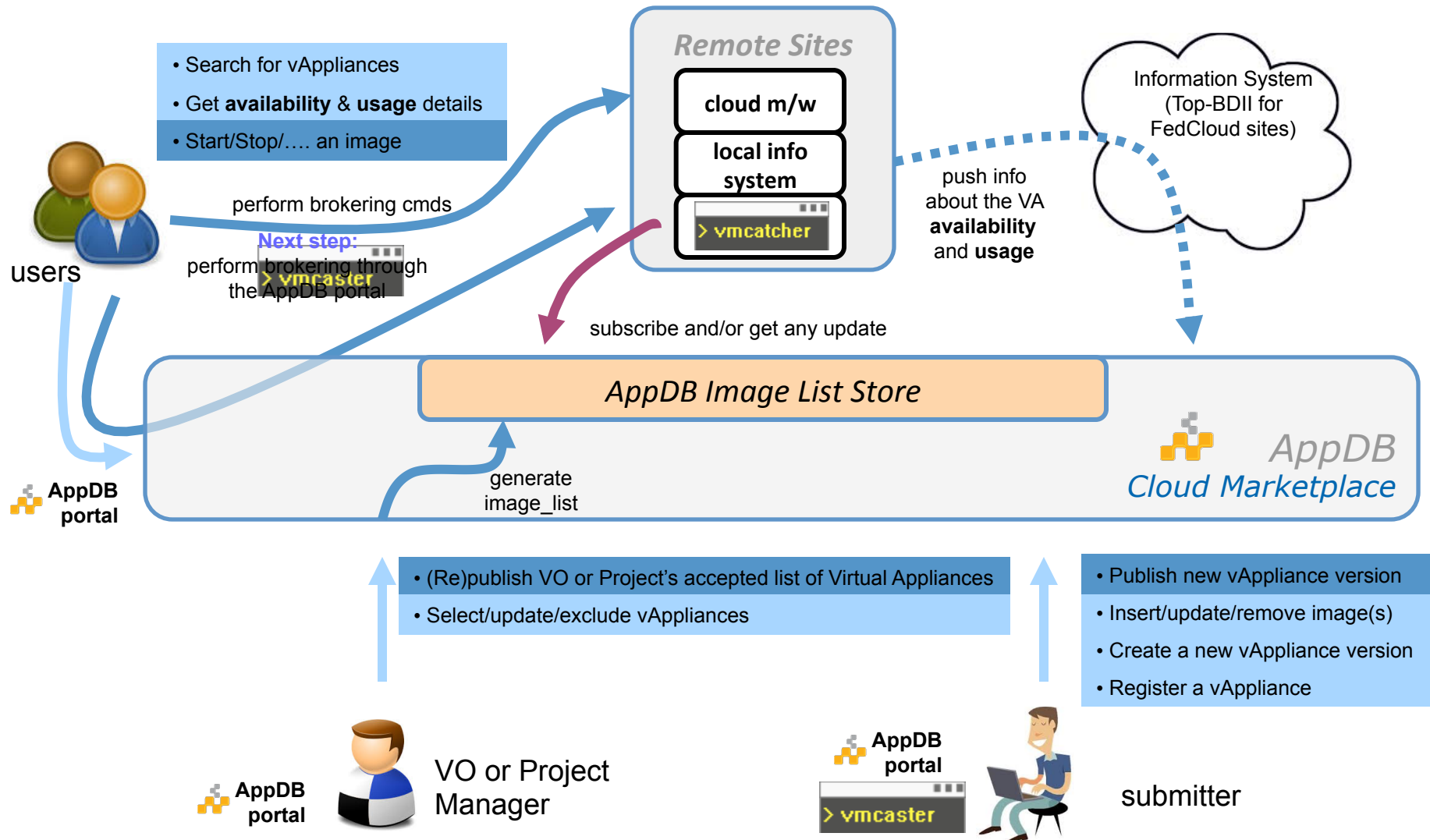
fetching sites dynamic information

Perun and EGI Operations Portal

for VO related details + Authorization attributes

OpenAIRE: retrieving FP7 and H2020 projects and linked Organizations, persons and in the near future... publications

- Holds, manages and populates “**Virtual Appliances**”
 - Virtual Appliance => bundle of one or more Virtual Machine Images ready to be used either on a cloud infrastructure (i.e. EGI FedCloud) or even to be downloaded for personal use
- Ability to the user to create “**Software Appliances**”
 - Software Appliance: a pair of a Virtual Appliance and a contextualization script which can be used for adhoc software installations at the Virtual Appliance during its boot time i.e. a base Virtual Appliance of a Centos6 and a contextualization script that downloads and installs the latest version of octave, on boot time.
- **VO or Project Managers:** able to **select** which of the registered Virtual Appliances, should be pushed to the sites of their responsibility.
 - The system takes the responsibility for **distributing** the “**chosen**” ones to the respective Resource providers/sites.
 - AppDB **gets feedback from the sites** that the Virtual Appliances **are indeed at the site side** and are ready to be used by the users.
- ... well, lets see some of these in a more graphical way.



- **Scope:** Provide information with regards to the ELIXIR reference datasets exist/replicated within the European Grid Infrastructure.

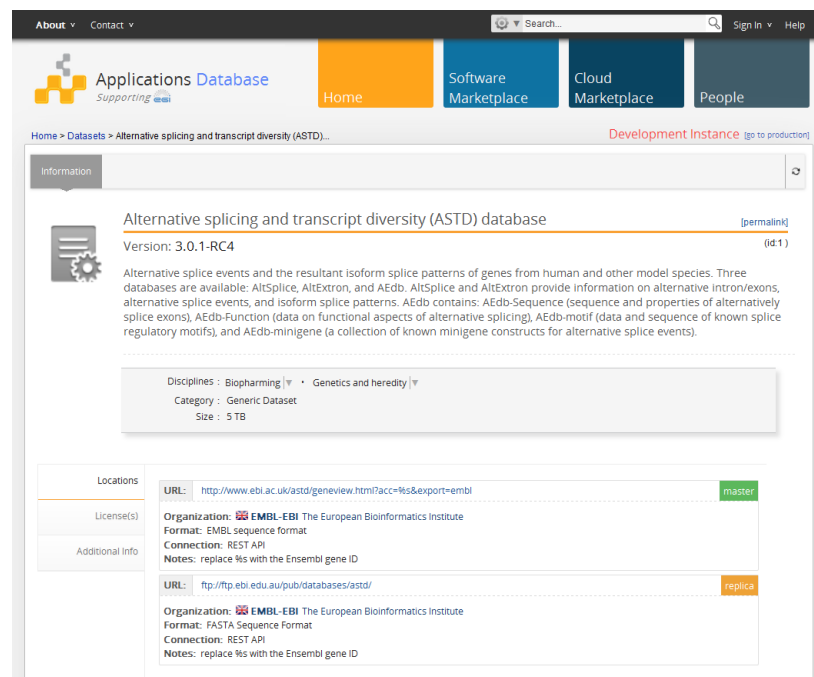
Activity supported by an on going pilot project, expected to be finalized in upcoming June.

Details **and credits**, can be found [here](#).

Tasks:

- Identify existing life science datasets in EGI
- Identify reference datasets for replication
- EGI AppDB extension to a dataset registry
- Identify and propose tools for data replication
- Analysis tools to work with data replicas
- Integration with ELIXIR Registry

The activity has been already endorsed by two ELIXIR nodes: **Greek** and **Slovenian**



The screenshot shows the Applications Database (AppDB) interface. The main content area displays the 'Alternative splicing and transcript diversity (ASTD) database' entry. The entry includes the following information:

- Version:** 3.0.1-RC4
- Description:** Alternative splice events and the resultant isoform splice patterns of genes from human and other model species. Three databases are available: AltSplice, AltExtron, and AEdb. AltSplice and AltExtron provide information on alternative intron/exons, alternative splice events, and isoform splice patterns. AEdb contains: AEdb-Sequence (sequence and properties of alternatively spliced exons), AEdb-Function (data on functional aspects of alternative splicing), AEdb-motif (data and sequence of known splice regulatory motifs), and AEdb-minigene (a collection of known minigene constructs for alternative splice events).
- Disciplines:** Biopharming | Genetics and heredity
- Category:** Generic Dataset
- Size:** 5 TB

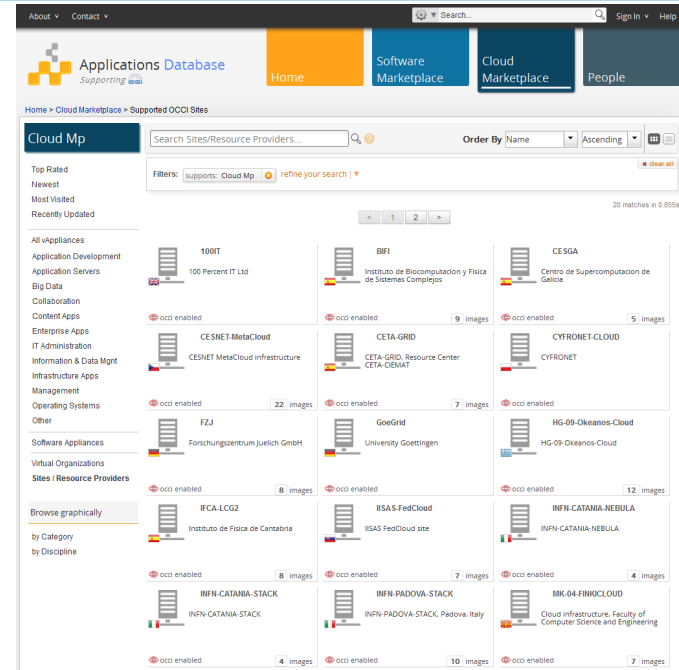
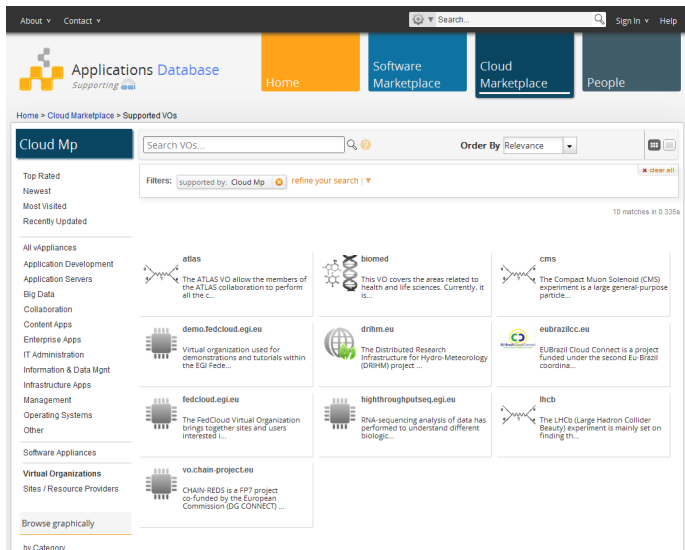
The 'Locations' section shows two entries:

- Master:** URL: <http://www.ebi.ac.uk/astd/geneview.html?acc=Hs&export=embl>
- Replica:** URL: <http://ftp.ebi.edu.au/pub/databases/astd/>

Both entries are provided by EMBL-EBI The European Bioinformatics Institute, in EMBL sequence format, and use REST API for connection. Notes for both entries indicate to replace %s with the Ensembl gene ID.

Some metrics:

20 Cloud providers already support the AppDB Cloud Marketplace



•10 Virtual Organizations and/or Projects are making use of the system

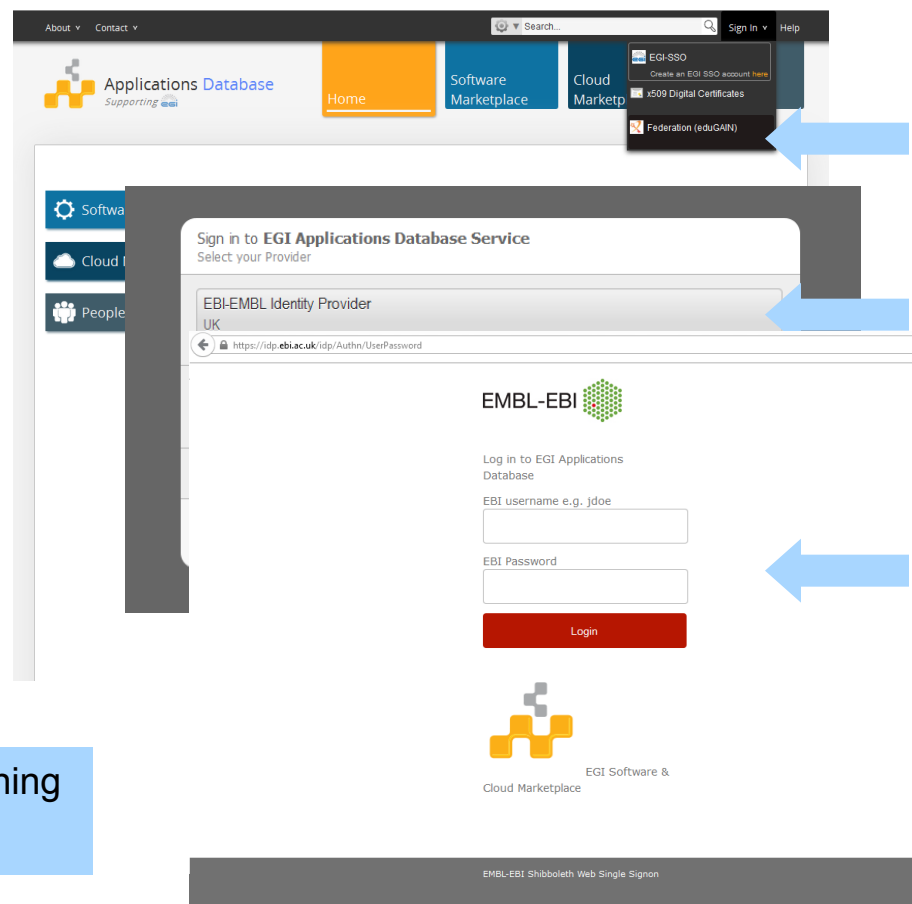
Tests with ebi-project are on-going.

Since the 1st quarter of 2014, the AppDB supports authentication using federated Identity providers through eduGAIN.

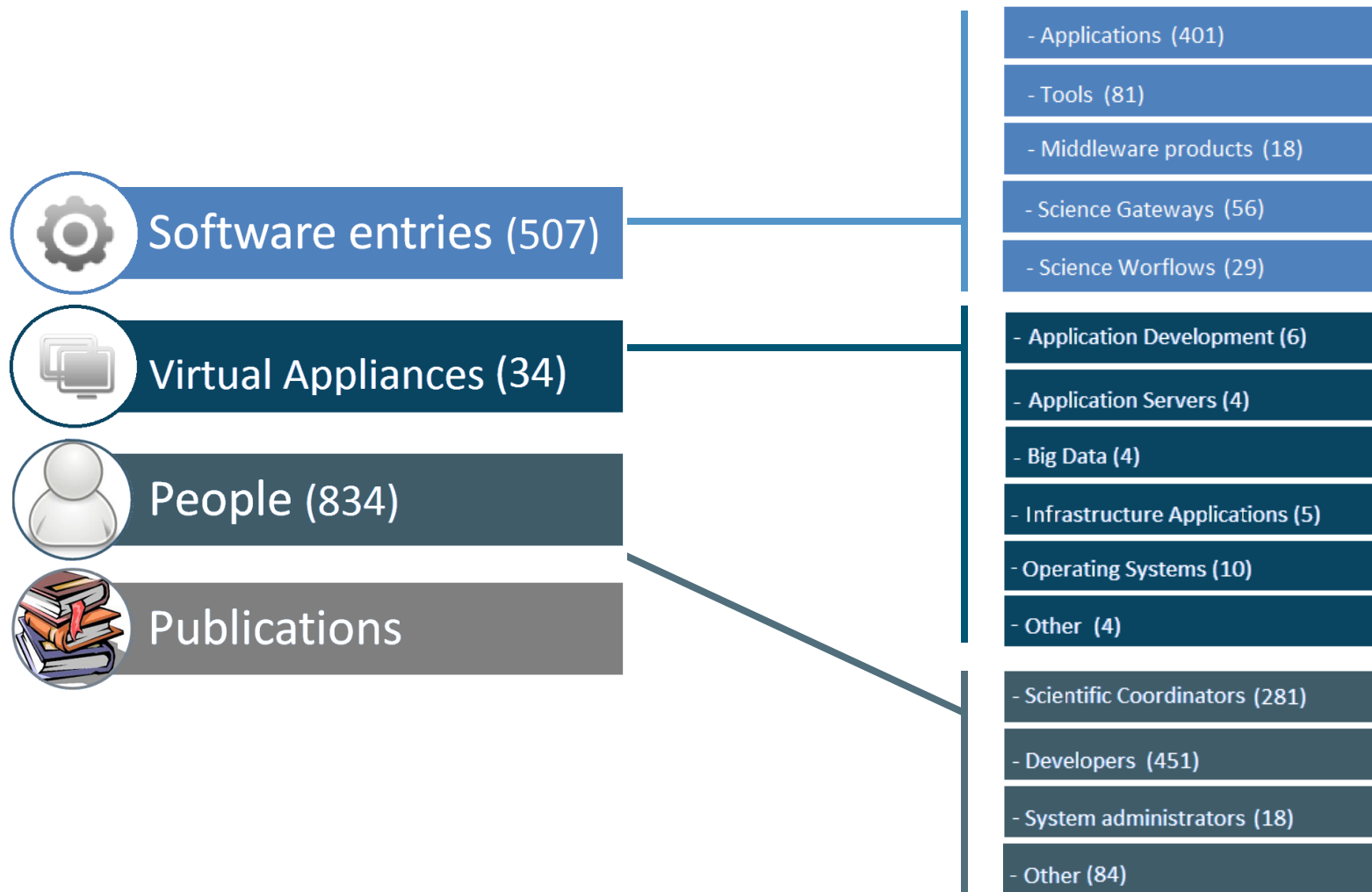
Thus, EBI-EMBL users can:

- *take advantage of the mechanism*
 - *select the EBI-EMBL IdP*
 - *and login into the AppDB*

Liberates the users from the necessity of either obtaining a new account or for using X509 certificates.



The screenshot shows the EGI Applications Database Service interface. At the top, there is a navigation bar with 'About', 'Contact', a search bar, and 'Sign In' and 'Help' links. Below this, there are several menu items: 'Applications Database', 'Home', 'Software Marketplace', and 'Cloud Marketplace'. A dropdown menu is open under 'Sign In', showing options for 'EGI-SSO', 'x509 Digital Certificates', and 'Federation (eduGAIN)'. A blue arrow points to the 'Federation (eduGAIN)' option. Below the navigation, there is a 'Sign in to EGI Applications Database Service' section with the instruction 'Select your Provider'. A dropdown menu is open, showing 'EBI-EMBL Identity Provider' and 'UK'. A blue arrow points to the 'EBI-EMBL Identity Provider' option. Below this, the browser address bar shows 'https://idp.ebi.ac.uk/idp/Authn/UserPassword'. The main content area displays the EBI-EMBL logo and the text 'Log in to EGI Applications Database'. There are two input fields: 'EBI username e.g., jdoe' and 'EBI Password'. A red 'Login' button is positioned below the password field. A blue arrow points to the 'Login' button. At the bottom of the page, there is a logo for 'EGI Software & Cloud Marketplace' and a footer that reads 'EMBL-EBI Shibboleth Web Single Signon'.



- Web: <https://appdb.egi.eu>
- Contact: appdb-support@iasa.gr
- Documentation: <https://wiki.appdb.egi.eu>