

Access to the IRF for CLEF Participants

The Information Retrieval Facility (IRF) is a new international not-for-profit foundation based in Vienna. The IRF seeks to promote and facilitate research on large scale information retrieval. A specific focus of the IRF is patent and intellectual property search, but the facility is open to researchers in all relevant areas of IR and related disciplines.

At present the IRF provides access to the following supercomputing facility:

- SGI Altix 4700
 - 80 cores → 40 CPUs (Itanium - IA-64) 1,4 GHz
 - ~ 300GB Memory (307)
 - 4FPGAs (Type: RC100)
- High performance XFS file-systems
 - ~40 TB Storage
- Software:
 - Linux
 - Lemur/Indri, Terrier, diverse JavaSDKs (Sun, BEA)
 - SGI NUMA Tools
 - Caché (object oriented database)
- Data
 - Patent Corpora (only text, no pictures and drawings) :
 - USPTO ~ 103 GB (with XML tags, ~68 GB without tags) EPO ~ 134 GB

About 2.6 million documents

To the user the facility appears to be a normal Linux systems accessed across the internet. Individual users can be given sufficient storage to work with large test collections.

At present the IRF welcomes applications all interested researchers via a simplified application process. All you need to do is send a short (approximately one paragraph) description of you proposed project, a short specification of the computing facilities required, disc, cpu, userids required, etc., and how long you need access for. Email this to the Chief Scientific Officer of the IRF.: john.tait@ir-facility.org.

Users must be committed to the Open Science philosophy of the IRF: they must be willing to share the results of their research publicly.

There are a number of other caveats and warning about the IRF.

- 1) The facility is of a finite size and supports several different kinds of usage simultaneously: users must give notification (to Oliver Falk – o.falk@matrixware.com) of critical deadlines – it cannot be assumed we know about CLEF deadlines. We will do our best to accommodate you, but this cannot be guaranteed.
- 2) The constitution of the IRF requires us to preferentially support projects with the highest scientific merit. Although we will support projects through to their agreed completion date we may at any time move to a competitive application process requiring more detailed applications and peer review at any time without notice.