The Atomic Data and Analysis Structure

—

An overview with focus on the database

Allan Whiteford, Martin O'Mullane and Hugh Summers

3rd October 2007

Contents

- Overview of ADAS
- Interactive ADAS
- Callable ADAS
- Extended-ADAS
- Offline ADAS

} For background information (11 slides)

- The ADAS database
- OPEN-ADAS

} Main focus of talk (22 slides)

Overview

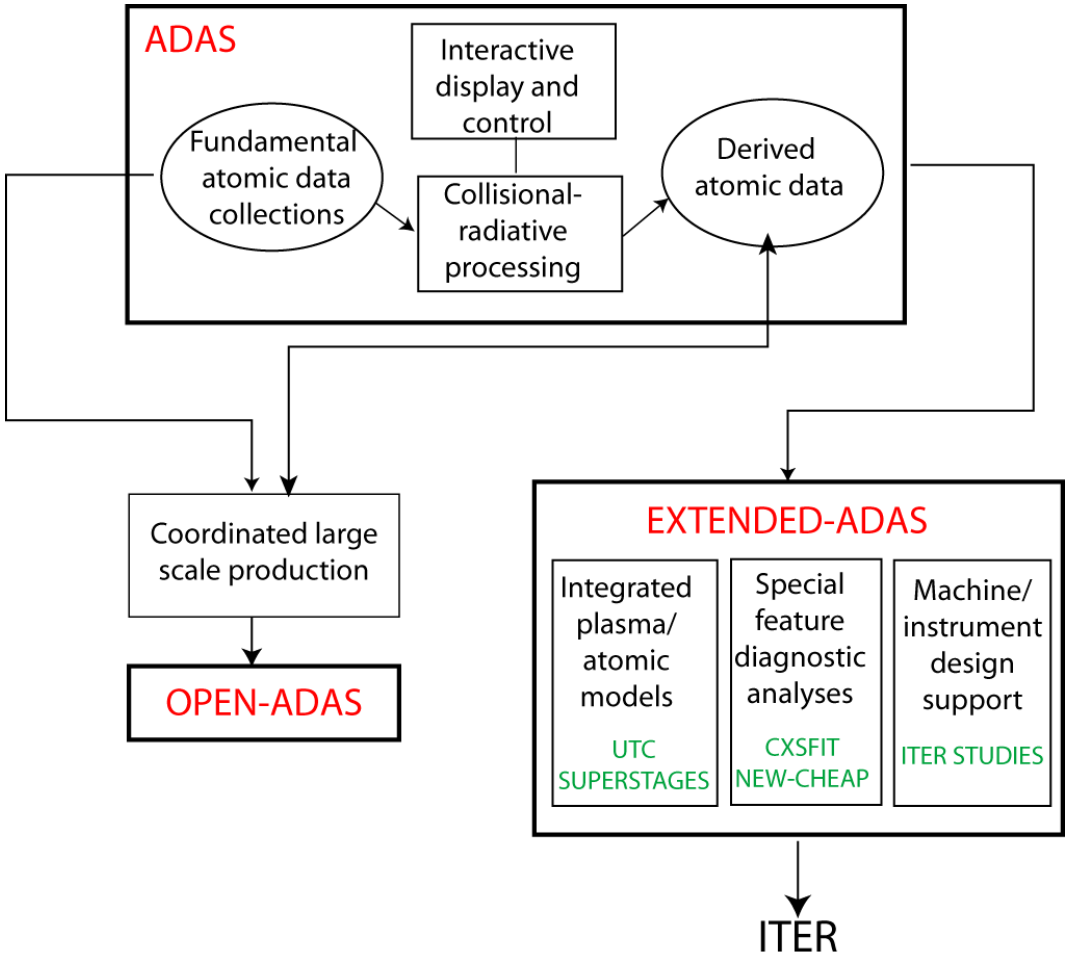
- The ADAS Project is a self-funding (i.e. funded by participants) project consisting of most major fusion laboratories along with other astrophysical and university groups. In its present incarnation it is over ten years old but the roots in JET go back twenty years.

- As an implementation, it is an interconnected set of computer codes and data collections for modelling the radiating properties of ions and atoms in plasmas.

- Historical roots are in fusion (JET) and so are the bulk of the users/members. Has also been extensively applied to astrophysics.

- Is governed by a steering committee coming from its members. Day to day running and implementation is done by the University of Strathclyde.

Schematic of ADAS Organisation



Size and Scope of use

- Rough idea of the computational 'size' of ADAS:
  - 2.8GB of data in 19,518 distinct files (discussed in detail later),
  - 422,967 lines of Fortran and 400,776 lines of IDL,
  - also contains C, Perl, csh and Matlab code,
  - tentative plans to expand to support Python and C++ integration.
- The Project currently has 22 members across North America, Europe and Asia.
- Integrated into key fusion transport codes (Strahl, JETTO, EDGE2D etc.).
- Data used for some real time control/safety systems.
- Being used extensively for ITER diagnostic design and some more general estimates of allowable leak rates etc.

Interactive ADAS

- Interactive ADAS consists of around eighty distinct, but interlinked codes.
- There are eight distinct series covering different themes (e.g. series 3 covers beam related modelling and data).
- The interactive ADAS system allows:
  - generation of all ADAS *derived* atomic data files*,
  - generation of many of the ADAS *fundamental* atomic data files*,
  - interactive inspection of the data,
  - limited plasma modelling capabilities.
- All interfaces are graphical and implemented in IDL.

* — distinction between derived and fundamental data will be discussed later.

Interactive ADAS Example: ADAS208

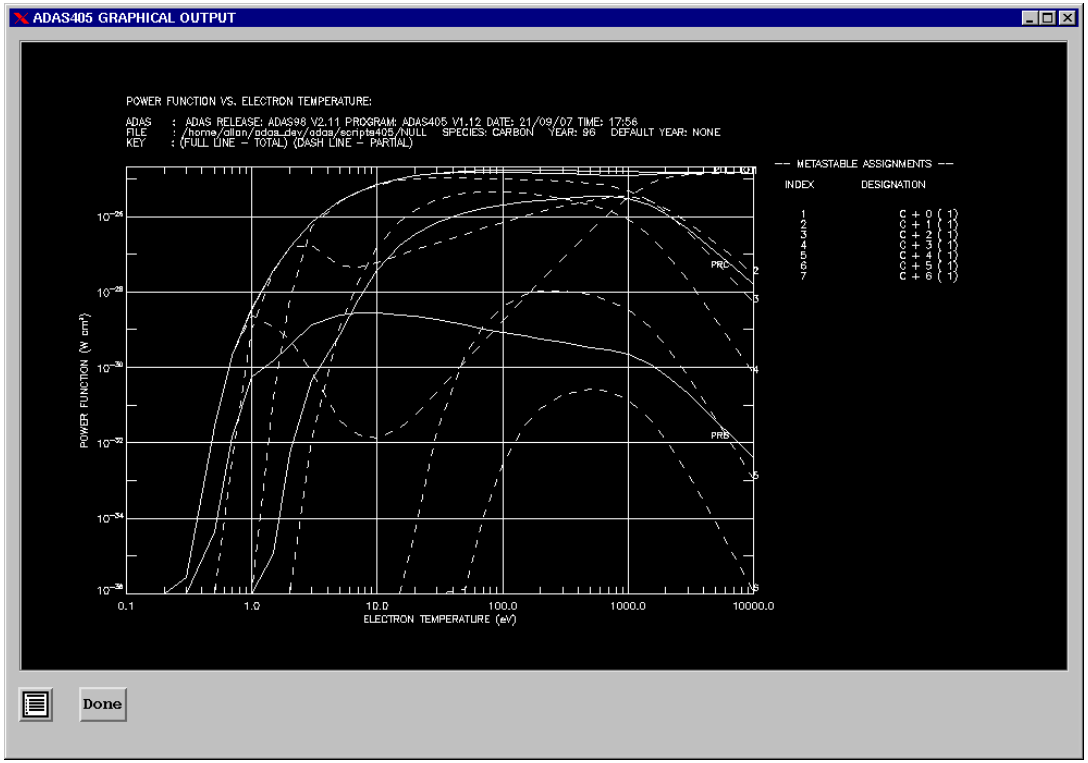
ADAS208: Generalised Collisional–Radiative (GCR) modelling of individual ions in a finite density plasma.

- Includes ionisation and recombination to adjacent ionisation stages.
- Full density dependence.
- Full metastable resolved coefficients returned.

Interactive ADAS Example: ADAS405

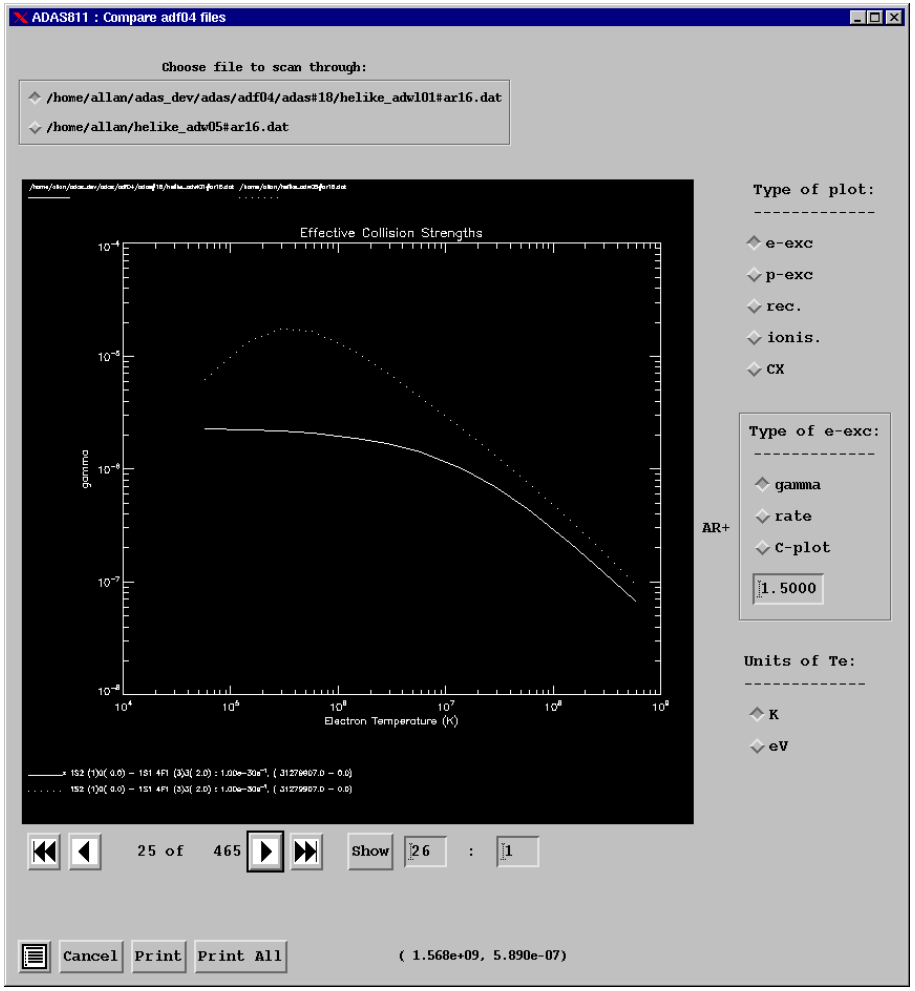
ADAS405: Equilibrium ionisation balance and cooling rate calculations.

- Assumes plasma with no transport.
- Calculates cooling curves in addition to ionisation fractions.
- Full density dependence.
- Fully metastable resolved.



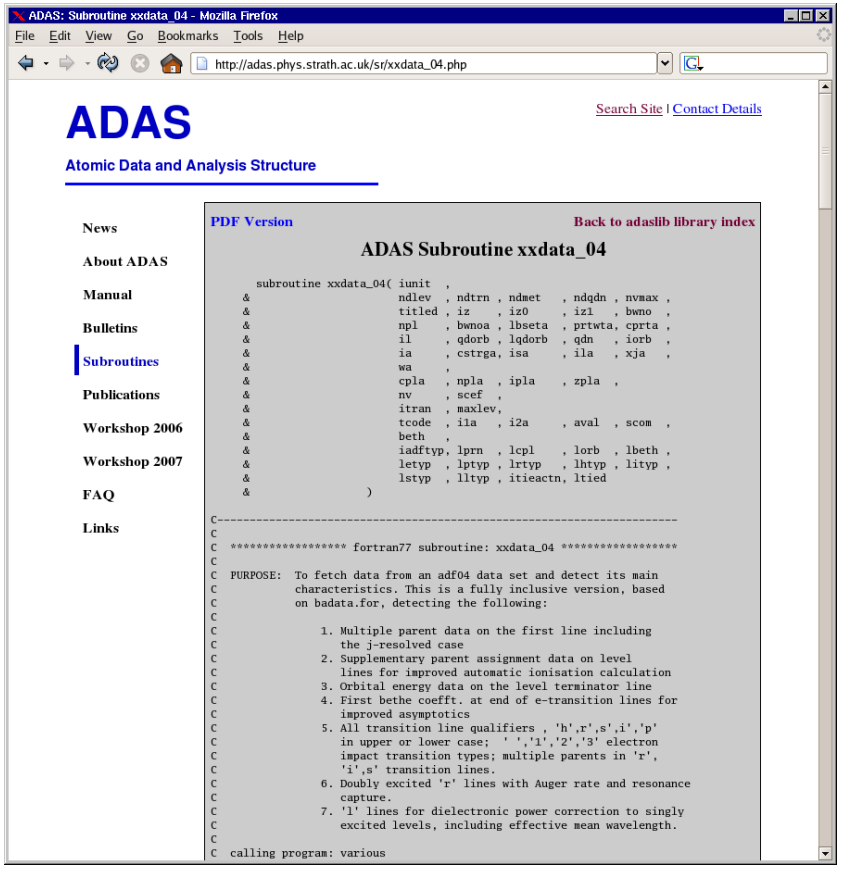
Interactive ADAS Example: ADAS811

- Comparison of (effective) collision strengths and related quantities
- One of a number of data display/comparison programs
- Can plot rates, (effective) collision strengths or cross-sections and can plot in real space or Burgess-space.



Callable ADAS

- ADAS Comes with a Fortran library of over 750 routines:
  - all documented in ~ 1500 page appendix to the user manual,
  - also documented on-line.
- Also comes with extensive IDL library:
  - far easier to use for interactive work at the command line,
  - self documenting; almost all routines accept a '`/help`' keyword.

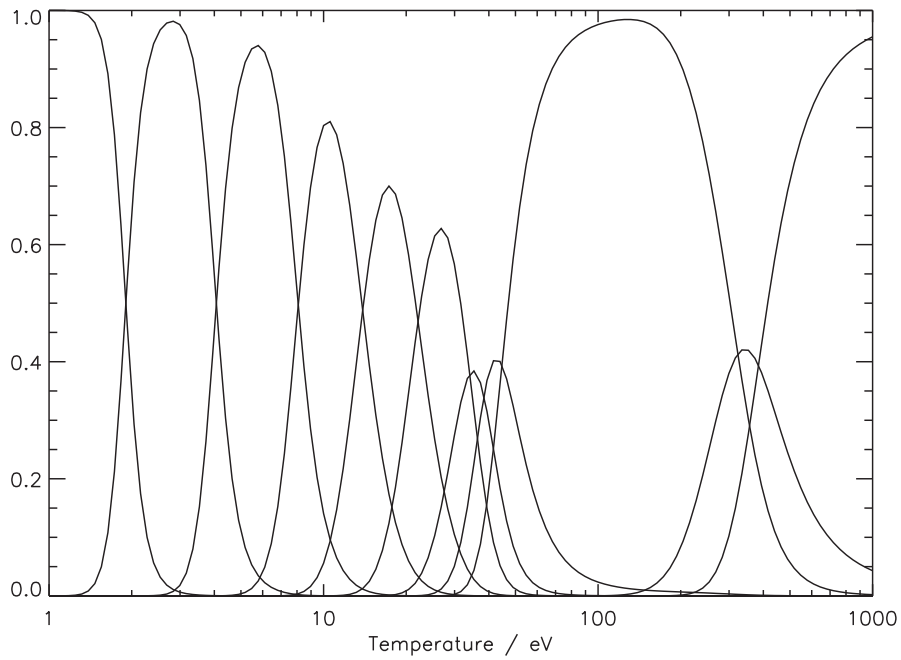


Partial/planned support for C, C++, Matlab, Perl and Python access.

Callable ADAS: A simple example

```
temp = 10d0^(findgen(127)/42)
dens = findgen(127) + 1e13

run_adas405,uid='adas', $
            year='96', $
            elem='ne', $
            te=temp, $
            dens=dens, $
            frac=frac
```




```
plot,temp,frac.ion[*,0],/xlog,xtitle='Temperature / eV'

for i=1,18 do oplot,temp,frac.ion[*,i]
```

Extended ADAS

- Covers a suite of codes which pass the *local atomic boundary*:
  - SANCO: $1\frac{1}{2}$-D impurity transport code,
  - UTC: Error propagating data analysis package for impurity transport,
  - FFS*: Generalised feature synthesis and fitting,
  - CXSFIT: Advanced spectral analysis of charge exchange measurements,
  - New-CHEAP*: Beam attenuation and modelling code for charge exchange.

- Very much tied to specific experimental analysis rather than applied atomic physics in the more general sense.

- Maintained and co-developed by the ADAS Project but not considered part of the core interactive ADAS series of programs.

* — early development/planning stage.

Offline ADAS

- Set of codes designed for running on large parallel computer systems
  - ADAS7#1: Collection of scripts for AUTOSTRUCTURE DR, RR, PI and PE calculations.
  - ADAS8#1: Automation of Cowan code for iso-nuclear calculations
  - ADAS8#2: Iso-nuclear automation of collisional–radiative model
  - ADAS8#3: Automation of R-matrix calculations.

- Usually requires expert tuning for specific machine architectures and compilers so isn't distributed as standard with ADAS.

- Do not make use of IDL (core language is Perl driving Fortran codes).

The ADAS database

- All data are strictly formatted according to ADF numbers
- Currently 50 distinct ADF numbers (some placeholders).
- Files all sit flat on a conventional UNIX filesystem.
- With release 2.12, 2.8GB of data in 19,518 distinct files.
- All data is stored as ASCII and in quasi-human readable file formats:
  - allows manual creation of datasets by users,
  - database has survived for $>$ 20 years; ASCII was a good choice.
- We make a distinction between derived, fundamental and driver data.

Derived, fundamental and driver data

- **Fundamental data** are core atomic data necessary for modelling: A-values, cross sections, effective collision strengths etc.,
  - some generated in-house but many come from literature, data centres etc.
- **Derived data** are data tailored for modelling: effective emission coefficients, effective ionisation/recombination rates etc.,
  - most of these data are unique to ADAS and is one of the main differences between ADAS and other data centres.
- **Driver data** allow complete regeneration of all ADAS derived data (and some fundamental data) in conjunction with the various ADAS codes,
  - completely unique to ADAS and of no use/interest to non-ADAS users.

Fundamental data highlights

- ADF01 — Charge exchange cross sections:
  - fully n, nl or nlj resolved,
  - underpins all of the ADAS CX modelling.
- ADF04 — Resolved specific ion data collections:
  - key data class for thermal plasmas,
  - effective collision strengths, A-values and energy levels.
- ADF09 — Dielectronic recombination coefficients:
  - fully state resolved by initial and final (potentially excited) state,
  - allows full density dependent treatment of DR.

File formats like ADF04 and ADF09 are widely used outside of ADAS (e.g. ORNL CFADC).

Derived data highlights

- ADF11 — Iso-nuclear master files:
  - gives metastable resolved effective ionisation and recombination coefficients,
  - core data class for all dynamic impurity transport modelling.
- ADF12 — Charge exchange emission coefficients:
  - gives charge exchange emission as function of plasma and beam parameters,
  - fully density dependent modelling subsequent transitions after capture.
- ADF15 — Photon emissivity coefficients:
  - gives emission of an arbitrary ion in an arbitrary plasma,
  - fully density dependent and metastable resolved.

File formats like ADF11, ADF12 and ADF15 are embedded in many key fusion codes.

Error database

- We use Monte-Carlo sampling to propagate errors in effective collision strengths through to a finite-density population structure, leading to the error in plasma models when compared with experiment.

- A complementary method shows sensitivity of specific populations to specific transitions in T_e/n_e space. This allows transitions which need refinement to be easily identified.

- We aspire* to setting up a parallel error database, for every `.dat` file we will have a corresponding `.err` file.

- This technique critically relies on having estimates of uncertainty on fundamental atomic data.

* So far only a proof of concept with one `.err` file — for neutral helium.

OPEN-ADAS

- OPEN-ADAS is a joint development between the IAEA and the ADAS Project.
- Main goals are:
  - to index the data contained within the ADAS database,
  - to provide a searching system for these data,
  - to re-work the documentation and data status,
  - to provide access to the data freely via the web.
- With the exception of the last point, all of the above have benefit to ADAS Project members.
- Primary goal, from the point of view of non-ADAS members is to disseminate key data to a wider user base.

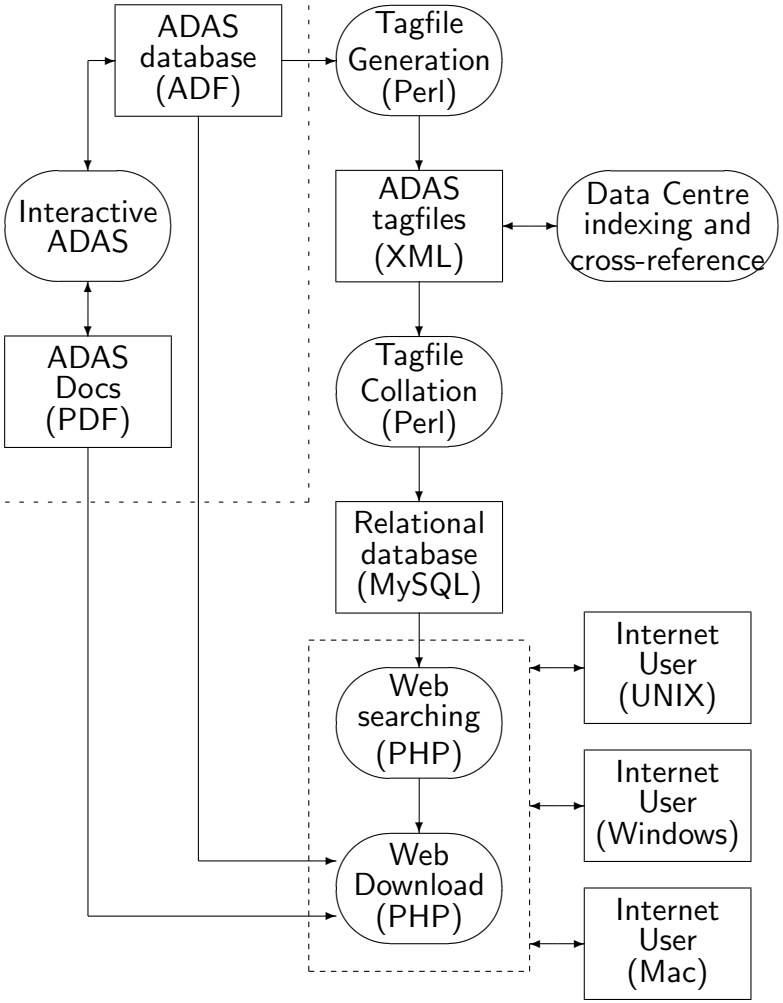
Scope of OPEN-ADAS

- OPEN-ADAS is limited to a selection of key data classes:
  - key diagnostic data classes for fusion are targeted,
  - opacity (and related data) already available,
  - no point in releasing driver files.
- OPEN-ADAS will not release any of the ADAS code, only data,
  - exception is code necessary for reading the data.
- New developments with flexible partitioning will not be included:
  - still in development,
  - need to be tuned to transport characteristics etc.,
  - best used with close support from ADAS personnel.

OPEN-ADAS data classes

Class	Description	Files	Size
ADF01	Charge exchange cross sections	118	3.0 MB
ADF04	Resolved specific ion data collections	1078	404 MB
ADF07	Electron impact ionisation coefficients	67	1.8 MB
ADF08	Radiative recombination coefficients	100	2.6 MB
ADF09	Dielectronic recombination coefficients	1531	1.0 GB
ADF11	Iso-nuclear master files	343	50 MB
ADF12	Charge exchange emission coefficients	45	2.0 MB
ADF13	Ionisation per photon coefficients	153	38 MB
ADF15	Photon emissivity coefficients	173	77 MB
ADF21	Effective beam stopping coefficients	218	4.4 MB
ADF22	Effective beam emission coefficients	402	7.6 MB
	Total	4228	1.58 GB

Implementation



Tagging system

- Perl is used to scan each ADF file:
  - each data class is done in turn,
  - custom code is written for each data class,
  - each part of the file is examined:
    - filename,
    - data contents,
    - comments section.
  - consistency is checked and warnings are flagged.
- A corresponding '`.tag`' file is generated alongside the original '`.dat`' file.
- Tag files are actually fully compliant XML files.

Example of tag file header

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<!DOCTYPE ADAS PUBLIC "-//ADAS//DTD ADAS 1.0//EN"
              "http://adas.phys.strath.ac.uk/adas.dtd">
<adf15>
        <file>
                <type>ADF15</type>
                <filename>pec96#c_pjr#c4.dat</filename>
                <directory>adf15/pec96#c</directory>
                <tagged_on>2006-07-21</tagged_on>
                <tagged_by>Allan Whiteford</tagged_by>
                <md5sum>722bda2a220dc12fe8a74c8c227bf725</md5sum>
        </file>
        <ion>
                <z0>6</z0>
                <z1>4</z1>
        </ion>
```

Transition entry in a tag file

```
<transitions>
        <transition>
                <z1>4</z1>
                <lambda units="A">40.7</lambda>
                <upper>
                        <level>3</level>
                        <cfg>1S1 2S1</cfg>
                        <m>1</m>
                        <l>0</l>
                        <j>.0</j>
                </upper>
                <lower>
                        <level>1</level>
                        <cfg>1S2</cfg>
                        <m>1</m>
```

```
                        <l>0</l>
                        <j>.0</j>
                </lower>
                <driving>
                        <level>1</level>
                        <cfg>1S2</cfg>
                        <m>1</m>
                        <l>0</l>
                        <j>.0</j>
                </driving>
                <type>Excitation</type>
        </transition>

</transitions>
```

Predecessors and contributors

- All ADAS data files contain information on:
  - Predecessors — other files in the ADAS database used in generation,
  - Contributors — people who worked on the generation of the file.
- Up until now these links had to be followed on a case by case basis by hand:
  - no automatic way of knowing where a possible error in, e.g. an ADF04 file, had propagated to elsewhere in the database,
  - in reality, these things were always tracked down by someone.
- The indexing system extracts this information from each file.
- Searchable tree is then built which forms a key part of the database generated from the '`.tag`' files.

Relational database

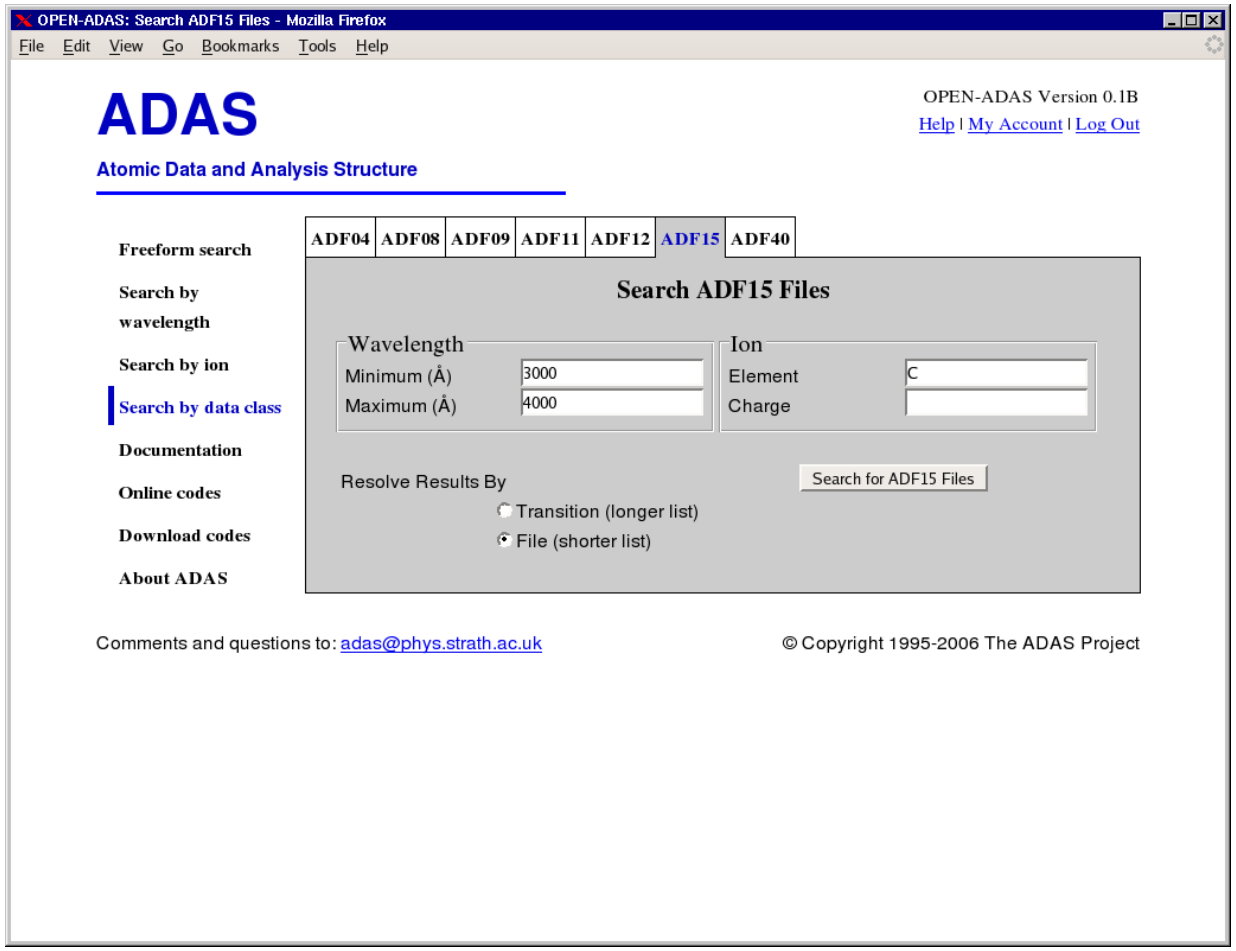
- Each '.tag' file is re-read by a different Perl script and inserted into a database.
- Using MySQL due to support within Linux and price (free!) but others (e.g. Oracle) could be used.
- Information archived according to relational database type standards:
  - data for each file is split across different tables,
  - links between files are done by unique IDs rather than filenames,
  - names of contributors are stored in separate table.
- Indexing/cross-referencing possible with data centres.
- MySQL also has ease of integration with web based software.

Web Interface

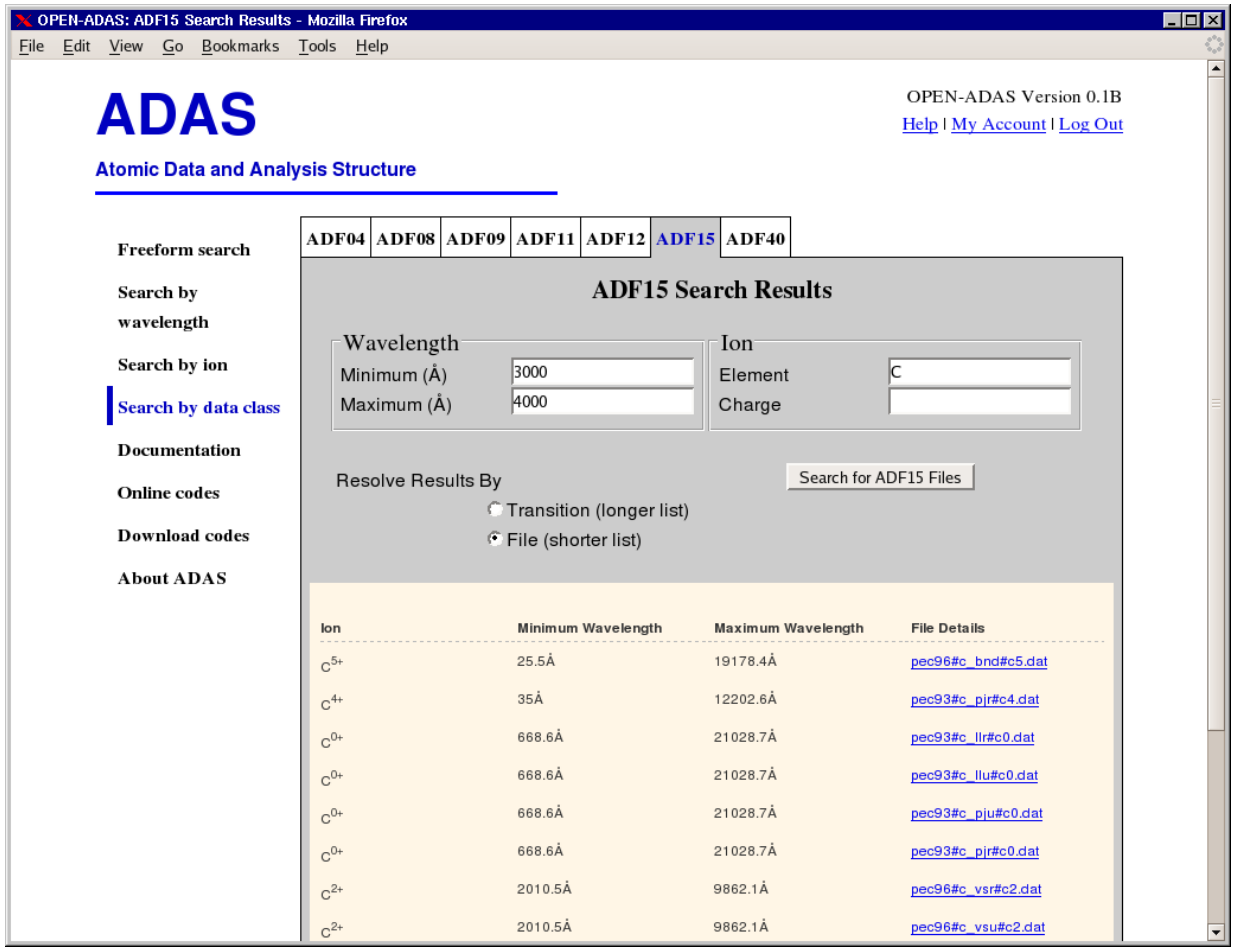
- Web interface written in PHP and outputs XHTML.
- Allows interactive searching of the database via a number of routes:
  - searching by data class,
  - cross-data class searching by ion,
  - cross-data searching by wavelength,
  - general free-form search*.
- Gives information on contents of each data file.
- Option to download the file, reading routines or documentation.

$\star$ — Awaiting implementation.

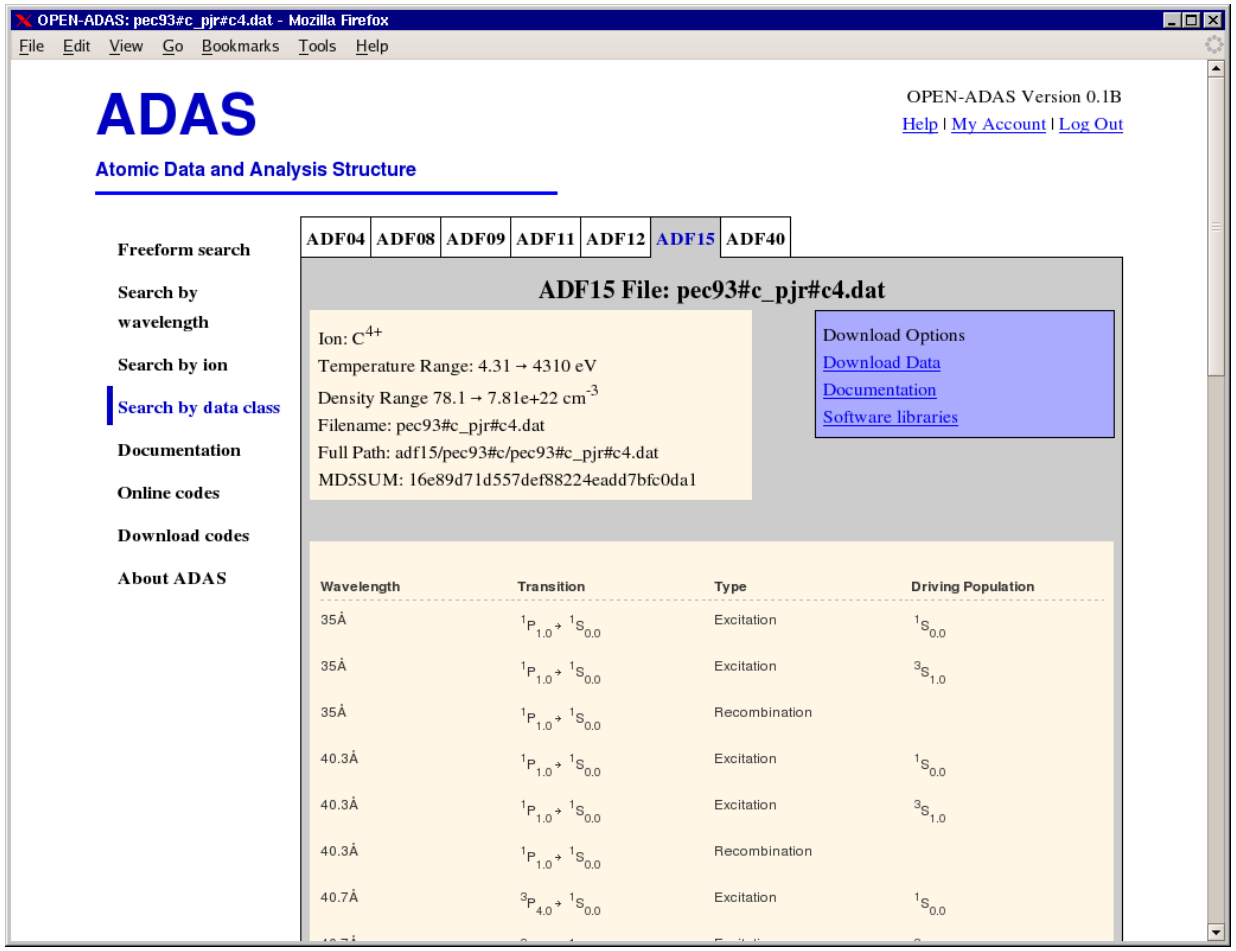
Searching ADF15 files



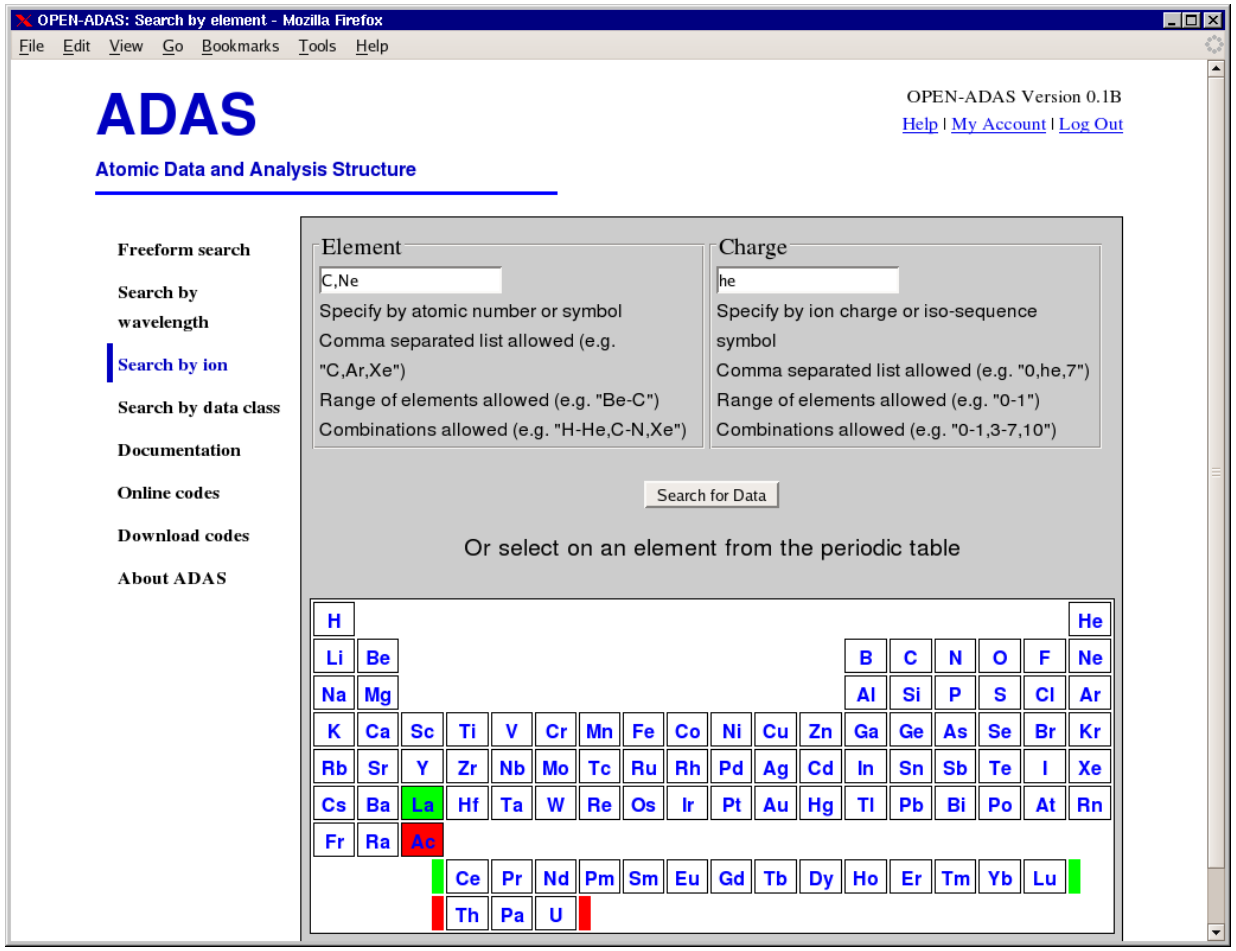
Results of searching ADF15 files



Information on an ADF15 file



Cross-class search by ion



Next steps for OPEN-ADAS

- Expansion to remaining data formats.
- Freeform search.
- Review of the design aspects of the web interface.
- Preparation of software libraries, including documentation.
- Setting up a live site and specific deployment considerations.
- Possibilities of world-wide mirrors: Japan and USA?
- We expect OPEN-ADAS to be up and running in 2008.

Some perspective on OPEN-ADAS

- ADAS is of course an atomic modelling system, linked primarily to fusion laboratories and so not a data centre as such. The fusion laboratories and other members of the ADAS Project determine its direction.

- It became clear more than fifteen years ago that active steps had to be directly by the project to improve the fundamental atomic data availability for ADAS.

- This has been very effective, has involved many workers and has provided a huge amount of high quality data. This production continues as a dispersed activity amongst many.

- OPEN-ADAS will make available these data, along with key derived data required for basic plasma modelling, to all. Most of these data are already in the public domain or freely available but not in such an organised manner.

Conclusions

- Talk covered:
  - background and non-database aspects of ADAS,
  - focused on database with some highlights,
  - OPEN-ADAS was discussed at length.

- The ADAS database is not the only part of ADAS but it is arguably the most important and certainly core to everything which ADAS does.

- We think that continued involvement with the IAEA and other DCN data centers is highly beneficial to the ADAS Project.

- We would hope that you agree the opposite is also true; we think that our database complements rather than competes with other atomic physics databases.