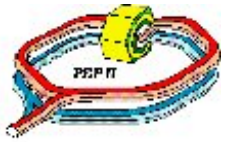


SLAC
NATIONAL ACCELERATOR LABORATORY

Outlook on outreach

3rd Workshop on Data Preservation and Long Term Analysis in HEP

M. Bellis

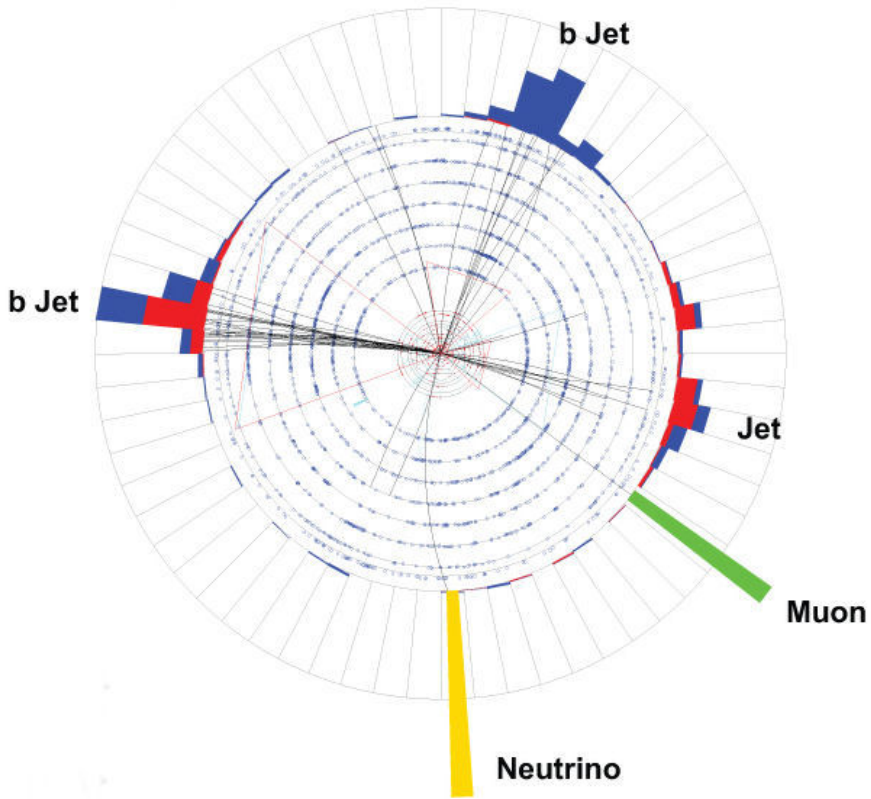
Department of Physics
Stanford University

Dec 8^{th}, 2009

Outline

Motivation

- Particle physics complexity is increasing.
- Difficult to explain to the public.
- Students have gaps in their experience.
- Can HEP data be used to attack these issues?



Audience

- Our audience.

Audience

- Our audience.
  - General public.

Audience

- Our audience.
  - General public.
  - High school students.
    - *Technically savvy public (Slashdot crowd).*

Audience

- Our audience.
  - General public.
  - High school students.
    - *Technically savvy public (Slashdot crowd).*
  - University students.

Audience

- Our audience.
  - General public.
  - High school students.
    - *Technically savvy public (Slashdot crowd).*
  - University students.
    - Particle physics course.
    - *Computing course?*
    - *Statistics course?*

Audience

- Our audience.
  - General public.
  - High school students.
    - *Technically savvy public (Slashdot crowd).*
  - University students.
    - Particle physics course.
    - *Computing course?*
    - *Statistics course?*
  - Future particle physicists.
    - *Useful for theorists?*

Audience

- Our audience.
  - General public.
  - High school students.
    - *Technically savvy public (Slashdot crowd).*
  - University students.
    - Particle physics course.
    - *Computing course?*
    - *Statistics course?*
  - Future particle physicists.
    - *Useful for theorists?*
- All present different challenges.

Audience

- Our audience.
  - General public.
  - High school students.
    - *Technically savvy public (Slashdot crowd).*
  - University students.
    - Particle physics course.
    - *Computing course?*
    - *Statistics course?*
  - Future particle physicists.
    - *Useful for theorists?*
- All present different challenges.
- First step, focus on one group.
- Effort in this will naturally filter to other groups.

Physics education

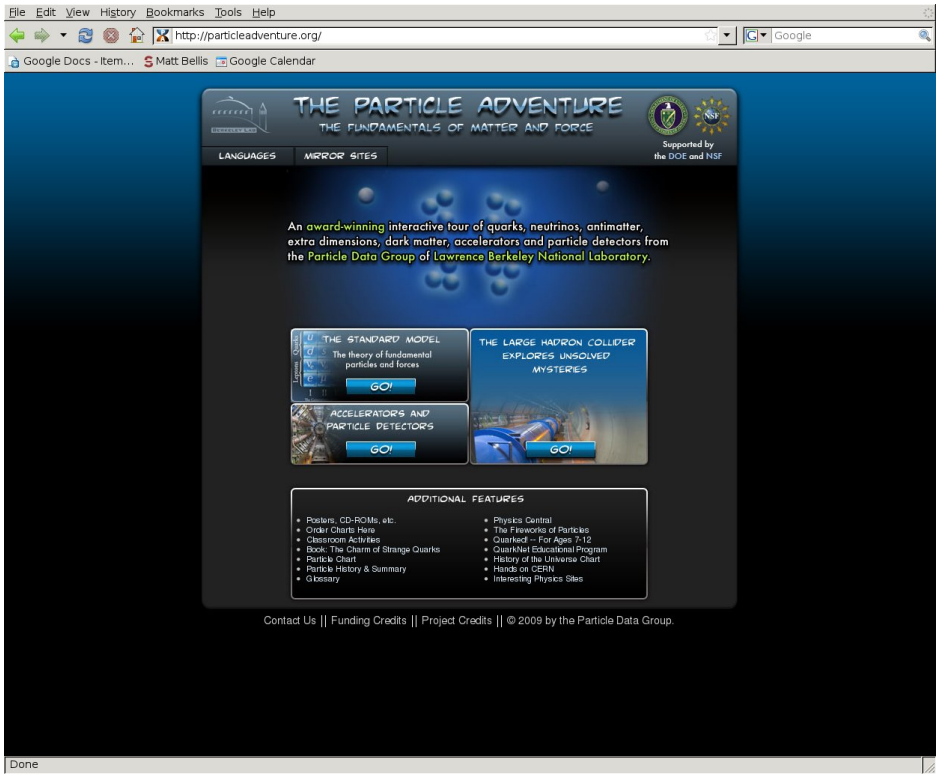
- *CPEP*
- Contemporary Physics Education Project.
- Demos and information.
- Classroom materials.

http://www.cpepweb.org/

Physics education

- *CPEP*
- Contemporary Physics Education Project.
- Demos and information.
- Classroom materials.
- Particle adventure.

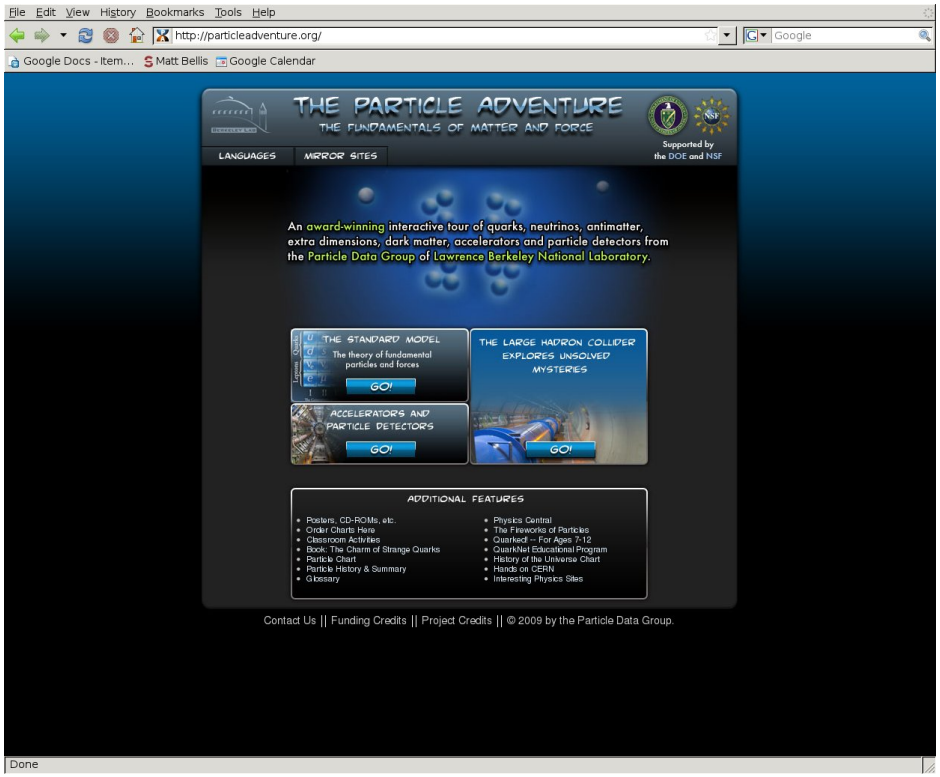
http://www.cpepweb.org/
http://particleadventure.org/a



Physics education

- *CPEP*
- Contemporary Physics Education Project.
- Demos and information.
- Classroom materials.
- Particle adventure.
- What we're proposing is *complementary* to these projects.

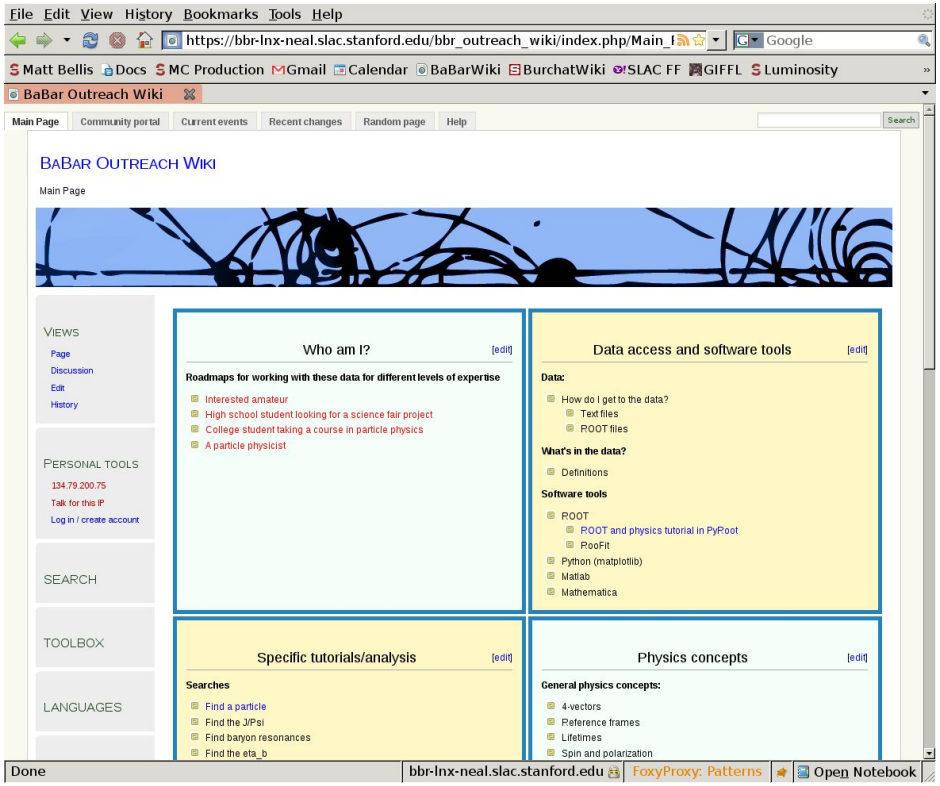
http://www.cpepweb.org/
http://particleadventure.org/a



BaBar (HEP?) outreach wiki

https://bbr-lnx-neal.slac.stanford.edu/bbr_outreach_wiki/index.php/Main_Page

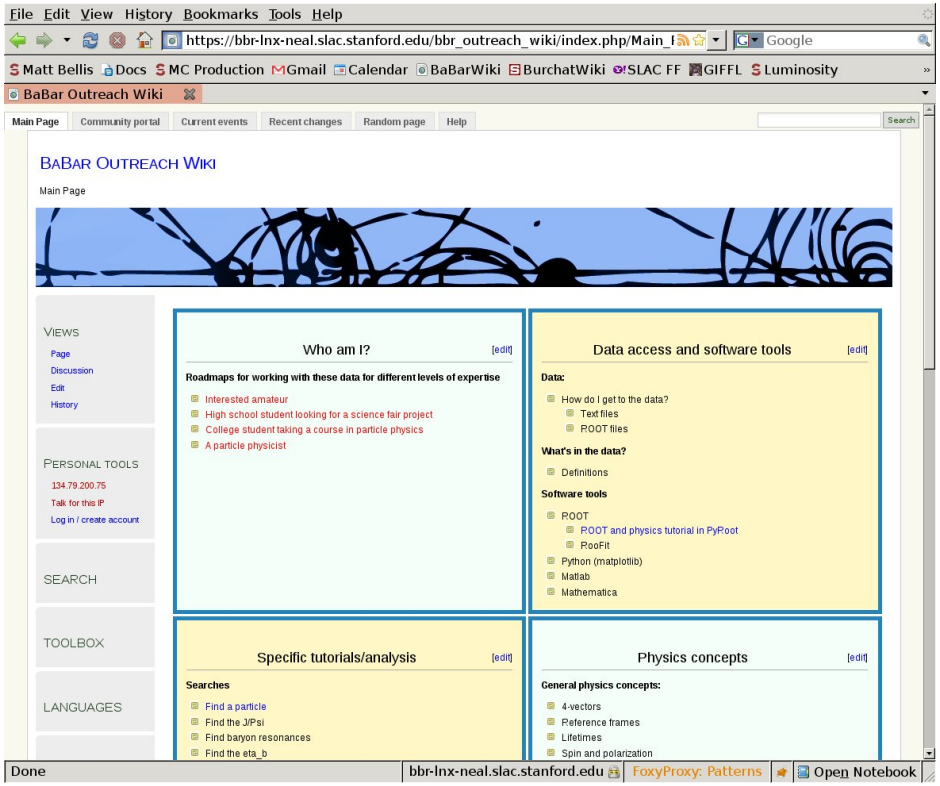
- Wiki (Mediawiki)
- Only accessible from SLAC right now...this will change.



BaBar (HEP?) outreach wiki

https://bbr-lnx-neal.slac.stanford.edu/bbr_outreach_wiki/index.php/Main_Page

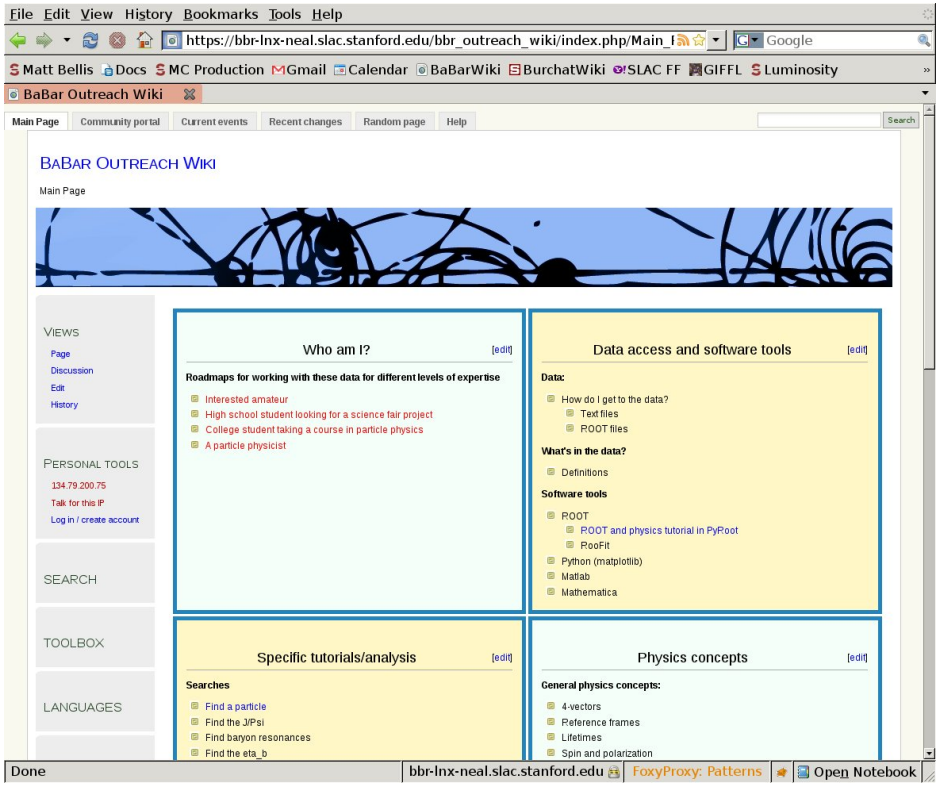
- Wiki (Mediawiki)
- Only accessible from SLAC right now...this will change.
- Data (samples), tutorials, lectures.



BaBar (HEP?) outreach wiki

https://bbr-lnx-neal.slac.stanford.edu/bbr_outreach_wiki/index.php/Main_Page

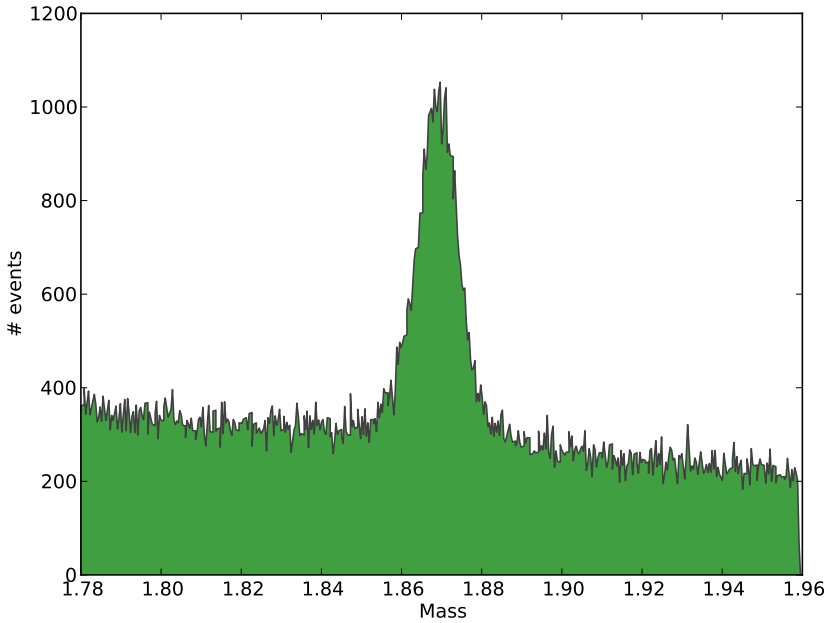
- Wiki (Mediawiki)
- Only accessible from SLAC right now...this will change.
- Data (samples), tutorials, lectures.
- Try this out in Stanford undergraduate particle physics course.
  - Pat Burchat



BaBar (HEP?) outreach wiki

https://bbr-lnx-neal.slac.stanford.edu/bbr_outreach_wiki/index.php/Main_Page

- Wiki (Mediawiki)
- Only accessible from SLAC right now...this will change.
- Data (samples), tutorials, lectures.
- Try this out in Stanford undergraduate particle physics course.
  - Pat Burchat
  - Text files
  - Python (matplotlib)
  - http://matplotlib.sourceforge.net/
  - Make mass plots, cut and count, etc.

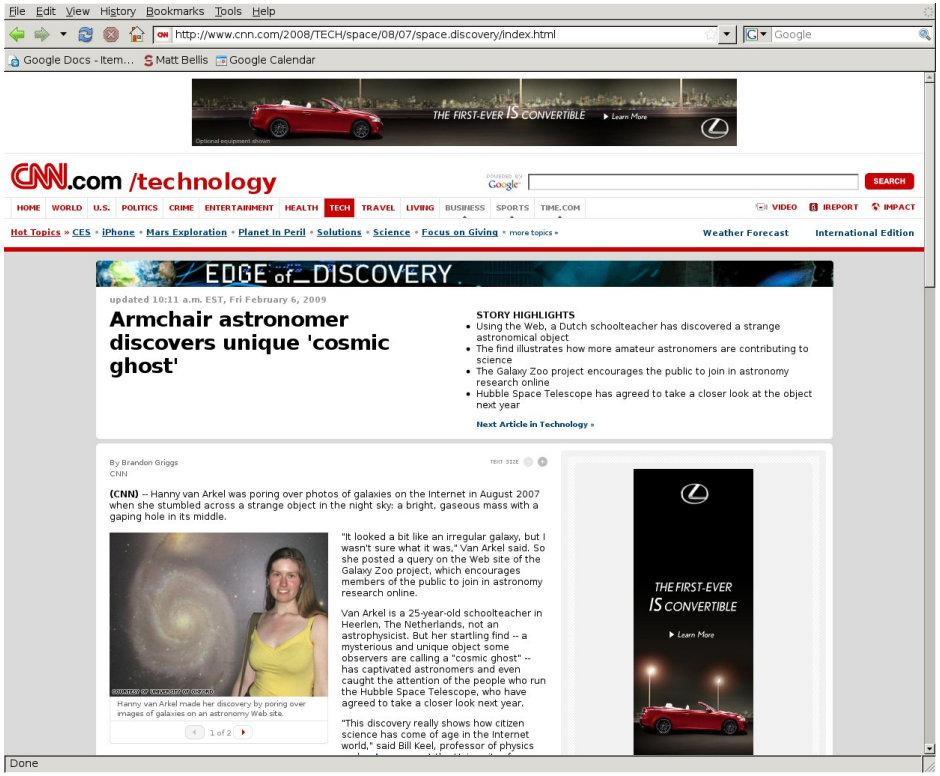


Galaxy Zoo

- Recent discussion with Galaxy Zoo folks.

Galaxy Zoo

- Recent discussion with Galaxy Zoo folks.
- Hanny van Arkel and her discovery.



Galaxy Zoo

- Recent discussion with Galaxy Zoo folks.
- Hanny van Arkel and her discovery.
- Phone conversation:
  - Chris Lintott (Oxford, Galaxy Zoo)
  - Michael Barnett (LBL, PDG/Atlas outreach)
  - Justin Albert (Victoria, BaBar)
  - Homer Neal (SLAC, BaBar)

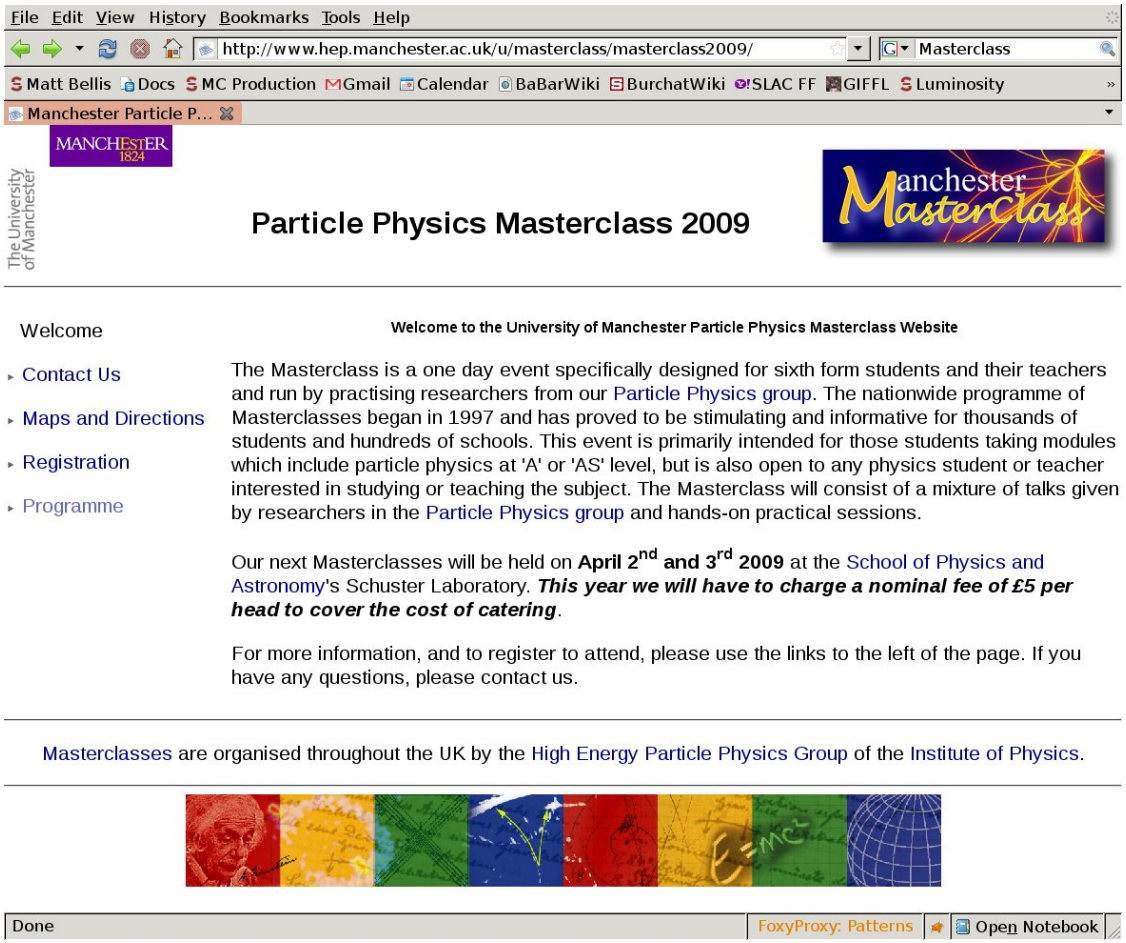
Galaxy Zoo

- Recent discussion with Galaxy Zoo folks.
- Hanny van Arkel and her discovery.
- Phone conversation:
  - Chris Lintott (Oxford, Galaxy Zoo)
  - Michael Barnett (LBL, PDG/Atlas outreach)
  - Justin Albert (Victoria, BaBar)
  - Homer Neal (SLAC, BaBar)
- Justin set up a CERN group.
  - `pub-interface-HEP-data`
- Galaxy Zoo (Event Zoo) approach may not be the right one for HEP...but that's only *my* opinion.

Other groups

- Masterclass (Manchester)

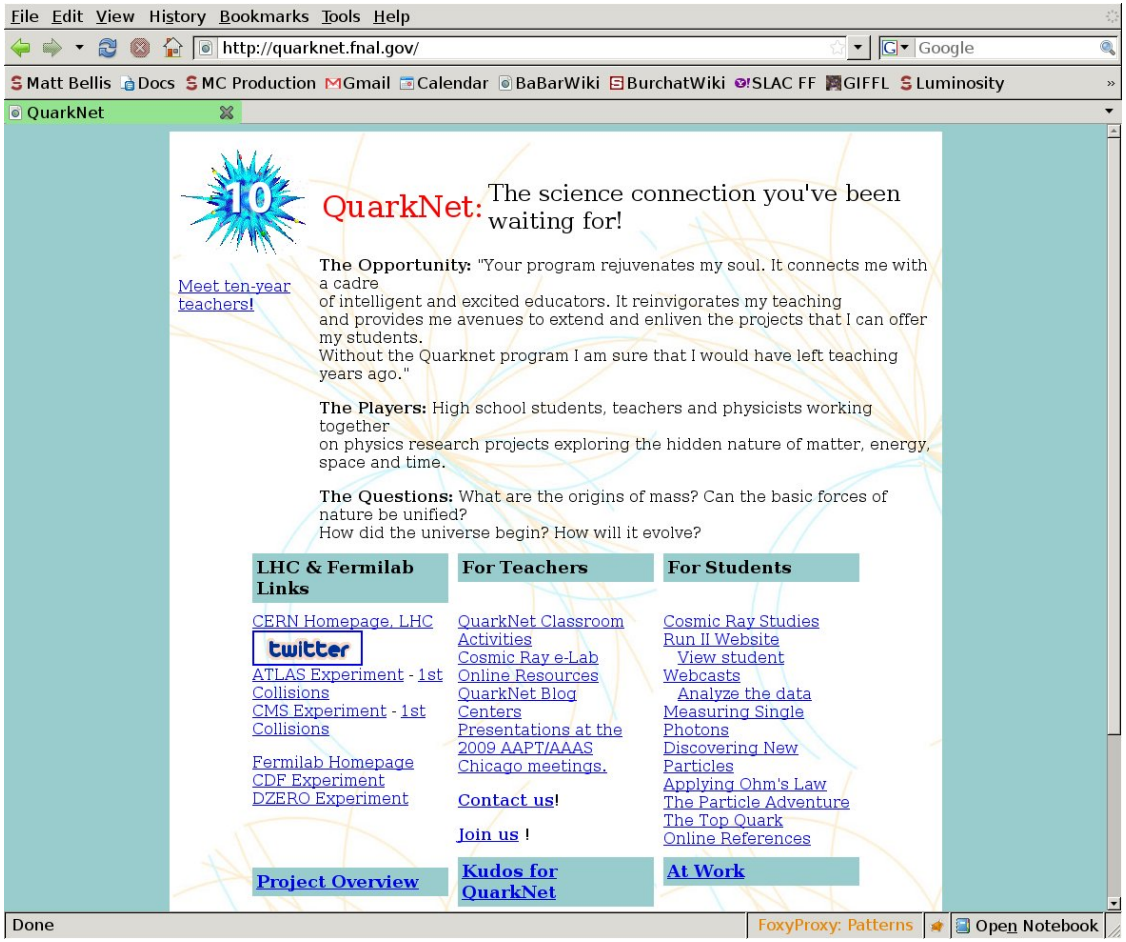
http://www.hep.manchester.ac.uk/u/masterclass/masterclass2009/



Other groups

- Masterclass (Manchester)
- Quarknet (NSF/DOE, Fermilab)

http://www.hep.manchester.ac.uk/u/masterclass/masterclass2009/
http://quarknet.fnal.gov/



Other groups

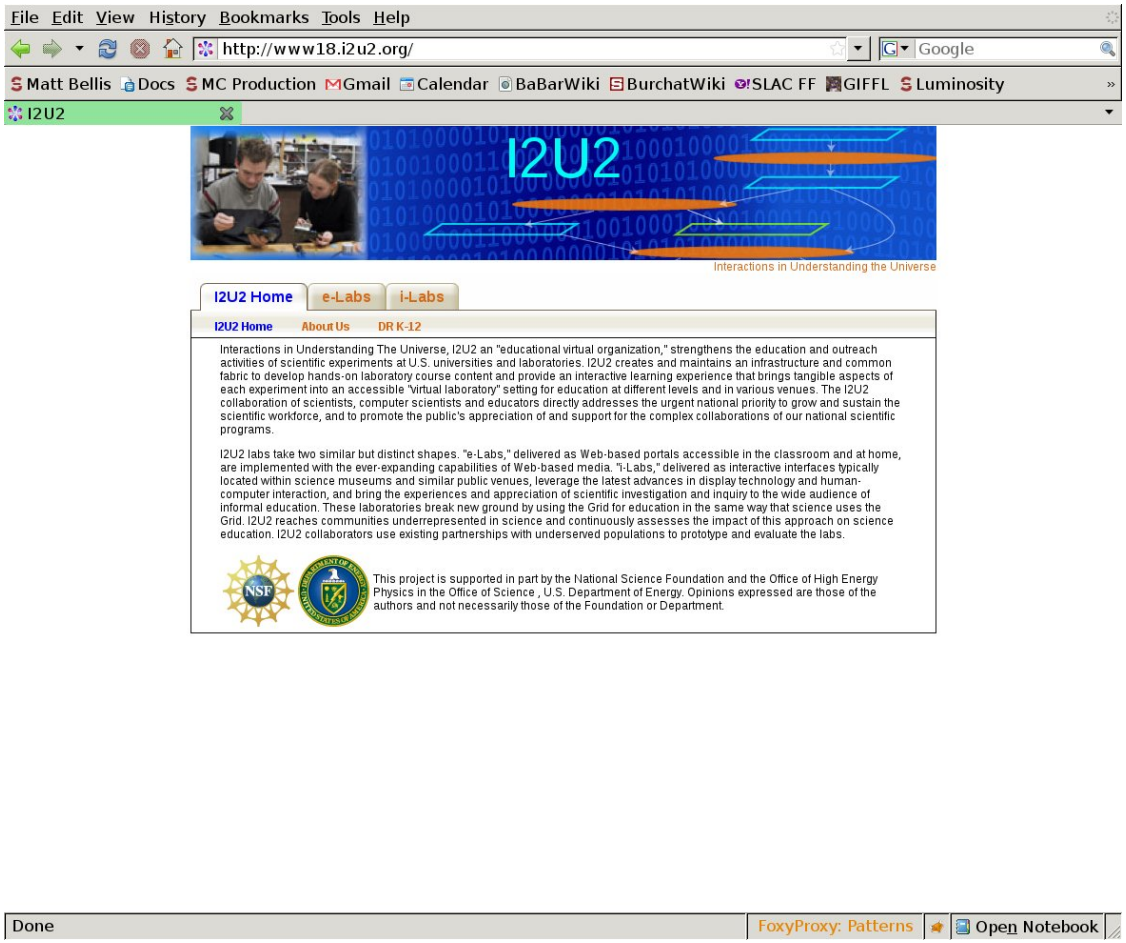
- Masterclass (Manchester)
- Quarknet (NSF/DOE, Fermilab)
- B-Lab (Belle)

http://www.hep.manchester.ac.uk/u/masterclass/masterclass2009/
http://quarknet.fnal.gov/
http://belle.kek.jp/b-lab/b-lab-english/

Other groups

- Masterclass (Manchester)
- Quarknet (NSF/DOE, Fermilab)
- B-Lab (Belle)
- I2U2 (NSF/DOE, LIGO, CMS)
- All seem to be targeting grades 6-12.

http://www.hep.manchester.ac.uk/u/masterclass/masterclass2009/
http://quarknet.fnal.gov/
http://belle.kek.jp/b-lab/b-lab-english/
http://www18.i2u2.org/



Current effort

- Outreach Wiki
- Tools and demos:

Current effort

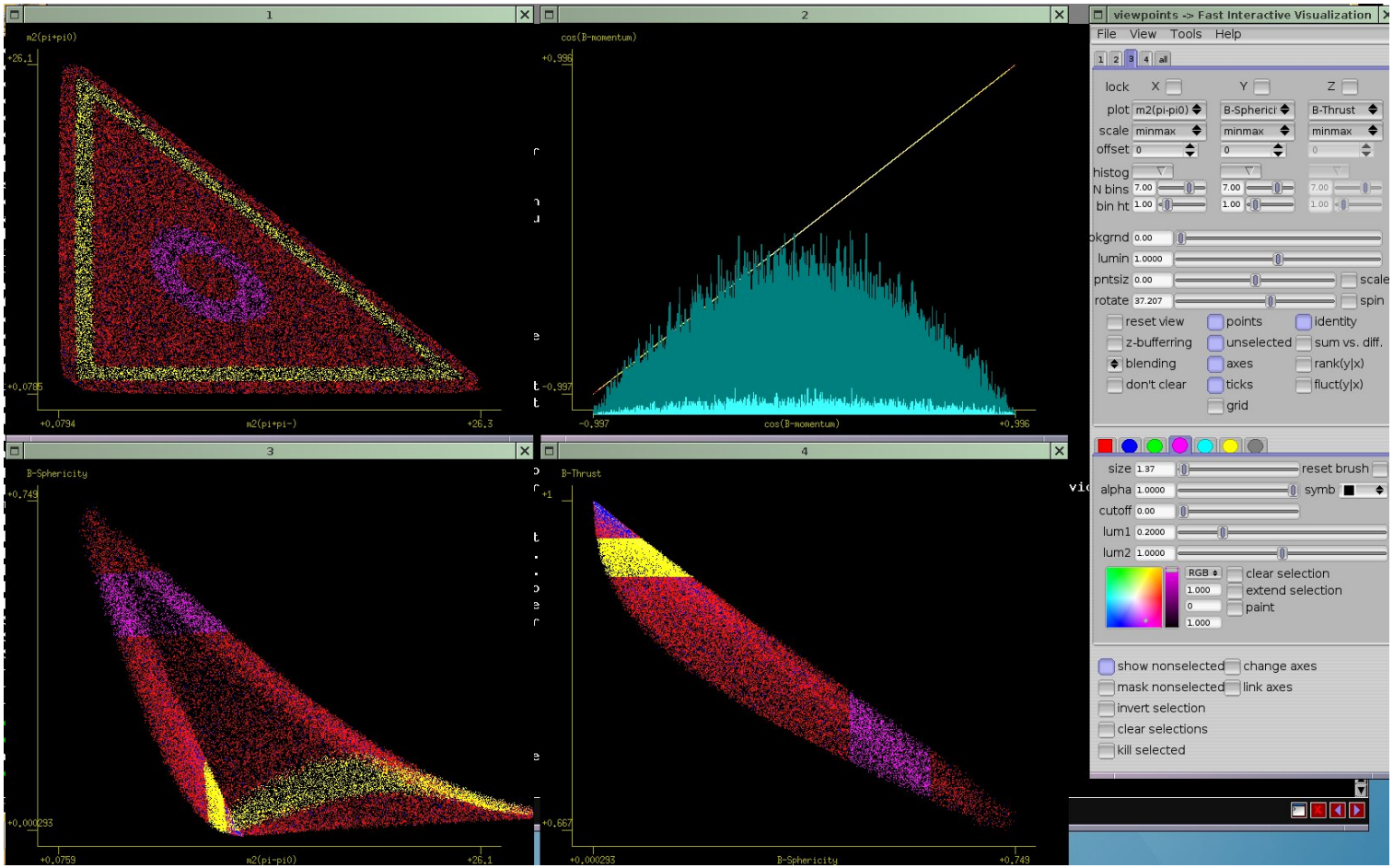
- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."

Current effort

- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."
    - What *I* heard: "ROOT would benefit from competition (outside ideas)."
  - Viewpoints. (NASA)

Current effort

http://astrophysics.arc.nasa.gov/~pgazis/viewpoints.htm
http://www.slac.stanford.edu/~bellis/viewpoints_demo.html



Current effort

- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."
    - What *I* heard: "ROOT would benefit from competition."
  - Viewpoints. (NASA)

Current effort

- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."
    - What *I* heard: "ROOT would benefit from competition."
  - Viewpoints. (NASA)
  - Python (matplotlib)

Current effort

- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."
    - What *I* heard: "ROOT would benefit from competition."
  - Viewpoints. (NASA)
  - Python (matplotlib)
  - Google visualization API with PDG data.
  - Homebrew movies for physics processes.
    - http://www.slac.stanford.edu/~bellis/HEP_data.html

Current effort

- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."
    - What *I* heard: "ROOT would benefit from competition."
  - Viewpoints. (NASA)
  - Python (matplotlib)
  - Google visualization API with PDG data.
  - Homebrew movies for physics processes.
    - http://www.slac.stanford.edu/~bellis/HEP_data.html
- File formats.
  - Data for students.

Current effort

- Outreach Wiki
- Tools and demos:
  - Other plotting visualization packages?
  - Conversation from 2^{nd} workshop at SLAC.
    - Rene Brun: "gcc benefitted when there was more competition from other compilers (icc, egcs, etc)."
    - What *I* heard: "ROOT would benefit from competition."
  - Viewpoints. (NASA)
  - Python (matplotlib)
  - Google visualization API with PDG data.
  - Homebrew movies for physics processes.
    - http://www.slac.stanford.edu/~bellis/HEP_data.html
- File formats.
  - Data for students.
  - Experimenting with LASS data (30+ years old!)
  - Revived by David Aston (SLAC/LASS)
    - http://www.slac.stanford.edu/~bellis/data_formats.html

Data formats

- File formats.
- Consider documentation.
  - Having a good `man` around can be a lifesaver.

Data formats

- File formats.
- Consider documentation.
  - Having a good `man` around can be a lifesaver.
  - Better yet, a well organzed `man`.

Data formats

- File formats.
- Consider documentation.
  - Having a good `man` around can be a lifesaver.
  - Better yet, a well organzed `man`.
    - `man` pages in use starting in 1971.
    - Huge boon to always have documentation at fingertips.
    - Imposed pseudo-requirements on programmers.

Data formats

- File formats.
- Consider documentation.
  - Having a good `man` around can be a lifesaver.
  - Better yet, a well organzed `man`.
    - `man` pages in use starting in 1971.
    - Huge boon to always have documentation at fingertips.
    - Imposed pseudo-requirements on programmers.
  - FITS (Flexible Image Transport System)
    - Used in astro community.
    - Header with information about the image.
    - Many implementations: Fortan, C/C++, Perl, Python, Java, etc.
    - Can even read with gimp or Photoshop.
  - I want a file format that carries its own documentation around with it.

Data formats

- How can I learn from this?
- Try converting LASS data.

Data formats

- How can I learn from this?
- Try converting LASS data.
  - Fixed target SLAC experiment.
  - 11 GeV K^+/K^- beams on hydrogen target.
  - 1977-1978, 1981-1982
  - $\sim$ 100M triggers.
  - David Aston (SLAC) revived the data.
    - " The original format (in fact, still is the format, since the files are just straight bit-copies of the originals) is IBM VBS – "Variable Blocks Spanned" – format, processed with VM/CMS; big-endian with IBM's mainframe floating point format (not IEEE)."

Data formats

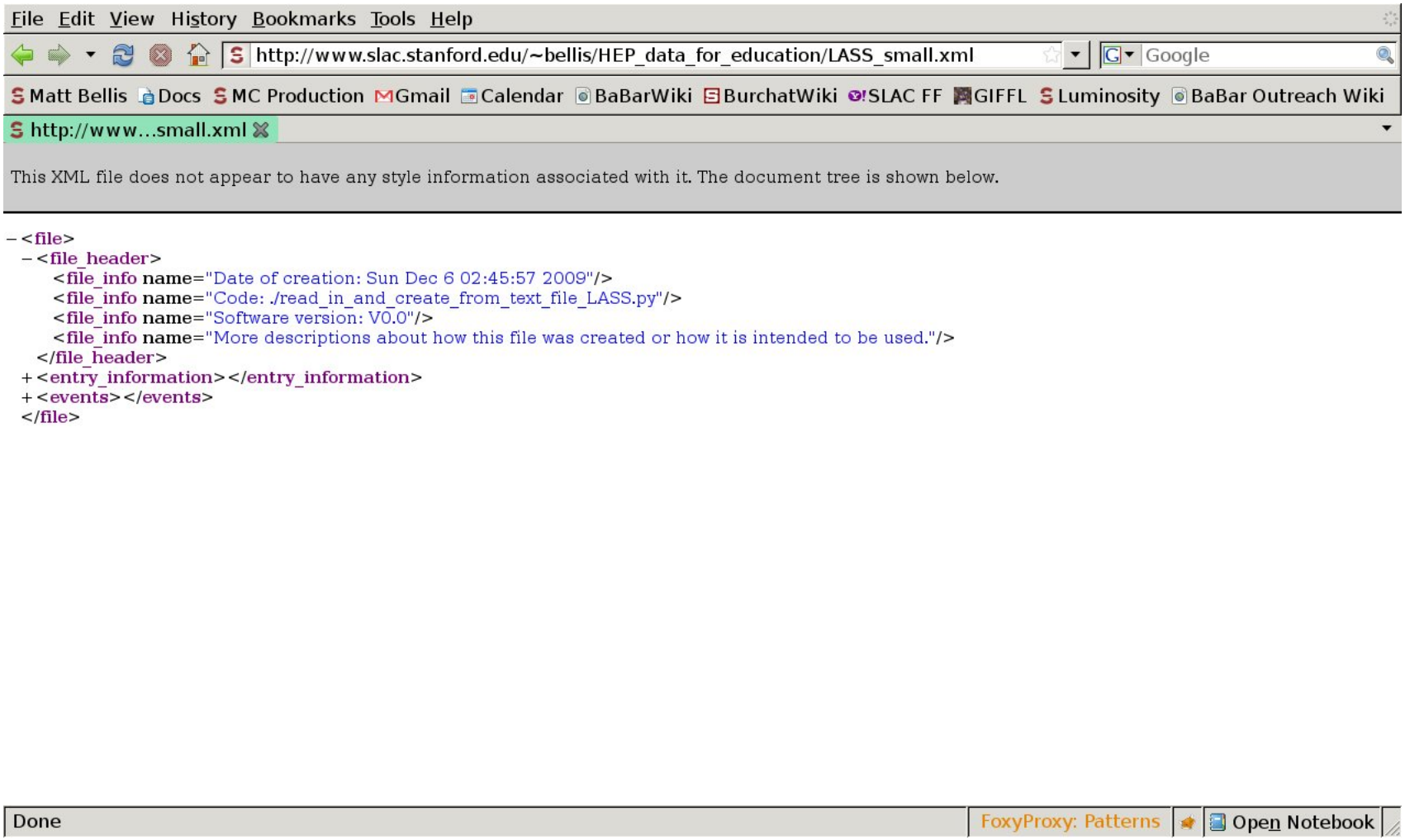
- How can I learn from this?
- Try converting LASS data.
  - Fixed target SLAC experiment.
  - 11 GeV K^+/K^- beams on hydrogen target.
  - 1977-1978, 1981-1982
  - $\sim$ 100M triggers.
  - David Aston (SLAC) revived the data.
    - " The original format (in fact, still is the format, since the files are just straight bit-copies of the originals) is IBM VBS – "Variable Blocks Spanned" – format, processed with VM/CMS; big-endian with IBM's mainframe floating point format (not IEEE)."
  - David gave me text files...

LASS text files

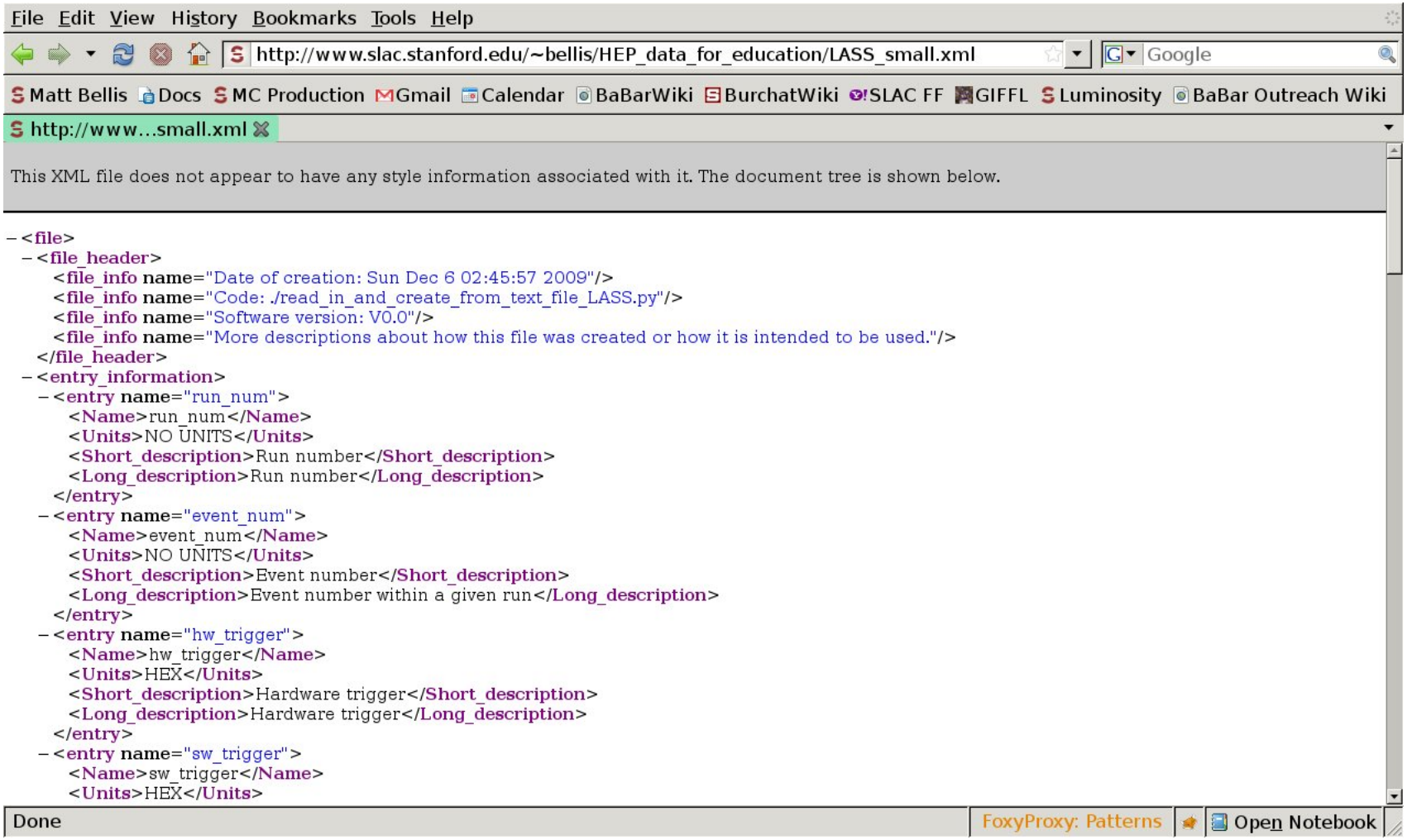
```
 new EVENT: run event hw & sw triggers    9550      3    18  8E18
 Topology :  1
 Vertices  tracks:  1 5
  Primary vtx: x y z d^2   -0.763999999 -0.40200001  81.5400009  0.0351999998
 Beam charge px py pz :  1 -0.00964925718  0.00366006303  10.9029951
 tracks/charge px py pz :
3  0.00761500327  0.309128046  8.70612812
5  0.133273423 -0.16253081  0.656466305
2  0.0946368724 -0.149057582  0.53865546
-1 -0.0939974785  0.214401409  0.689579844
 Topology :  2
 Vertices  tracks:  2 4
  Primary vtx: x y z d^2   -0.786000013 -0.375999987  80.5699997  0.0177999996
 Beam charge px py pz :  1 -0.00964925718  0.00366006303  10.9029951
 tracks/charge px py pz :
3  0.00738755707  0.309133559  8.70612812
2  0.0962994769 -0.147988901  0.53865546
0  0.037703827  0.0509431213  1.34604621
  Secondary vtx: x y z d^2   -0.858938336 -0.303011149  81.9899979
 0.00257943221
 daughters/charge px py pz :
5  0.13265042 -0.163039669  0.656466305
-1 -0.094946593  0.213982791  0.689579844
  new EVENT: run event hw & sw triggers    9550     91    18  CE1C
 Topology :  1
 Vertices  tracks:  1 5
```

Convert to XML...

LASS XML files



LASS XML files



LASS XML files

File Edit View History Bookmarks Tools Help

http://www.slac.stanford.edu/~bellis/HEP_data_for_education/LASS_small.xml

Google

Matt Bellis Docs MC Production Gmail Calendar BaBarWiki BurchatWiki SLAC FF GIFFL Luminosity BaBar Outreach Wiki

http://www...small.xml

```
  </file_header>
 + <entry_information></entry_information>
 - <events>
   - <event>
       <header run_num="9550" event_num="3" hw_trigger="24" sw_trigger="36376"/>
     - <beams>
         <beam beam_tv_map="1100" beam_charge="1" beam_X="-0.00964926" beam_Y="0.00366006" beam_Z="10.90299510"/>
         <beam beam_tv_map="2100" beam_charge="1" beam_X="-0.00964926" beam_Y="0.00366006" beam_Z="10.90299510"/>
       </beams>
     - <vertices>
         <vertex vertex_tv_map="1000" vertex_chi2="0.03520000" vertex_X="-0.76400000" vertex_Y="-0.40200001" vertex_Z="81.54000090"/>
         <vertex vertex_tv_map="2000" vertex_chi2="0.01780000" vertex_X="-0.78600001" vertex_Y="-0.37599999" vertex_Z="80.56999970"/>
         <vertex vertex_tv_map="2100" vertex_chi2="0.00257943" vertex_X="-0.85893834" vertex_Y="-0.30301115" vertex_Z="81.98999790"/>
       </vertices>
     - <vertex_map_info>
         <v_map topology="1" vertex_type="0" num_tracks_with_vertex="4"/>
         <v_map topology="2" vertex_type="0" num_tracks_with_vertex="3"/>
         <v_map topology="2" vertex_type="1" num_tracks_with_vertex="2"/>
       </vertex_map_info>
     - <tracks>
         <track track_info="1000" track_id="3" track_X="0.00761500" track_Y="0.30912805" track_Z="8.70612812"/>
         <track track_info="1000" track_id="5" track_X="0.13327342" track_Y="-0.16253081" track_Z="0.65646631"/>
         <track track_info="1000" track_id="2" track_X="0.09463687" track_Y="-0.14905758" track_Z="0.53865546"/>
         <track track_info="1000" track_id="-1" track_X="-0.09399748" track_Y="0.21440141" track_Z="0.68957984"/>
         <track track_info="2000" track_id="3" track_X="0.00738756" track_Y="0.30913356" track_Z="8.70612812"/>
         <track track_info="2000" track_id="2" track_X="0.09629948" track_Y="-0.14798890" track_Z="0.53865546"/>
         <track track_info="2000" track_id="0" track_X="0.03770383" track_Y="0.05094312" track_Z="1.34604621"/>
         <track track_info="2100" track_id="5" track_X="0.13265042" track_Y="-0.16303967" track_Z="0.65646631"/>
         <track track_info="2100" track_id="-1" track_X="-0.09494659" track_Y="0.21398279" track_Z="0.68957984"/>
       </tracks>
     </event>
   - <event>
       <header run_num="9550" event_num="91" hw_trigger="24" sw_trigger="52764"/>
```

Done FoxyProxy: Patterns Open Notebook

Data formats

- Can we replicate this in ROOT?
- Suggestion from Rene Brun: `GetUserInfo` in TTree.

Data formats

- Can we replicate this in ROOT?
- Suggestion from Rene Brun: `GetUserInfo` in TTree.
- Uses `TList` to hold anything derived from `TObject`.
- Define my own standard: `TList`'s of TString's

Data formats

- Can we replicate this in ROOT?
- Suggestion from Rene Brun: `GetUserInfo` in TTree.
- Uses `TList` to hold anything derived from `TObject`.
- Define my own standard: `TList`'s of TString's
  - 0^{th} entry: Information about the file.
  - 1^{st} entry: Defines what information is held for each entries.
  - $2^{nd} - n^{th}$ entry: Information about the `TTree` entries.
- Only store `int`'s, `float`'s and arrays of these in the `TTree`.
- Provide PyRoot script to dump information.

LASS ROOT files

```
> ./dump_header_info.py -h
Usage: dump_header_info.py [options]

Options:
  -h, --help            show this help message and exit
  -a, --all             Dump all the information
  -t TREE_NAME, --tree-name=TREE_NAME
                        Name of the TTree object
  -i INFO, --info=INFO  Append these choices for specific information to dump
                        out. Note that only the # of the information is
                        necessary. e.g: -i 0 -i2
```

LASS ROOT files

```
> ./dump_header_info.py LASS_small.root
Information about how this file was generated
Date of creation: Sun Dec  6 02:45:58 2009
Code: ./read_in_and_create_from_text_file_LASS.py
Software version: V0.0
More descriptions about how this file was created or how it is intended to be used.

Description of entries in header
0: Name
1: Units
2: Short description
3: Long description
```

LASS ROOT files

```
> ./dump_header_info.py LASS_small.root -a
Information about how this file was generated
Date of creation: Sun Dec  6 02:45:58 2009
Code: ./read_in_and_create_from_text_file_LASS.py
Software version: V0.0
More descriptions about how this file was created or how it is intended to be used.

Description of entries in header
Name
Units
Short description
Long description

run_num
                Name: run_num
               Units: NO UNITS
   Short description: Run number
    Long description: Run number

event_num
                Name: event_num
               Units: NO UNITS
   Short description: Event number
    Long description: Event number within a given run

hw_trigger
                Name: hw_trigger
               Units: HEX
   Short description: Hardware trigger
    Long description: Hardware trigger
```

Data formats

- ROOT looks promising.
- Use LASS as test-bed and try to revive entire dataset ($\sim$ 1% so far).
- Still exploring and would welcome suggestions.

Data formats

- ROOT looks promising.
- Use LASS as test-bed and try to revive entire dataset ($\sim$ 1% so far).
- Still exploring and would welcome suggestions.
- I'm just throwing this out there...
- Could ROOT's file format be separated like FITS?

Data formats

- ROOT looks promising.
- Use LASS as test-bed and try to revive entire dataset ($\sim 1\%$ so far).
- Still exploring and would welcome suggestions.
- I'm just throwing this out there...
- Could ROOT's file format be separated like FITS?
- Could this become a standard for the community?
- Could require *any function in the analysis chain to return documentation for storage in the data stream.*
- Could require *unit testing models with all functions*.
- Could allow more flexibility to adapt to new languages, hardware, etc.

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.
  - Wiki (Mediawiki software just like Wikipedia)
  - Allows many people to contribute more easily.

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.
  - Wiki (Mediawiki software just like Wikipedia)
  - Allows many people to contribute more easily.
  - Maybe a wrapper web interface (CMS: Drupal?).

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.
  - Wiki (Mediawiki software just like Wikipedia)
  - Allows many people to contribute more easily.
  - Maybe a wrapper web interface (CMS: Drupal?).
  - Forums?

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.
  - Wiki (Mediawiki software just like Wikipedia)
  - Allows many people to contribute more easily.
  - Maybe a wrapper web interface (CMS: Drupal?).
  - Forums?
  - Standardize some issues to draw others in.

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.
  - Wiki (Mediawiki software just like Wikipedia)
  - Allows many people to contribute more easily.
  - Maybe a wrapper web interface (CMS: Drupal?).
  - Forums?
  - Standardize some issues to draw others in.
  - Audience focus will filter both up and down with the right framework.

Near future

- Summary and near future efforts.
  - Education and outreach Wiki.
  - *Not* accessible to outside world...yet.
  - Wiki (Mediawiki software just like Wikipedia)
  - Allows many people to contribute more easily.
  - Maybe a wrapper web interface (CMS: Drupal?).
  - Forums?
  - Standardize some issues to draw others in.
  - Audience focus will filter both up and down with the right framework.
  - Set up something that can be maintained by small, distributed core group.

Summary

Thanks for your time

BACKUP SLIDES

University courses

- History of particle physics.
- Repository?

1) Last week we learned about the prediction of a 4th flavour of quarks and the development of the CKM matrix.

Go to the website www.slac.stanford.edu/BaBar_data and download some sample of bulk events.

Look up the primary hadronic decay modes of the D meson.

Find evidence of the D meson.

University courses

- History of particle physics.
- The BaBar dataset contains much of this history,
- Repository?

2) The $\Upsilon(4S)$ is a vector meson and is produced in a polarized state in e^+e^- colliders.

Predict the angular distribution if this state decays to two spin-0 particles.

Go to the website www.slac.stanford.edu/BaBar_data and download the appropriate skimmed sample for this problem.

Reconstruct B mesons in the $J/\psi K$ final state.

Plot the angular distribution. Does this agree with your prediction?

General public/high school students

- Put the data in people's hands.

General public/high school students

- Put the data in people's hands.
- Let people look for new particles.

General public/high school students

- Put the data in people's hands.
- Let people look for new particles.
- New correlations?

General public/high school students

- Put the data in people's hands.
- Let people look for new particles.
- New correlations?
- Encourage people to send back their plots.
  - Forum?
  - Plot of the month?
  - Have physicists to comment on the most interesting ones?

Training ground

- Could HEP data be used as a training ground?
- Online school?
- Summer program?

Training ground

- Could HEP data be used as a training ground?
- Online school?
- Summer program?
- Perfect for theorists looking for experimental experience.
- **Do an analysis start to finish?**
  - Measure branching fractions.
  - Extract spin information.
  - Measure $\sin(2\beta)$

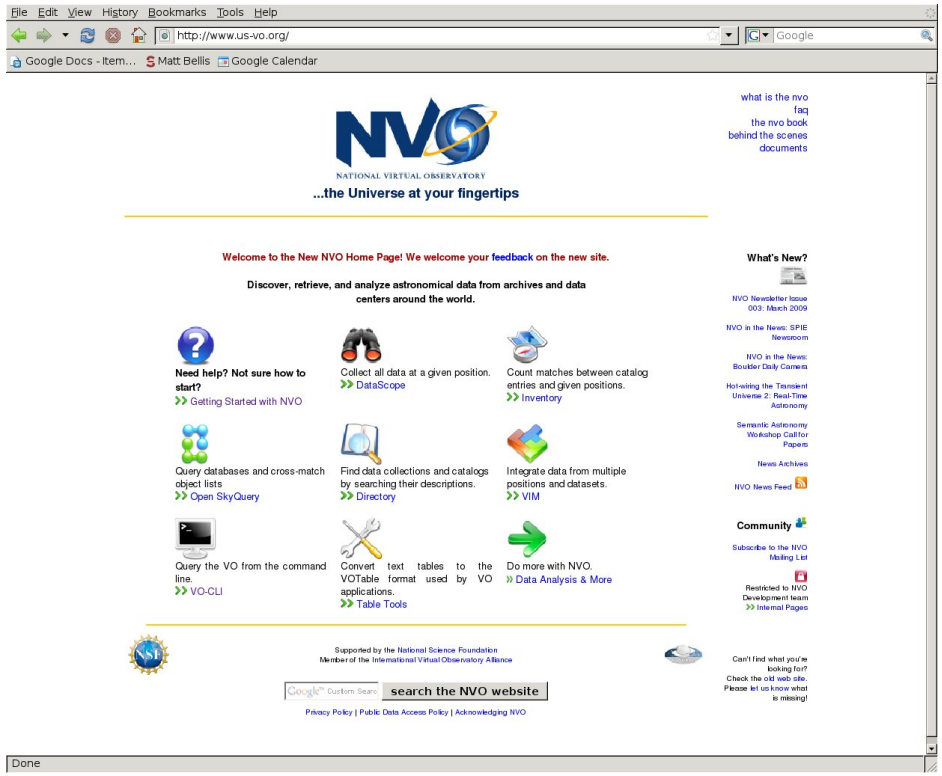
Training ground

- Could HEP data be used as a training ground?
- Online school?
- Summer program?
- Perfect for theorists looking for experimental experience.
- **Do an analysis start to finish?**
  - Measure branching fractions.
  - Extract spin information.
  - Measure $\sin(2\beta)$
- Needs a well-calibrated dataset (BaBar, others?) and an analysis flow chart,
- **Could you do a publishable analysis in a summer school?**

Physics education

- National Virtual Observatory.
- *Access to real data!*

http:www.us-vo.org

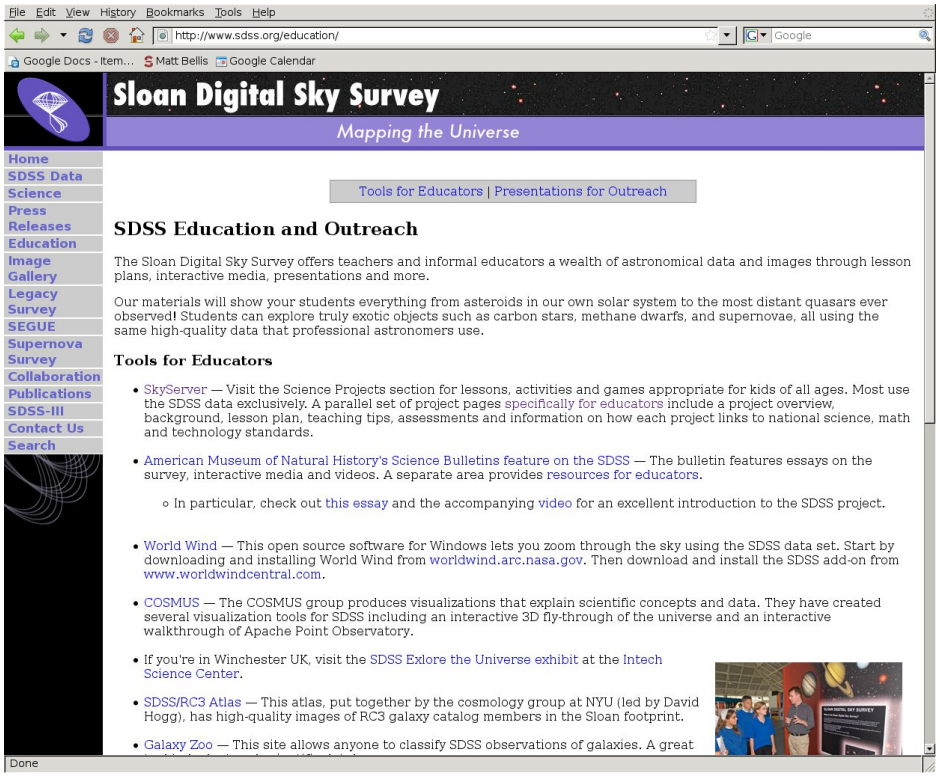


Physics education

- National Virtual Observatory.
- *Access to real data!*
- Sloan Digital Sky Survey

http:www.us-vo.org

http://www.sdss.org/education/



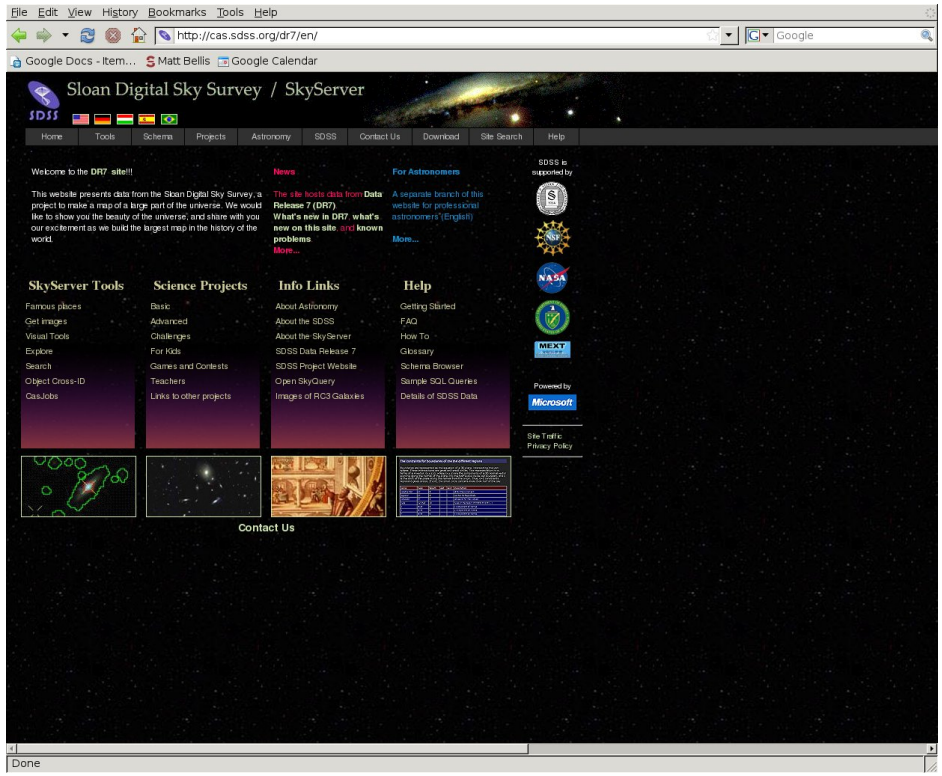
Physics education

- National Virtual Observatory.
- *Access to real data!*
- Sloan Digital Sky Survey
- SDSS SkyServer

http:www.us-vo.org

http://www.sdss.org/education/

http://cas.sdss.org/dr5/en/



Physics education

- National Virtual Observatory.
- *Access to real data!*
- Sloan Digital Sky Survey
- SDSS SkyServer
- Large Synoptic Survey Telescope

http:www.us-vo.org

http://www.sdss.org/education/

http://cas.sdss.org/dr5/en/

http://www.lsst.org/lsst/public/outreach