



# Particle Data Group

#### **Juerg Beringer**

Physics Division
Lawrence Berkeley National Laboratory



#### **Outline:**

- Overview of PDG
- Interaction with SPIRES
- PDG computing upgrade makes innovations possible
- Opportunities to collaborate



## PDG in a Nutshell



- PDG is an international collaboration charged with summarizing Particle Physics, as well as related areas of Cosmology and Astrophysics
- Review of Particle Physics (RPP) includes
  - A compilation and evaluation of measurements of the properties of the elementary particles ("Listings")
  - Review articles covering most of particle physics ("Reviews")
  - Summary tables
- Review of Particle Physics available in print and online
  - By far most cited article in HEP (total of over 32,000 citations over all editions)



## The PDG Collaboration



- 170 authors from 20 countries and 108 institutions
- Plus 700 consultants in the HEP community
- Close collaboration with working groups at LEP, Tevatron, B-factories, ...





## 2008 Edition of the Review



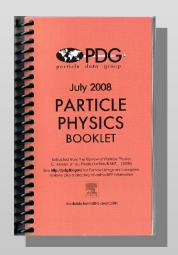
#### 2,778 new measurements from 645 new papers

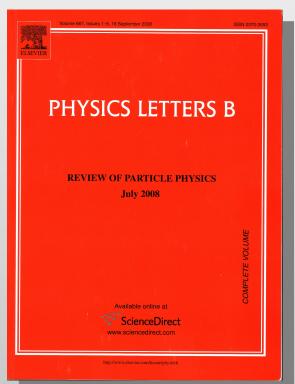
(in addition to 24,559 measurements from 7,104 papers that first appeared in previous editions)

#### **108** reviews written by experts

RPP: 1344 pages

Booklet: 294 pages



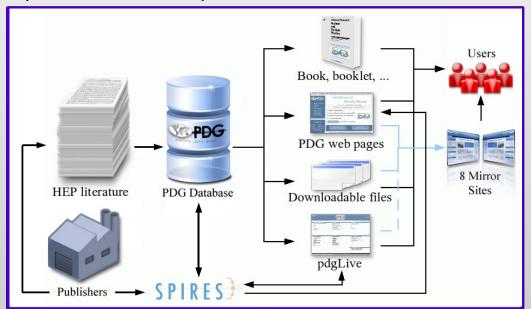




## **Data Evaluation**



- Primary goal of PDG is to evaluate and review available published data in order to give authoritative answers endorsed by the experts in the field
  - We rely on publishers and libraries (in particular through SPIRES) for access to published data



 As part of doing its work, PDG naturally catalogs published results, but this is not our primary goal



## Interaction with SPIRES



#### So far have interacted with SPIRES in two ways:

- Harvesting of information based on paper reference
  - Authors and titles
  - Used e.g. to display paper titles in popup windows in pdgLive

VALUE (GeV)	DOCUMENT ID TECN COMMENT
171.2 ± 2.1	OUR EVALUATION See comments in the header above.
$170.8 \pm 2.2 \pm 1.4$ $176.2 \pm 9.2 \pm 3.9$	1, 2 AALTONEN 071 CDF lenton + jets (h-tag)  3 ABAZOV AALTONEN 20071: Physical Review
179.5 ± 7.4 ± 5.6 164.5 ± 3.9 ± 3.9	3 ABAZOV Letters 99 (2007) 182002 4, 2 ABULENCIA Lepton+Jets Topology at CDF II
$180.7 + \frac{1}{1}3.\frac{5}{4} \pm 8.6$	5 ABULENCIA U7J CDF lepton + jets

#### Cross-referencing between SPIRES and PDG

- From pdgLive to SPIRES, and from SPIRES to pdgLive
- From internal pages for encoders to SPIRES (as a convenient way to point encoders to relevant papers)
- Based on "CODEN, Volume, Page" where possible, on SLAC IRN otherwise



## Interaction with SPIRES



#### From pdgLive to SPIRES



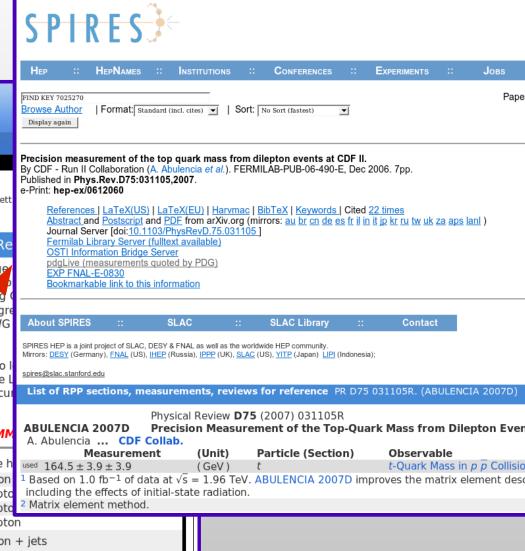
#### t-Quark Mass in $p \overline{p}$ Collisions

 $170.2 \pm 4.1 \pm 1.2$ 

OUR EVALUATION of  $171.2 \pm 1.2 \pm 1.8$  GeV (TEVEWWG 2008A) is an average from Tevatron Run-I (1992 – 1996) and Run-II (2001-present) that were this Review. This average was provided by the Tevatron Electroweak Working correlated uncertainties properly into account and has a  $\chi^2$  of 10.6 for 10 degree most recent unpublished top mass measurements from Run-II, the TEV WWG of 172.6  $\pm$  0.8  $\pm$  1.1 GeV (TEVEWWG 2008). See the note "The Top Quark".

For earlier search limits see PDG 1996, Physical Review **D54** (1996) 1. We no indirect top mass determinations from Standard Model Electroweak fits in the L can be found in the Listings of the 2007 partial update). For a discussion of cur "The Top Quark" and "Electroweak Model and Constraints on N w Physics."

VALUE (GeV)	DOCUMEN	IT I TECN	сомм
171.2 ± 2.1	OUR EVALUATION	e comments	s in the h
$170.8 \pm 2.2 \pm 1.4$	1, 2 AALTONE	07I CDF	lepton
$176.2 \pm 9.2 \pm 3.9$	<sup>3</sup> ABAZOV	07W D0	dilepto
$179.5 \pm 7.4 \pm 5.6$	The state of the s	07W_D0	dilepto
$164.5 \pm 3.9 \pm 3.9$	4 , 2 ABULENCIA	A 07D CDF	dilepton
$180.7 + \frac{1}{1}3 = \frac{5}{4} \pm 8.6$	5 ABULENCIA	A 07J CDF	lepton +

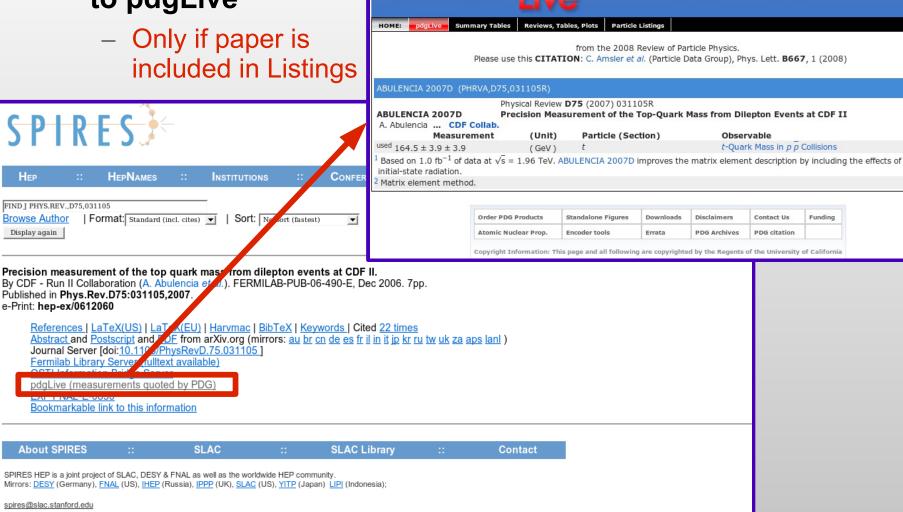




#### Interaction with SPIRES



#### From SPIRES to pdgLive



Funding



# **Urgent Computing Upgrade**



- Presently used PDG system dates back to late eighties
  - This is before the web was born!
  - Hardware upgraded from original VAX to now Linux PCs
  - Software philosophy still dates back to single-user data entry on an ASCII terminal



- Became clear that required PDG computing upgrade cannot be accomplished without additional resources
- Developed plan and asked for funding from DOE and NSF
- Comprehensive review by DOE last Fall

Written in 2006

# High-Level Requirements and Roadmap for PDG Computing

Juerg Beringer
Particle Data Group
Lawrence Berkeley National Laboratory

This document summarizes the high-level requirements for the upgraded PDG computing system and proposes a roadmap for completing the upgrade. It is intended to serve as a starting point for a cost estimate for the completion of the upgrade project.



## **Outcome**



- Vital role of PDG is reaffirmed
  - "The PDG publications are crucial to the field ..." (DOE reviewer)
  - "The work of the PDG is absolutely necessary for rapid progress of elementary particle physics." (NSF reviewer)
- Funding from DOE and NSF for computing upgrade
  - 2 FTE for 3 years
  - 0.5 FTE of a programmer for ongoing support after initial development
- Work on computing upgrade started at the end of last year with team of experienced computing professionals



# **Goals for New PDG System**



- A modern, modular, extendable, easy-to-use, maintainable and well-documented computing infrastructure for PDG
- Production quality system PDG data must be correct
  - Extensive error-checking and cross-checking built into system
- Support all areas of our work, including in particular:
  - Decentralized, web-based data entry and verification for Listings
  - Interaction with over 100 review authors
  - Monitoring of progress in RPP production
  - Programs for evaluation of data (fits, averages, plots, ...)
  - Expert tools for editor, including creation of book manuscript and static web pages (PDF files)
  - Interactive browsing of PDG database similar to pdgLive
- See backup slides for technical details on new system



## **New Possibilities**



- Immediate and primary goal of the PDG computing upgrade is to ensure PDG can continue to function well
  - This has absolute priority over any fancy extensions
- New computing system will also provide platform where innovative new features can be implemented easily
- Planned new features will include for example
  - Search functionality for pdgLive
  - Downloading of PDG data in machine-readable format
  - "pdgLive" on Blackberry / iPhone / ... mobile devices
- Other ideas are under discussion
  - RSS feed, Twitter, or e-mail lists for announcements?
  - Web-based user participation in PDG ("Web 2.0")?
  - ...



## **PDG** and "Web 2.0"?



Are there areas, where it would be useful if users could give input to PDG via the web?

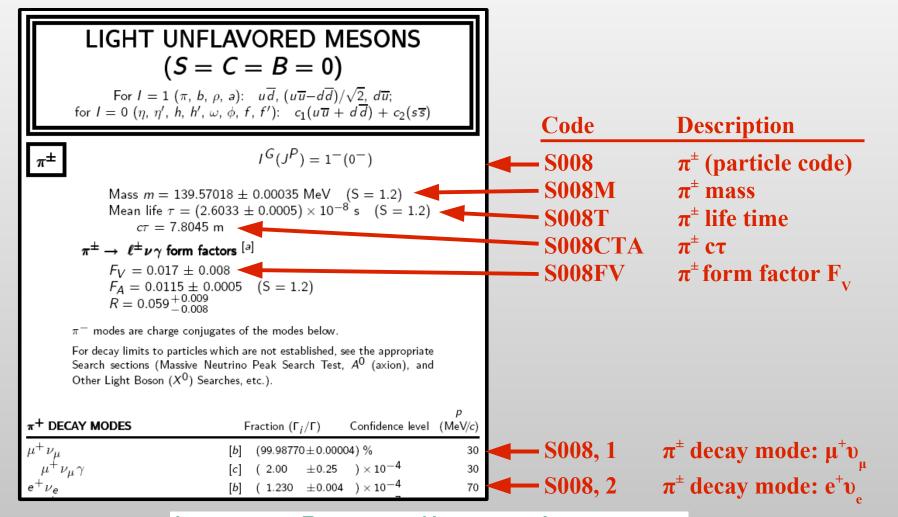
- Possibility to comment on reviews or sections in the Listings?
  - In public, not just via e-mail or as feedback on pdgLive
  - On pdgLive pages? Via public discussion lists?
- Could current manual scanning of HEP literature by PDG be augmented by community input via web?
- Tagging of HEP articles based on PDG classification?



## **PDG Classification**



#### Behind the scene, the PDG Listings define a wellestablished classification scheme for HEP data:



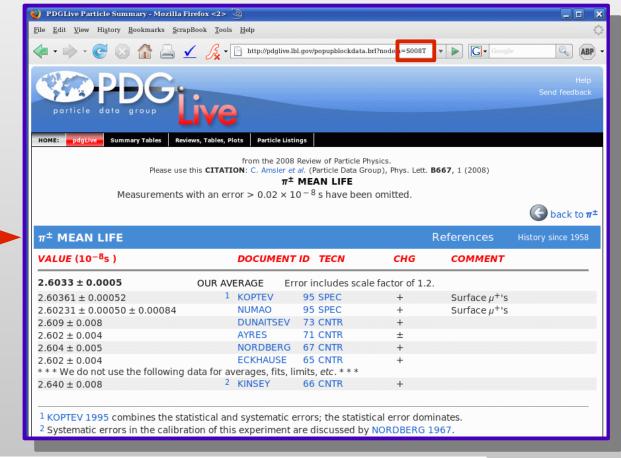


**S0087** 

#### From Node to Full Data Block



- The classification code is a unique string that leads directly to the corresponding PDG data
  - Similarly for PDG reviews



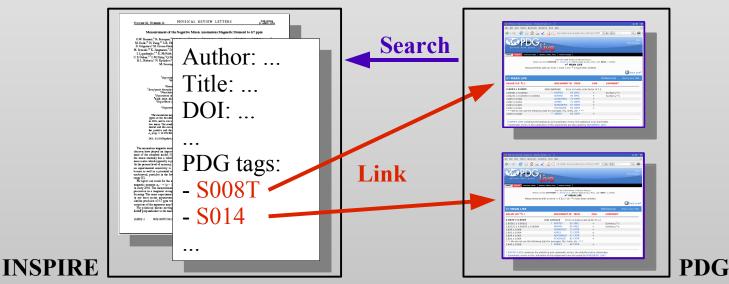


## **Common Classification**



#### Proposal:

- Agree on fine-grained classification scheme for HEP results, similar to the one presently used by PDG
- Allow classification codes in INSPIRE's article metadata
- Ask authors to tag their articles with classification codes through a convenient interface, e.g. when submitting a preprint to arXiv
  - Tags might be entered by navigating web pages similar to pdgLive
  - Might allow others to submit or vote on tags



AWDENDE BERKELEY NATIONAL LABORATE



## **Common Classification**



#### Benefits:

- Allows searches within INSPIRE and pdgLive based on a well-known classification
  - "What articles contain information about the top mass?"
- INSPIRE can directly link from article to relevant PDG data
  - By pointing at PDG pages using classification tags
  - Independently of whether article has been included in RPP
  - Can point to both PDG Listings and Reviews
- pdgLive can point user to articles that have not been included in PDG but might still be relevant for the user
  - By searching INSPIRE for the corresponding classification tag
  - Example: most recent unpublished top mass results
- Helps PDG to efficiently scan HEP literature for relevant results to be included into RPP



## Your Ideas for PDG?



- What can PDG provide to you?
- What are your ideas of how PDG should evolve?

Please contact us – during this Summit or by e-mail to JBeringer@lbl.gov

Thank you!



## **Conclusions**



- Computing upgrade will allow implementation of new innovative features
- Continue collaboration started long ago with SLAC with INSPIRE regarding
  - Cross-linking
  - Information interchange
- Propose to collaborate on and implement a common fine-grained classification scheme for HEP results



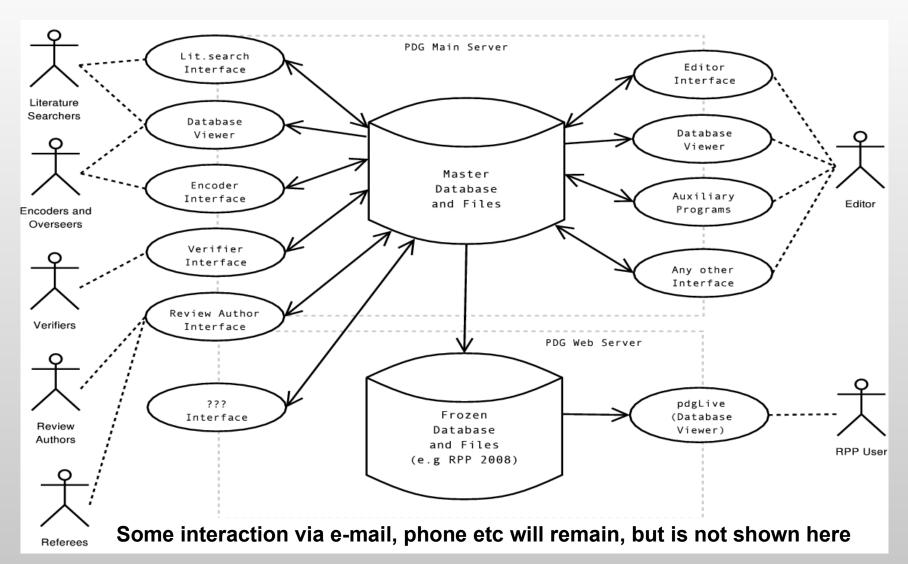


# **Backup Slides**



## **Architecture**

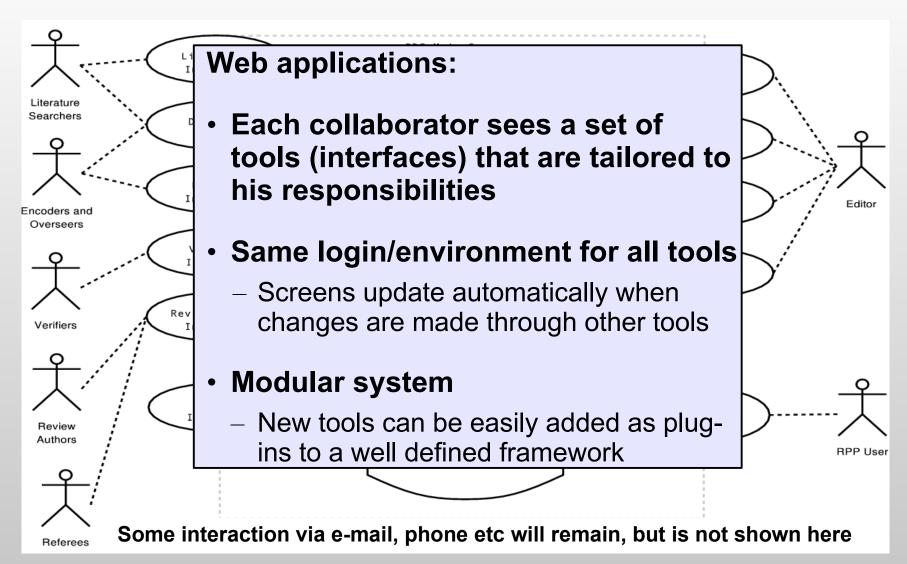






## **Architecture**







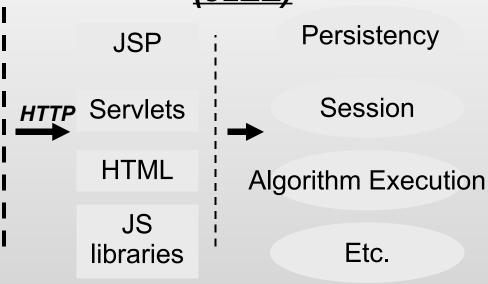
## **Three-Tier Web Application**



# Web Browser Web Application Server (J2EE)

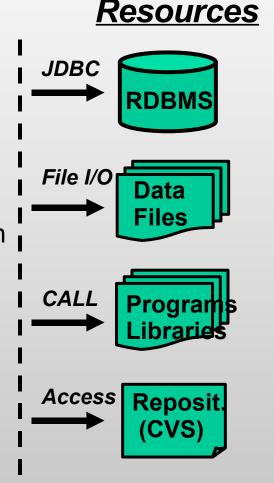
Execute
AJAX
enabled
Web pages
(HTML,
JavaScript)

- interact with user
- input processing
- dynamic page rendering



- generate dynamic HTML
- AJAX back end support
- static HTML pages
- JavaScript libraries
- rendering of formulas (isMath, mathML)

- object-relational mapping
- sessions support
- user authorization
- complex application logic
- interface to legacy code
- monitoring, etc.





## **Technologies**



#### J2EE-based web application framework

 Commonly used industry standard for building scalable, distributed web applications

#### AJAX-enabled web pages

User-friendly and highly interactive GUI behavior

#### Relational database (PostgreSQL)

O(100) database tables

#### Programming languages

- Java and JSP for web application framework backend
- JavaScript for client-side HTML (AJAX)
- Python API for programmatic access to database
- Legacy Fortran applications restructured as libraries