Status of the PDG Computing Upgrade

- Overall status
- Editor interface
- pdgLive
- Encoder interface
- Planning for phase 2
- Conclusions

Introduction

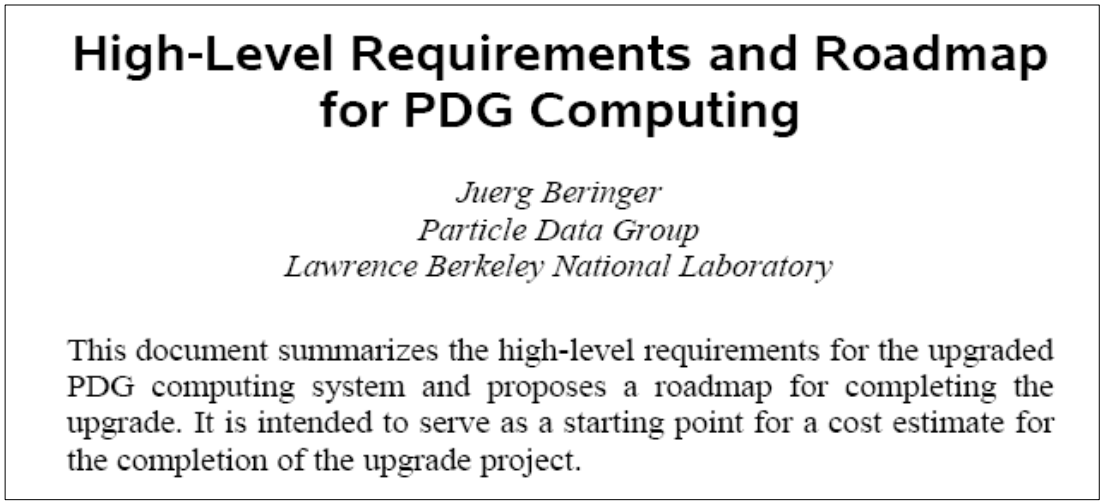
- "PDG Computing" refers to the infrastructure needed to **support the full process leading to the publication of the Review of Particle Physics**
  - In addition, must support diary production, ordering system, high-volume reliable web server w/mirrors, outreach activities (ATLAS, Particle Adventure, etc), institutional contacts database, reporting, statistics, monitoring ...
- Old computing system originated in early 1980s and was running on old hardware (without replacement system) and obsolete software
- Several upgrade attempts since mid 1990s did not converge
- LBNL-internal review in 2003 suggested giving highest priority to the completion of the "RPPs 1.5" prototype and let people work with it
  - In spite of concerns about some technology choices and maintainability
  - Upgrade will need to continue (with emphasis on long-term maintainability)
- Urgency of completing at least a partial upgrade increasingly evident in 2004

Upgrade Plan in 2004

- Obviously lacking sufficient resources to carry out a full upgrade on a short time-scale, we **decided in 2004 on a pragmatic (albeit not ideal) approach that would ensure our ability to continue producing the RPP**
- Defined 3 phases of the upgrade:
  - **Phase 1**: Switch to partially upgraded system in time for RPP 2006 production
    - Switch to modern hardware (Linux servers)
    - Reimplementation of existing Oracle/FORMS editor interface
    - If feasible, provide database viewer and initial version of encoder interface
  - **Phase 2**: Improve partially upgraded system in an incremental fashion
    - Address technology choices, long-term maintainability, and documentation
    - Improve or replace existing interfaces as necessary
    - Add new tools (e.g. for handling of Reviews)
  - **Phase 3** (if deemed necessary): Completely redesign database structure
    - Appears to be not necessary

Status Today

- **Phase 1 is essentially complete**
  - Switched RPP data entry and production to the new system on 9/13/2005 ...
  - ... and RPP 2006 was produced completely with the new system!
  - pdgLive – a database viewer with the full core functionaliy – became available at the same time as the "traditional" 2006 web edition
  - An advanced prototype of the encoder interface has been tested by selected PDG persons but remains to be debugged completely before being put into production
- **Phase 2: Planning and cost estimate are in progress**



High-Level Requirements and Roadmap for PDG Computing

Juerg Beringer
Particle Data Group
Lawrence Berkeley National Laboratory

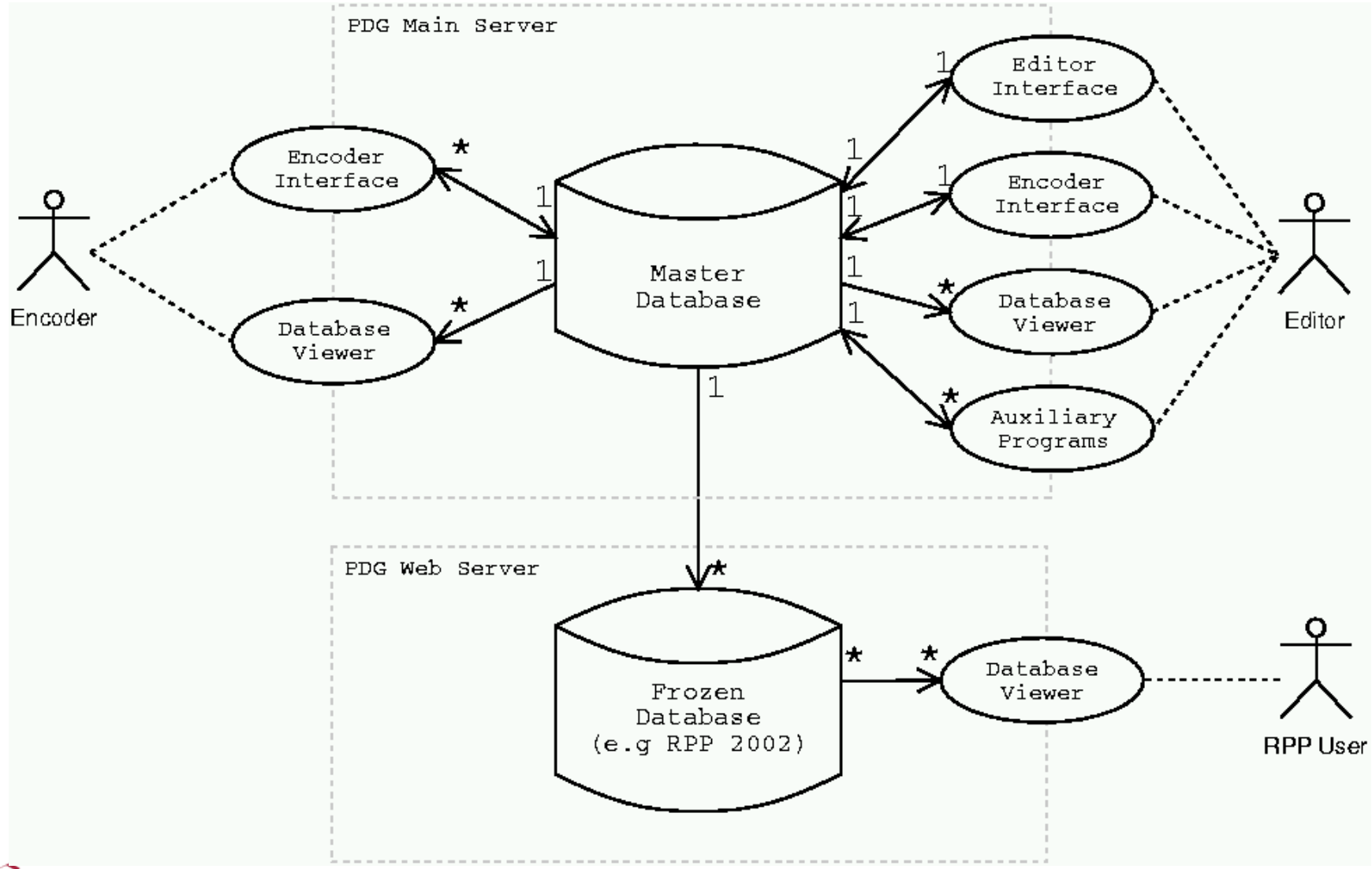
This document summarizes the high-level requirements for the upgraded PDG computing system and proposes a roadmap for completing the upgrade. It is intended to serve as a starting point for a cost estimate for the completion of the upgrade project.

Recent Contributors to PDG Computing Upgrade

- From COMPAS group, IHEP Protvino:
  - Kirill Lugovsky (web interfaces)
  - Slava Lugovsky (web interfaces)
  - Vitaly Lugovsky (core libraries, database, left in 2004)
  - Lyudmila Lugovskaya (documentation, left in 2004)
  - Vladimir Ezhela (group leader, retired)
  - Oleg Zenin (group leader, new)
- From LBNL:
  - Juerg Beringer (project leader, since March 2004)
  - Orin Dahl (auxiliary programs, Oracle/FORMS related work, retired)
  - Piotr Zyla (daily operation, production tasks, editor interface)

These are all part-time contributors, mostly at the 10% to 70% level

High-Level Architecture of Upgraded System



Implementation

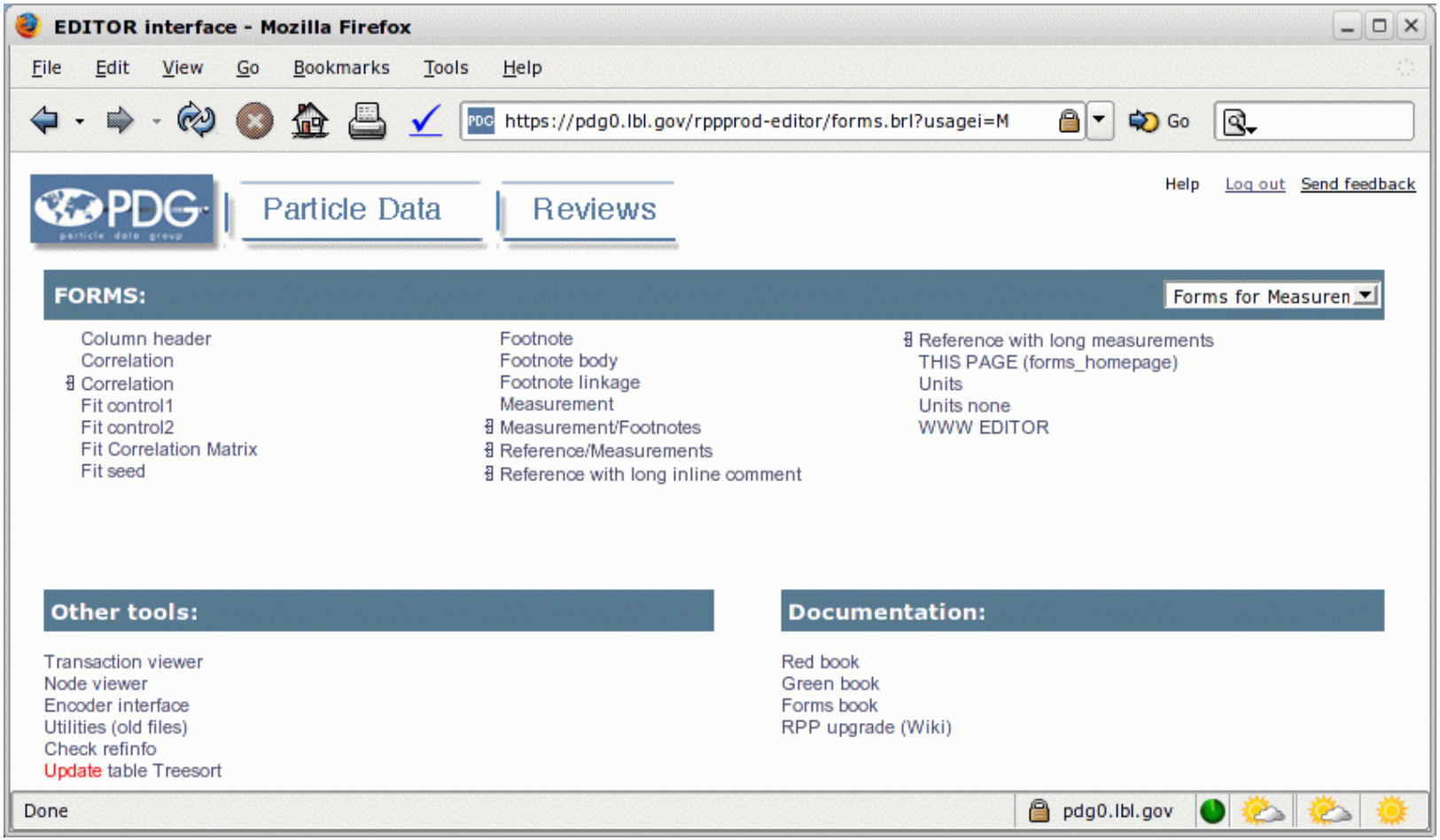
- 2 Linux servers (dual-CPU, 2GB RAM each, total 1 TB RAID disk)
- PostgreSQL – mature relational database
  - Interfaced through JDBC
  - Could easily switch to MySQL or Oracle
- Kawa 1.7 – Java based Scheme (Lisp) system
- BRL 2.2.1 – system to embed dynamic content into web pages
  - Similar in spirit to JSP, but uses the Scheme language instead of Java
- Apache Tomcat – servlet container
  - used to run BRL / Kawa within the web server
- Apache web server
- HTML and JavaScript
- Mimetex – standalone tool to generate gif images from TeX snippets

Maintainability Concerns

- **Primary issue is use of Kawa and BRL** regarding
  - Future support for these tools
  - Suitability of Scheme/Lisp for complex web applications
- Kawa (http://www.gnu.org/software/kawa/)
  - Actively maintained by author, small user community, active discussion lists
- BRL (http://brl.sourceforge.net)
  - Based on Kawa, no significant development since 2003
- Possible alternative similar Scheme-based tools exist, e.g. Bigloo, SISC, and SISCweb, but none of them widely used
- Use of Scheme/Lisp for writing web applications rather exotic
- Present code has evolved over many iterations with changing requirements and designs, and needs to be refactored (independently of language choice)

Editor Interface

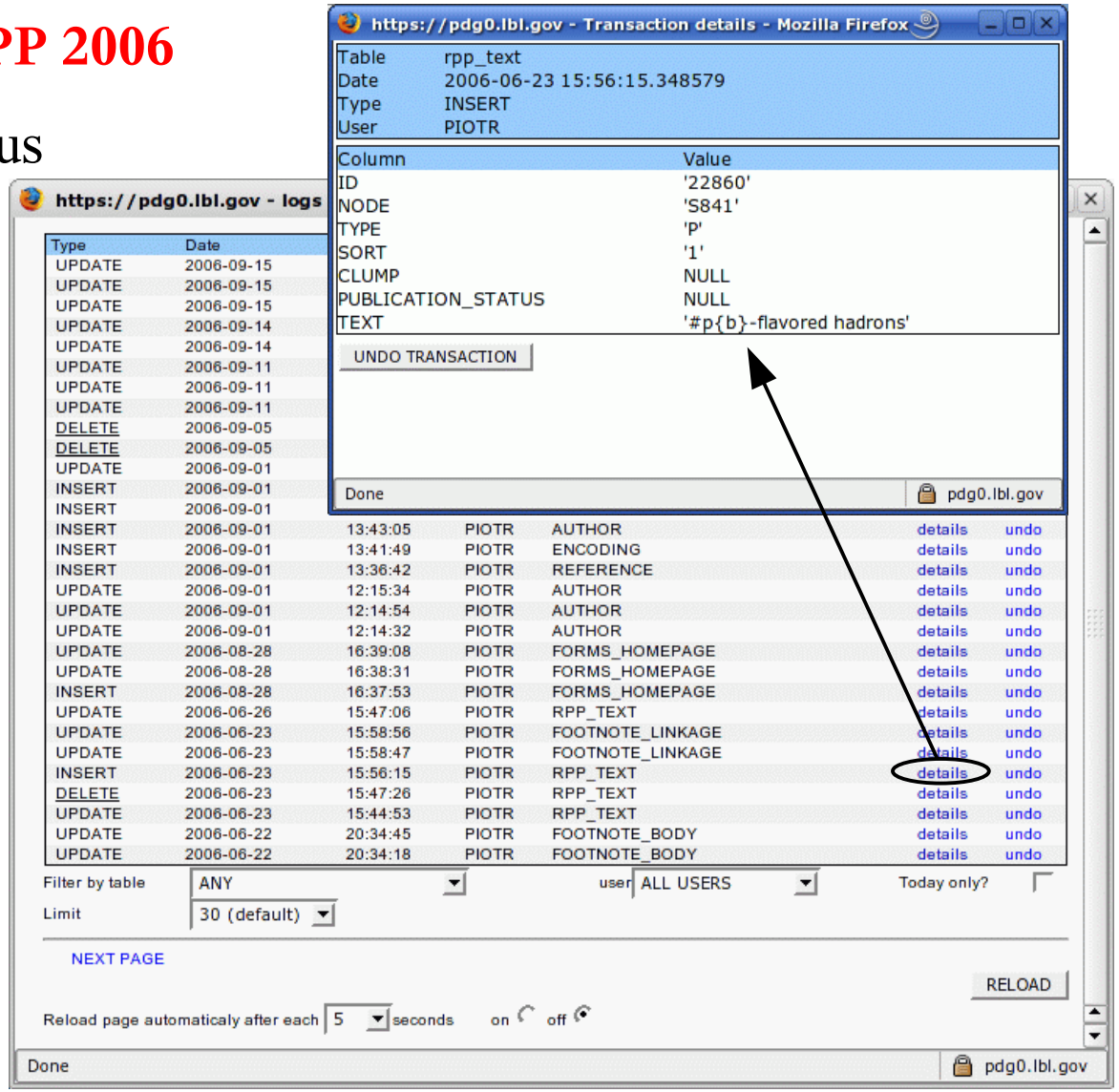
- An expert-only web-based GUI to edit the raw content of the PDG database
- Knows about connections between tables and constraints on input values



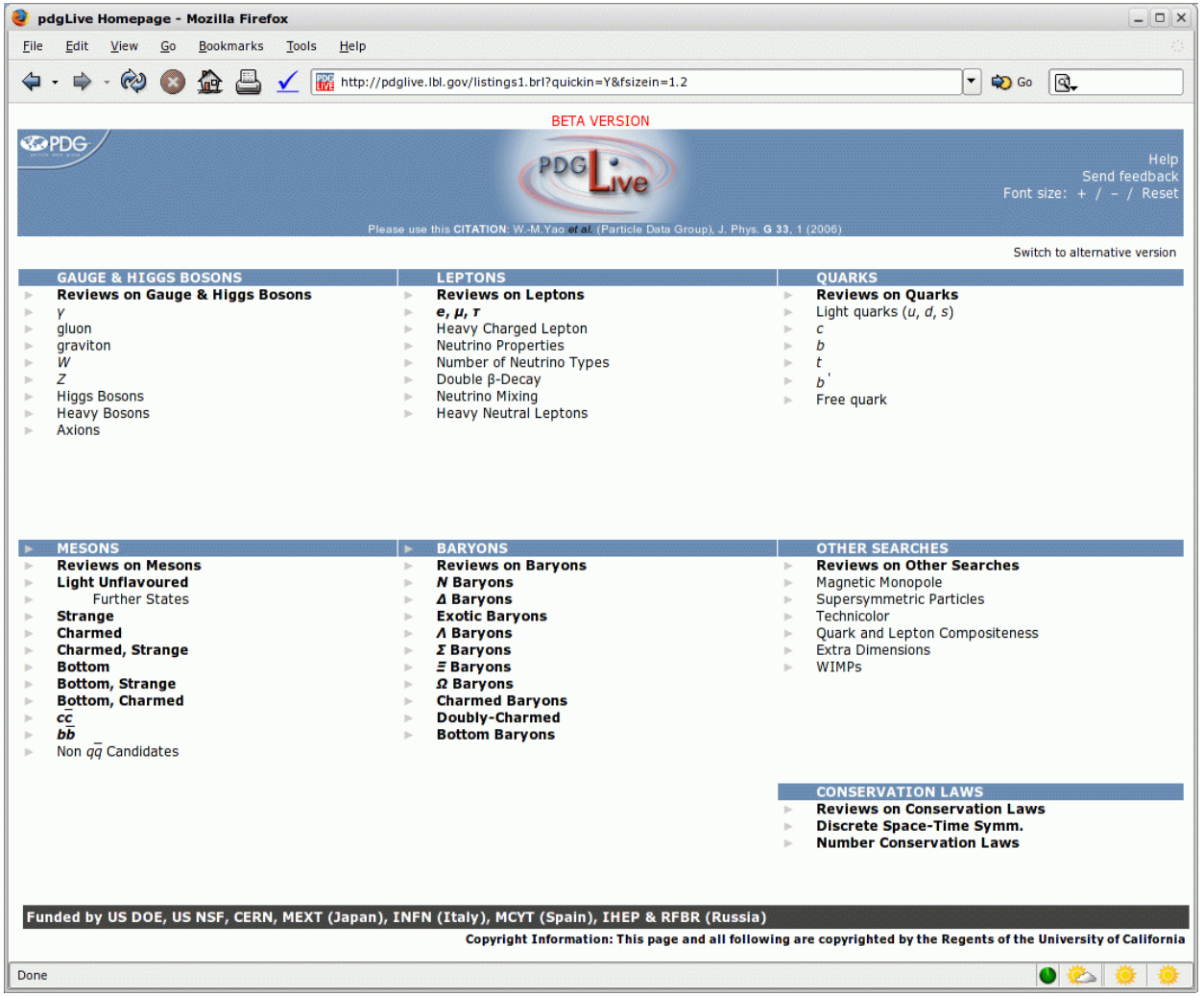

https://pdg0.lbl.gov - REFERENCE MEASUREMENT form - Mozilla Firefox
1. Search for a publication
2. Choose desired entry
3. Display and edit measurements
save
Find
reset
QUERY
Done
pdg0.lbl.gov

Status of Editor Interface

- **Completed in time for RPP 2006**
- Full replacement of previous Oracle/FORMS interface with several improvements:
  - Database transaction logger with “undo”
  - Improved access to tree table
  - Easily customizable to support new database tables
- Sufficient for short-term
- Maintenance issues



Database Viewer (pdgLive)



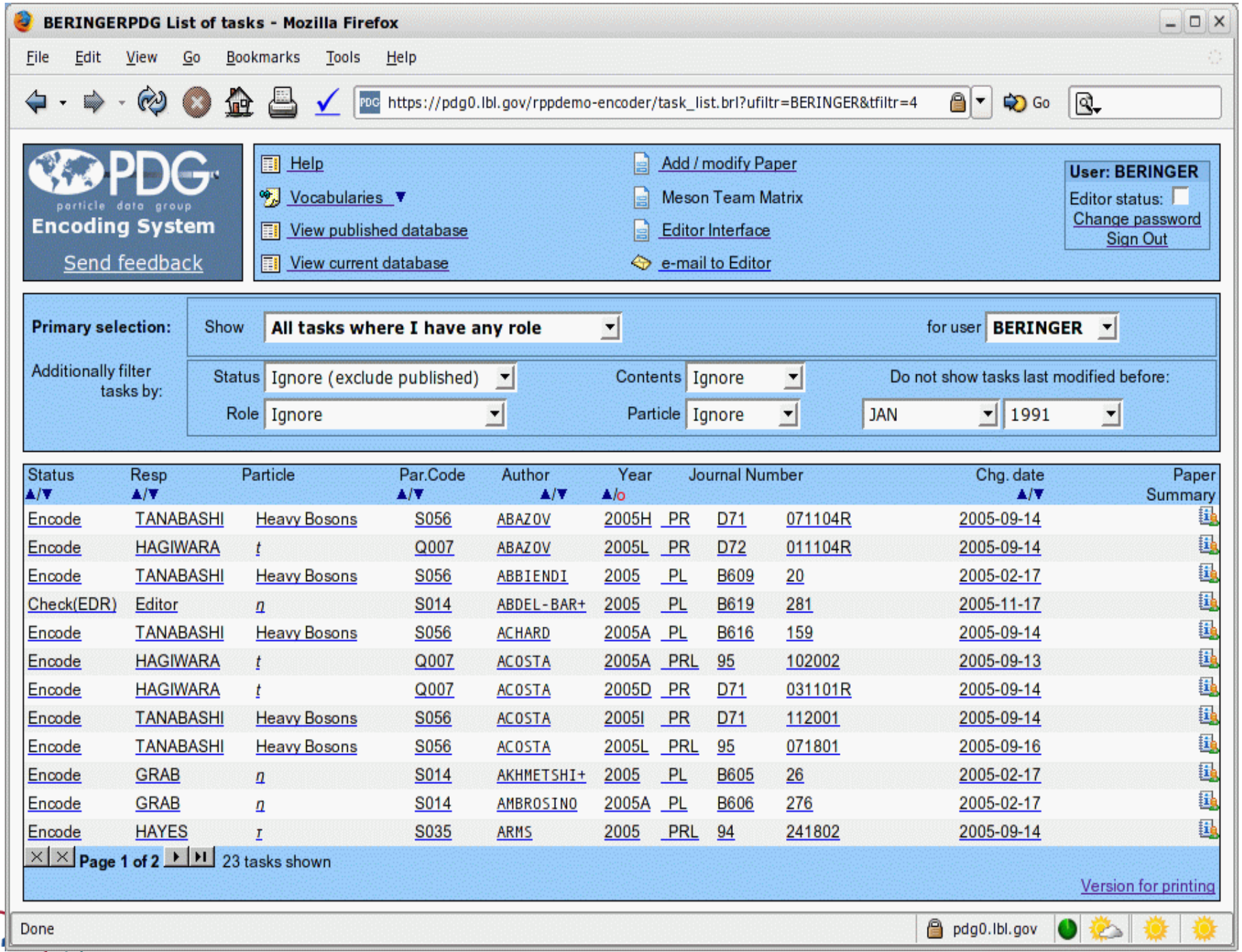
- A web application to browse the contents of the PDG database
- **Used for pdgLive, our public interface:** http://pdgLive.lbl.gov
  - See demonstration
- **Beta version of pdgLive announced publicly 7/18/2006**
- Used internally to view master database
- Why "beta"? Mainly maintenance concerns

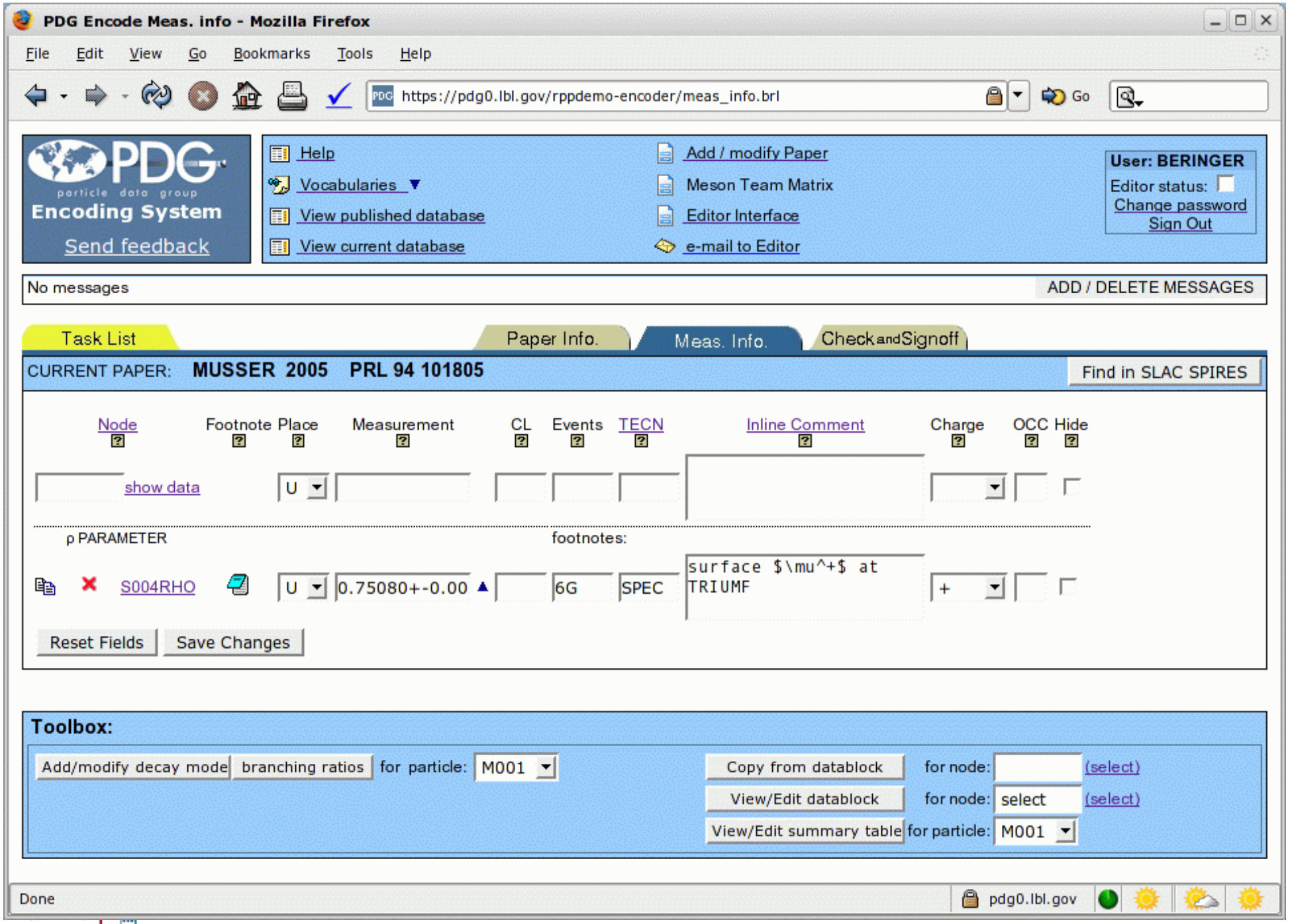
What Users Told Us About pdgLive

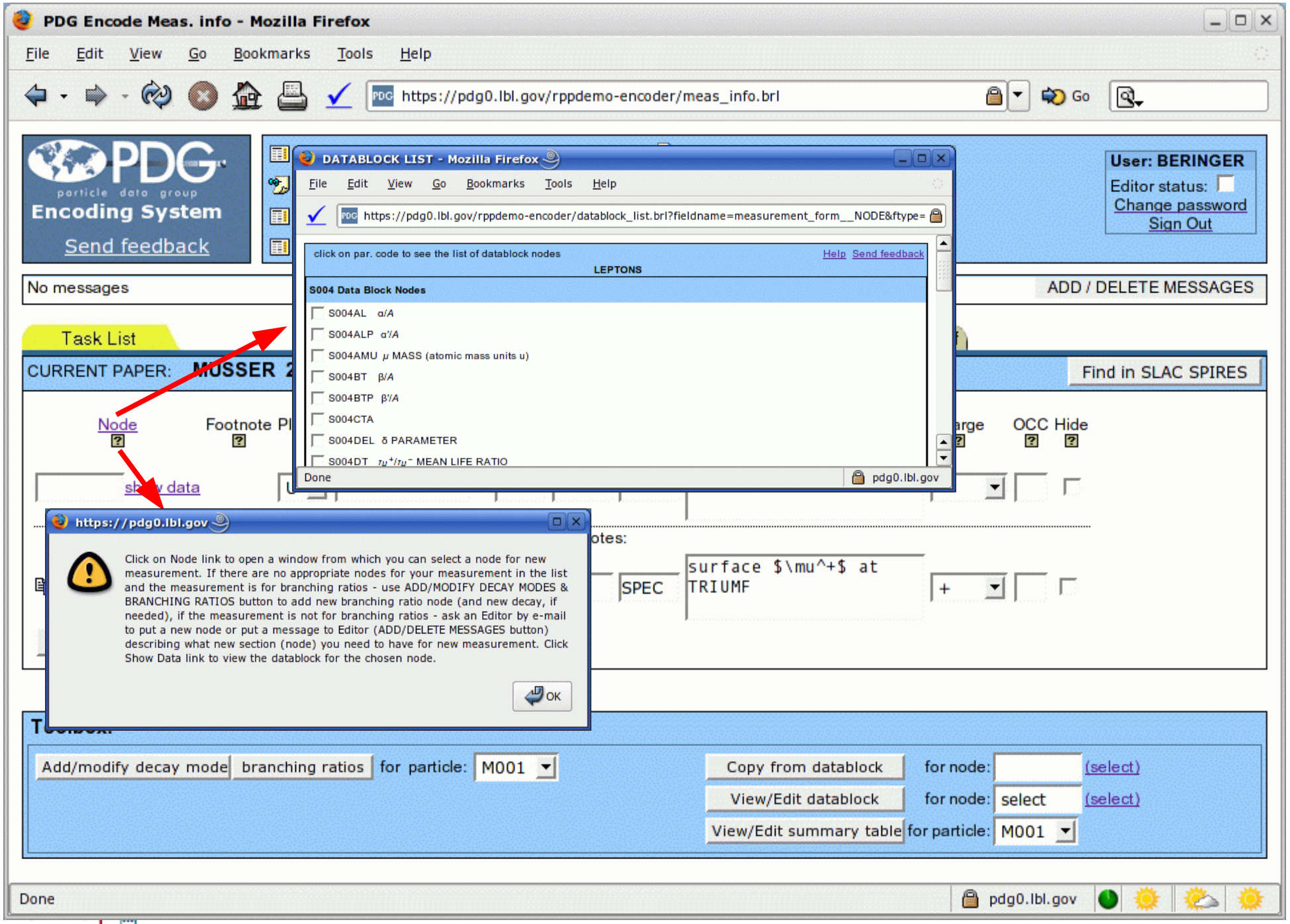
- "This is one of the nicest, clearest, and most useful Web pages in HEP. Superb job."
- "PDG Live is absolutely wonderful. ... Plus, it's really beautiful and lots of fun. I'm so excited I can't even tell you."
- "Bravo! This is the way I always hoped to browse PDG on the web!"
- "Question: What size army of graduate students was utilised to get pdglive accomplished?"
- ...
- Suggestions:
  - Add links to all reviews (not only the mini-reviews in the Listings)
  - Make decay products link to the corresponding particles
- No bugs reported so far ... 

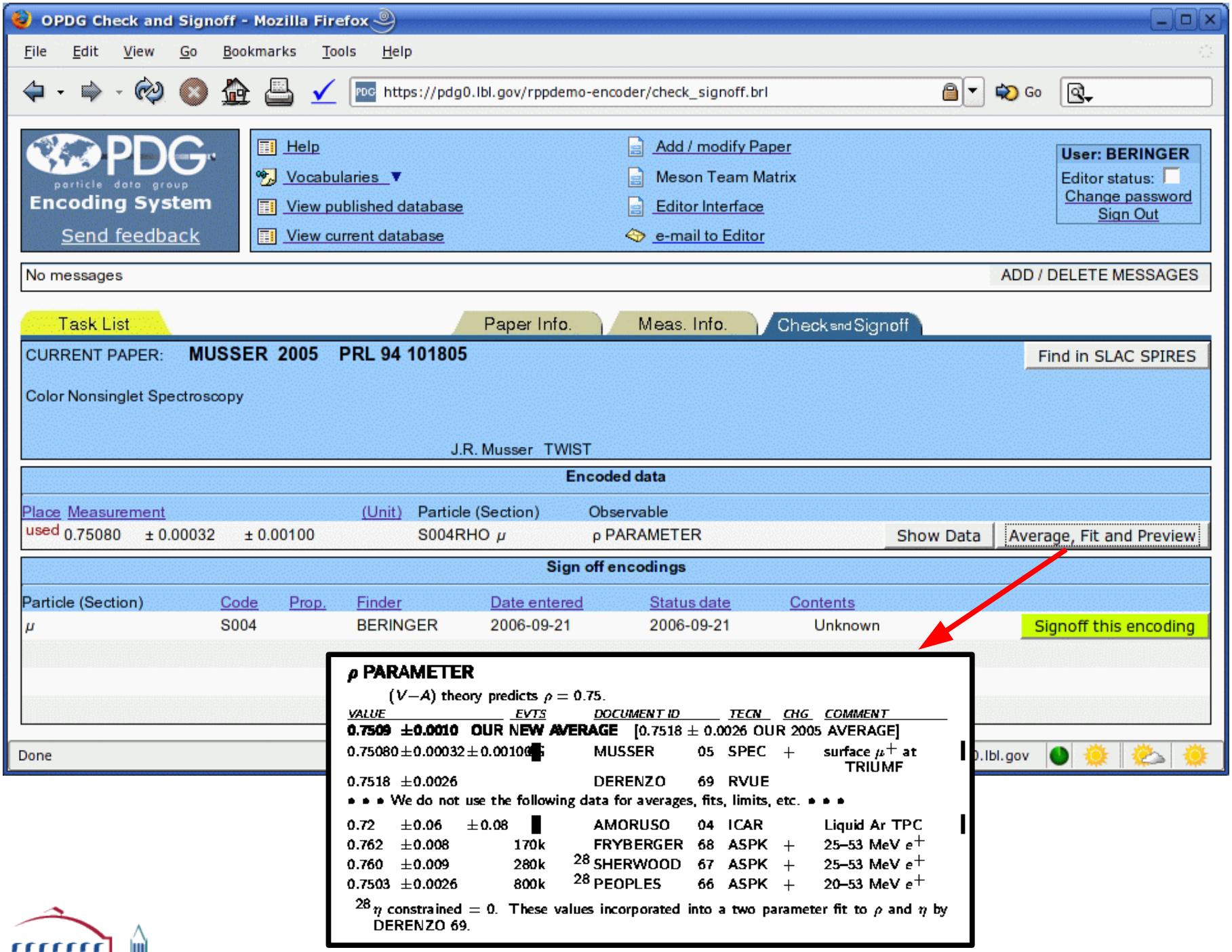
Encoder Interface

- Encoder interface is a web-based application that allows
  - Encoders to enter new measurements
  - Overseers to check, edit and sign off on new entries, fits and averages
  - Editor to check, edit and sign off on new entries
- Encoders can enter (most) encodings entirely themselves
  - **Task oriented system** guides user and knows what pieces of information need to be entered and how to store them into the underlying database
  - Immediate feedback on how it will look in the Listings
  - Allows editor to concentrate on his main tasks
- Simplifies encoding process by managing flow of information between encoders/overseers/coordinators/editor and (eventually) verifiers
- Make precise reports about actual status of encoding possible
- Special tools for meson team


BERINGERPDG List of tasks - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://pdg0.lbl.gov/rppdemo-encoder/task_list.brl?ufiltr=BERINGER&tfiltr=4
Go
PDG particle data group
Encoding System
Send feedback
Help
Vocabularies
View published database
View current database
Add / modify Paper
Meson Team Matrix
Editor Interface
e-mail to Editor
User: BERINGER
Editor status:
Change password
Sign Out
Primary selection:
Show All tasks where I have any role
for user BERINGER
Additionally filter tasks by:
Status Ignore (exclude published)
Contents Ignore
Do not show tasks last modified before:
Role Ignore
Particle Ignore
JAN 1991
Status Resp Particle Par.Code Author Year Journal Number Chg. date Paper Summary
Encode TANABASHI Heavy Bosons S056 ABAZOV 2005H PR D71 071104R 2005-09-14
Encode HAGIWARA t Q007 ABAZOV 2005L PR D72 011104R 2005-09-14
Encode TANABASHI Heavy Bosons S056 ABBIENDI 2005 PL B609 20 2005-02-17
Check(EDR) Editor η S014 ABDEL-BAR+ 2005 PL B619 281 2005-11-17
Encode TANABASHI Heavy Bosons S056 ACHARD 2005A PL B616 159 2005-09-14
Encode HAGIWARA t Q007 ACOSTA 2005A PRL 95 102002 2005-09-13
Encode HAGIWARA t Q007 ACOSTA 2005D PR D71 031101R 2005-09-14
Encode TANABASHI Heavy Bosons S056 ACOSTA 2005I PR D71 112001 2005-09-14
Encode TANABASHI Heavy Bosons S056 ACOSTA 2005L PRL 95 071801 2005-09-16
Encode GRAB η S014 AKHMETSHI+ 2005 PL B605 26 2005-02-17
Encode GRAB η S014 AMBROSINO 2005A PL B606 276 2005-02-17
Encode HAYES τ S035 ARMS 2005 PRL 94 241802 2005-09-14
Page 1 of 2 23 tasks shown
Version for printing
Done
pdg0.lbl.gov


PDG Encode Meas. info - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://pdg0.lbl.gov/rppdemo-encoder/meas_info.brl
Go
PDG particle data group
Encoding System
Send feedback
Help
Vocabularies
View published database
View current database
Add / modify Paper
Meson Team Matrix
Editor Interface
e-mail to Editor
User: BERINGER
Editor status:
Change password
Sign Out
No messages
ADD / DELETE MESSAGES
Task List
Paper Info.
Meas. Info.
CheckandSignoff
CURRENT PAPER: MUSSER 2005 PRL 94 101805
Find in SLAC SPIRES
Node
Footnote
Place
Measurement
CL
Events
TECN
Inline Comment
Charge
OCC
Hide
show data
U
ρ PARAMETER
footnotes:
S004RHO
U
0.75080+-0.00
6G
SPEC
surface μ^+ at TRIUMF
+
Reset Fields
Save Changes
Toolbox:
Add/modify decay mode
branching ratios
for particle:
M001
Copy from datablock
for node:
(select)
View/Edit datablock
for node:
select
(select)
View/Edit summary table
for particle:
M001
Done
pdg0.lbl.gov


PDG Encode Meas. info - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://pdg0.lbl.gov/rppdemo-encoder/meas_info.brl
PDG particle data group
Encoding System
Send feedback
DATABLOCK LIST - Mozilla Firefox
https://pdg0.lbl.gov/rppdemo-encoder/datablock_list.brl?fieldname=measurement_form__NODE&ftype=
click on par. code to see the list of datablock nodes
Help Send feedback
LEPTONS
S004 Data Block Nodes
S004AL α/A
S004ALP α'/A
S004AMU μ MASS (atomic mass units u)
S004BT β/A
S004BTP β'/A
S004CTA
S004DEL δ PARAMETER
S004DT τμ+/τμ− MEAN LIFE RATIO
Done
pdg0.lbl.gov
User: BERINGER
Editor status:
Change password
Sign Out
No messages
ADD / DELETE MESSAGES
Task List
CURRENT PAPER:
Find in SLAC SPIRES
Node
Footnote Pl
OCC Hide
https://pdg0.lbl.gov
Click on Node link to open a window from which you can select a node for new measurement. If there are no appropriate nodes for your measurement in the list and the measurement is for branching ratios - use ADD/MODIFY DECAY MODES & BRANCHING RATIOS button to add new branching ratio node (and new decay, if needed), if the measurement is not for branching ratios - ask an Editor by e-mail to put a new node or put a message to Editor (ADD/DELETE MESSAGES button) describing what new section (node) you need to have for new measurement. Click Show Data link to view the datablock for the chosen node.
OK
SPEC
surface μ^+ at TRIUMF
Add/modify decay mode
branching ratios
for particle:
M001
Copy from datablock
for node:
(select)
View/Edit datablock
select
View/Edit summary table
Done
pdg0.lbl.gov

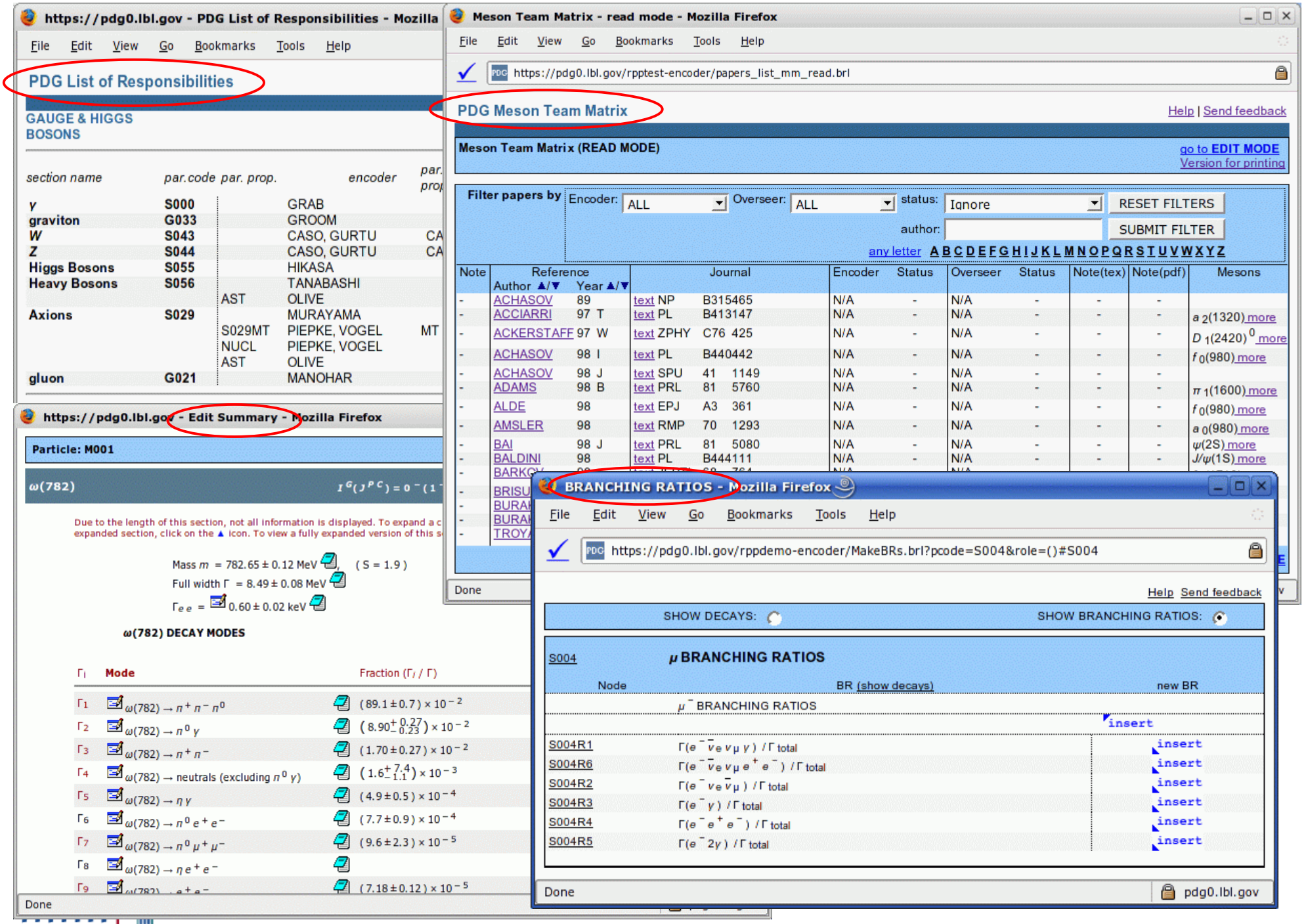


ρ PARAMETER

$(V-A)$ theory predicts $\rho = 0.75$.

VALUE	EVTS	DOCUMENT ID		TECN	CHG	COMMENT
0.7509 ±0.0010 OUR NEW AVERAGE [0.7518 ± 0.0026 OUR 2005 AVERAGE]						
0.75080±0.00032±0.00100		MUSSER	05	SPEC	+	surface μ^+ at TRIUMF
0.7518 ±0.0026		DERENZO	69	RVUE		
• • • We do not use the following data for averages, fits, limits, etc. • • •						
0.72 ±0.06 ±0.08		AMORUSO	04	ICAR		Liquid Ar TPC
0.762 ±0.008	170k	FRYBERGER	68	ASPK	+	25–53 MeV e^+
0.760 ±0.009	280k	[28] SHERWOOD	67	ASPK	+	25–53 MeV e^+
0.7503 ±0.0026	800k	[28] PEOPLES	66	ASPK	+	20–53 MeV e^+

[28] η constrained = 0. These values incorporated into a two parameter fit to ρ and η by DERENZO 69.

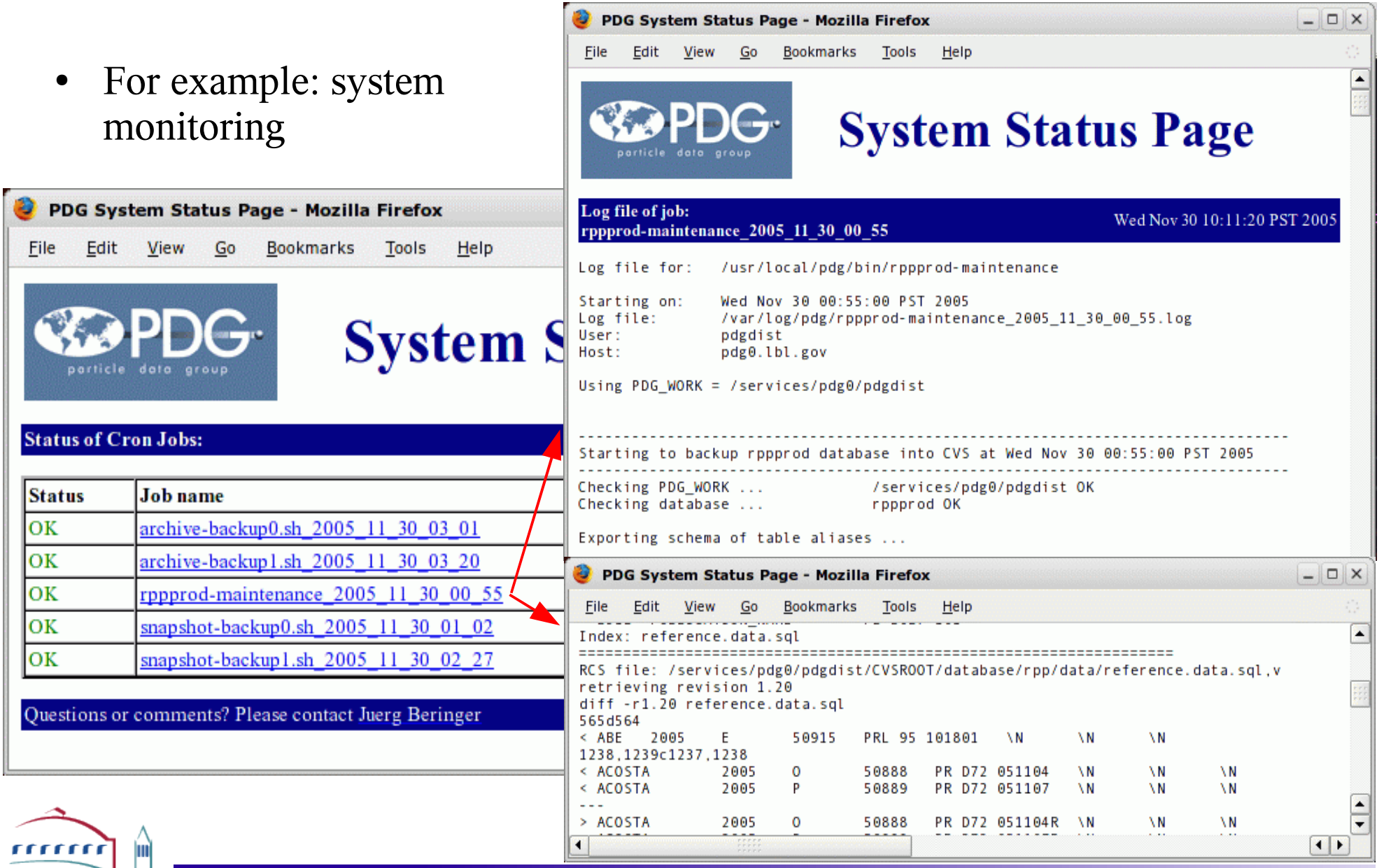

https://pdg0.lbl.gov - PDG List of Responsibilities - Mozilla
File Edit View Go Bookmarks Tools Help
PDG List of Responsibilities
GAUGE & HIGGS BOSONS
section name par.code par. prop. encoder par. prop
γ S000 GRAB
graviton G033 GROOM
W S043 CASO, GURTU CA
Z S044 CASO, GURTU CA
Higgs Bosons S055 HIKASA
Heavy Bosons S056 TANABASHI
AST OLIVE
Axions S029 MURAYAMA
S029MT PIEPKE, VOGEL MT
NUCL PIEPKE, VOGEL
AST OLIVE
gluon G021 MANOHAR
Meson Team Matrix - read mode - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://pdg0.lbl.gov/rpptest-encoder/papers_list_mm_read.brl
PDG Meson Team Matrix
Help | Send feedback
Meson Team Matrix (READ MODE)
go to EDIT MODE
Version for printing
Filter papers by Encoder: ALL Overseer: ALL status: Ignore RESET FILTERS
author: SUBMIT FILTER
any letter A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Note Reference Author Year Journal Encoder Status Overseer Status Note(tex) Note(pdf) Mesons
ACHASOV 89 text NP B315465 N/A N/A
ACCIARRI 97 T text PL B413147 N/A N/A
ACKERSTAFF 97 W text ZPHY C76 425 N/A N/A
ACHASOV 98 I text PL B440442 N/A N/A
ACHASOV 98 J text SPU 41 1149 N/A N/A
ADAMS 98 B text PRL 81 5760 N/A N/A
ALDE 98 text EPJ A3 361 N/A N/A
AMSLER 98 text RMP 70 1293 N/A N/A
BAI 98 J text PRL 81 5080 N/A N/A
BALDINI 98 text PL B444111 N/A N/A
Done
https://pdg0.lbl.gov - Edit Summary - Mozilla Firefox
Particle: M001
ω(782)
Due to the length of this section, not all information is displayed. To expand a c
expanded section, click on the ▲ icon. To view a fully expanded version of this s
Mass m = 782.65 ± 0.12 MeV (S = 1.9)
Full width Γ = 8.49 ± 0.08 MeV
Γee = 0.60 ± 0.02 keV
ω(782) DECAY MODES
Γi Mode Fraction (Γi / Γ)
Γ1 ω(782) → π+ π− π0 (89.1 ± 0.7) × 10−2
Γ2 ω(782) → π0 γ (8.90 +0.27 −0.23) × 10−2
Γ3 ω(782) → π+ π− (1.70 ± 0.27) × 10−2
Γ4 ω(782) → neutrals (excluding π0 γ) (1.6 +7.4 −1.1) × 10−3
Γ5 ω(782) → η γ (4.9 ± 0.5) × 10−4
Γ6 ω(782) → π0 e+ e− (7.7 ± 0.9) × 10−4
Γ7 ω(782) → π0 μ+ μ− (9.6 ± 2.3) × 10−5
Γ8 ω(782) → η e+ e−
Γ9 (7.18 ± 0.12) × 10−5
Done
BRANCHING RATIOS - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
https://pdg0.lbl.gov/rppdemo-encoder/MakeBRs.brl?pcode=S004&role=()#S004
Help Send feedback
SHOW DECAYS: SHOW BRANCHING RATIOS:
S004 μ BRANCHING RATIOS
Node BR (show decays) new BR
μ− BRANCHING RATIOS
insert
S004R1 Γ(e− ν̄e νμ γ) / Γ total insert
S004R6 Γ(e− ν̄e νμ e+ e−) / Γ total insert
S004R2 Γ(e− νe ν̄μ) / Γ total insert
S004R3 Γ(e− γ) / Γ total insert
S004R4 Γ(e− e+ e−) / Γ total insert
S004R5 Γ(e− 2γ) / Γ total insert
Done
pdg0.lbl.gov

Status of Encoder Interface

- **Prototype with major features available and awaits final debugging**
  - Already went through several rounds of user testing (and improvements)
  - Completing debugging/testing is main goal of this visit of Kirill and Slava
  - Hope to start using encoder interface in production system for a few select users before the end of this year
- Restrictions in the present version (ie things that must be done by editor)
  - Cannot create arbitrary new nodes (can create new nodes for branching ratios)
  - Cannot add new particles
  - Cannot change which nodes are used in fits
  - Deliberately do not allow some operations (delete, certain changes to tree)
- Main concerns:
  - Potential concurrency issues
  - Long-term maintainability

Many Other Improvements from Phase 1

- For example: system monitoring



A Few Lessons from Phase 1

- High-level design achieved desired isolation of components, greatly simplifying development and deployment
- The time Piotr (testing) and myself (project management, architectural work, testing) can spend on the upgrade is one of the key limiting factors
  - Led to postponing final debugging and deployment of encoder interface
  - In 2004/5 I spent far more time on PDG computing than allocated
    - Cannot be repeated - significant slowdown of progress this year
- While we did succeed in completing production-quality applications in spite of exotic programming language, ...
- Experience of debugging clearly demonstrates that rewriting of existing code base is unavoidable in the long term
- At database level, primarily need a better separation between scientific content and output format (need to define a meta-language)

So What's Next?

- Completion of phase 1 eliminates risk from possible failure of old system
- But it is only the first step:
  - **We do not yet have a system that will be maintainable in the long term**
  - Present system does not address the full functionality
    - No support for handling reviews, verification process, several other issues
- Extensive full-day review by LBNL Physics Division on December 6, 2005
  - Reviewers agreed with our assessment regarding long-term maintainability
  - Recommended replacing tools from phase 1 with new applications developed using industry standard tools and practices, based on proper requirements and design studies for each key component
    - Preliminary project estimate by LBNL IT Division (Jeff Willer) gives cost of about $45k for completing analysis, functional review, and proper cost estimate
  - Recognized the need for substantial additional resources

Planning for Phase 2

- High-level requirements need to include everything we'll ultimately need
- In addition to the primary applications (editor and encoder interface, auxiliary programs, database viewer/pdgLive), will ultimately need at least:

  - Review author interface
  - Interface for updating institution database
  - Mailing system
  - Product ordering system
  - System monitoring
  - Status reporting
  - User profile management

- **Taking into account dependencies among components, required effort and expected benefits, developed a preliminary roadmap for phase 2**
  - Roadmap document describes high-level requirements, milestones and tentative target dates for phase 2
  - **Actual progress of phase 2 will depend greatly on when and how much additional resources will become available**
    - Need to keep producing the book in parallel to work on computing upgrade!

Milestones and Tentative Schedule

NOTE: the dates below are merely a first guess and are extremely preliminary!

Tentative target date	*Milestone*
As soon as possible	Estimate cost, identify resources, and agree on schedule
November 2006	Complete debugging of and deploy existing encoder interface prototype
July 2007	Agree on database schema and meta-language
July 2007	Complete and deploy the new system for managing user information (including all relevant applications)
October 2007	Complete review author interface
July 2008	Complete and deploy new encoder interface
2008	Complete new database viewer
2008/9	Complete update of programs and editor interface (if needed)

Conclusions

- **Completion of all essential parts (and more!) of phase 1 in time for production of RPP 2006 with the upgraded system is a major success**
- Existing system plus soon to-be-debugged encoder interface is expected to be sufficient for us to function until upgrade will be complete
- Main purpose of phase 2 is to arrive at a computing system that
  - Will be maintainable in the long term (10 years), is based on sound technology choices and is well documented
  - Provides at least the same essential functionality available now in the editor interface, pdgLive and prototype encoder interface, but will hopefully be complemented by additional new applications helping us to streamline the book production and work more efficiently
- **Carrying out phase 2 will require a significant effort (4 FTE-years?) that will only be possible through additional resources**
  - Roadmap document provides high-level requirements and tentative milestones
  - Actively working on obtaining additional resources