

# Community White Paper Editorial Board Meeting

---

Introduction

2017-11-01

# First Draft

- Prepared by four people for efficiency in getting something out of the door in a timely fashion (Peter Elmer, John Harvey, Michel Jouvin, Graeme Stewart)
- First draft was released Friday 20th October
  - [Google Doc](#) is attracting comments - from ~15 people (hard to count that); we know more have read it
  - There is a [PDF](#) that can also be reviewed
- Graeme gave a talk on the current draft in the Scientific Computing Forum
  - <https://indico.cern.ch/event/663273/>
- Summary seems to be that no major problems have been identified so far
  - There are some larger questions that we should address as the EB however
    - Some of these we pulled out (see later slides), there may be others
  - Individual working groups have been asked to review the summary of their work **with urgency**

# Missing Pieces

- Event Generators
  - Driven so far by the generator community (MCnet) but no progress: it seems that there are some contentious issues
  - Should we take the ownership: focus on computing challenges linked to generators rather than generators themselves: possible input by experiments?
- Workload Management
  - Partially covered in Event Frameworks and Distributed Computing
  - There is a working draft in preparation - Kaushik is organising this and marshalling input from some others
    - Is this sufficient to represent a community consensus though?
- Security
  - Romain Wartel suggested we could add something based on the works going in WLCG & all
  - Will prepare something to be integrated in draft2: first draft expected next week

# Larger Questions

- Do we agree on the intended audience of the document?
  - HEP community itself
    - So no 'introduction to particle physics'
    - But we should avoid, or at least support understanding of, technical details (glossary?) by non-specialists
- Would people like to see the R&D proposals organised in a different way?
  - Currently alphabetical, which is very neutral
    - However, it's not so useful for identifying the more critical areas (e.g., track reconstruction vs. visualisation)
  - Many different dimensions to the work: technical, physics, software, distributed computing, criticality
    - So this would involve far more value judgement from us, but may provide a more useful organisation for the reader

# Larger Questions

- What assessment of priority can we or should we give?
  - Easy to identify critical *areas* of work
    - Done in Chapter 5, very much driven by global resource implications
  - Harder to achieve *community consensus* on the specific R&D proposals themselves
    - We are not a funding or a management body
      - So we have no money/effort resource envelope to guide us
    - Also the proposals themselves are speculative in many areas
      - And we have no idea what the eventual benefits might be (e.g., of a Data Lake or some ML application)
- Signatories
  - We want the document to be signed by those who give it their broad support
  - This partially comes from the composition of this board
    - But it should go wider and we should target specifically experiment and regional representatives

# Timeline To Completion

- Original goal was August this year
- We are cursed by all deadlines being soft
  - So we have to impose discipline on ourselves and others involved

## Proposal:

- Close first draft for comments at the end of next week - 10 November
  - *Your input as the Editorial Board is also really important!*
- Release second draft a week later - 17 November
  - Only possible if we begin to work on this now
- Close second draft after two weeks - 1 December
- Final version about two weeks after that - ~15 December

List of signatories: we need to do some work in parallel - at least gather some tentative approval in advance of the final version

- Pete started a [Google Doc](#) for that

# Practical Next Steps

- 2 potentially contradictory objectives
  - Keep the initial draft stable so that comments are consistent, based on the same version
  - Get a second draft out quickly... which involves a lot of work that cannot be done in a couple of hours
- Need to work on the second draft in parallel with people commenting the first draft
  - Implies working on a “shadow copy”
  - We need to make sure we don’t lose comments
  - Keeping track of who is addressing what
  - Achieving consensus on any controversial points

# Second Draft Proposal

Create a copy of the first draft asap: GoogleDocs doesn't allow to copy comments but a Chrome extension exists to do that

- <https://chrome.google.com/webstore/detail/copy-comments/ojfnkeolkonikgdlo/kpkpcmbffnmpbif>
- Managed to clone 1st draft as an illustration
  - <https://docs.google.com/document/d/1Eypu16Q3PdF7thkkCcixOpBW9FdtMmcB2SWLdEh0Zo/edit#heading=h.wq4wngymp8nh>
- Not perfect but good enough: in particular all comments attributed to the person who do the copy with the original author in the comment text

# Second Draft Proposal

- Begin to work on the copy, processing comments in the first draft
  - Close comments in the copy when processed
  - Use the Comments button in the first draft to get the comments by date (in reverse order): allow to easily identify comments done after a certain date
- Second draft release (almost identical to what was done for the 1st draft)
  - Copy the copy made to prepare the second draft without the comments
  - Announce the new URL with the same sharing as the first draft
  - Keep the draft 1 document (at its existing URL) after disabling changes by anybody (adjust sharing permissions)
- LaTeX for final version
  - This will be some work, but would result in a more polished final document
  - Some automation is possible (GDoc -> HTML/Doc/RTF -> LaTeX using Pandoc): test it in // with the second draft work?