General comments:

- The suggestion is to have a structure as follows:
    - Executive summary
    - Introduction
    - Description of challenges and needed effort to address them
    - Recommendations
- Content. At the moment the report reads as a text book describing End User analysis software, with a glossary, but the description of the challenges and R&D/development activity, or recommendations (beyond the idea, to be developed, of supporting personnel) is still very thin.
- Time scale. Should you explicitly say that this report targets future experiments/observations/calculations in the next 10-15 years rather than what is commonly known as future facilities?
- Audience. When iterating on the text, remember you are talking to colleagues across frontiers (who are not familiar with the S&C jargon) and the funding agencies (to whom you are telling what they should support, although without requesting sums of money for specific projects)
- International program. This is not the first time the community has discussed challenges and opportunities in S&C in the next 10-15 years. In my mind, Snowmass is an opportunity for the US community to decide what's important for our programs (domestic and international) and what pieces of the overall international S&C program we want to invest on, because they are critical to our US HEP interests, we have a domain expertise advantage, or funding is available. The report should have a contextualization of the challenges discussion and recommendations within the international HEP effort. This is connected with training, hiring and retention in critical areas of expertise for example.
- EDI. S&C has a reputation for not creating a welcoming and inclusive environment, and this goes beyond the poor representation of women and minorities in our teams. We will write about this in the CompF report so we would appreciate your input in the context of your TG domain, in the form of a paragraph in the main text, and may be a recommendation if you have concrete ideas.

Detailed comments:

Executive summary – missing

5.2 People do software work

- General: a brief introduction would help me understand what the content of this subsection will be.
- General: should there be a separate subsection about this or rather comments spread across the end user analysis description of challenges and needs, and then very clear recommendations on how to support common efforts and people? If you decide to have

a separate subsection, I suggest to have it after you have explained the challenges and work needed.

-

5.2.1 Problems

- General: Problems with what? It looks like you are diving directly into the collaboration, recognition, long-term support, funding problems without discussing the end user analysis challenges first.

-

5.2.2 Personnel support case studies

- General: again, it looks like you are diving into the funding or collaborative models of "external" collaborations developing common tools, before discussing end user analysis.

5.3 Analysis Ecosystems: Libraries, Languages, and Data Formats

- General: very interesting topic! I like your introduction to software ecosystems. Certainly, ROOT and Python are THE two major ecosystems in HEP relevant to End User Analysis (granted – relevant to many other things too), but you present them as if they were THE two in HEP; I am not sure this true.
- Line 31. I guess we can agree that ROOT is hosted by CERN, but we need to be careful about ignoring that US makes significant contributions, in particular to the area of I/O.
- Line 41. What I am hearing from colleagues, is that ROOT is spreading too thin in trying to cover two main objectives: I/O – event loop – data format – storage on one hand, and end user analysis on the other (histo-statistics-visualization-plotting). "Modern thinking", although I am aware there are different opinions, favor a ROOT evolution that enables the use of the very rich Python ecosystem for the latter objective. Are you going to say something about this? The US may continue contributing to ROOT I/O, but there is significant work these days in the context of IRIS-HEP (and also DOE research funding) to develop tools to take advantage of column-wise data formatting and Python tools (ServiceX, Coffee, etc).

5.3.1 Programming Languages

- General: very nice introduction to languages. I want to see how you connect everything with the description of the End User Analysis challenges before commenting in detail or advising whether this is the right level of detail.

5.3.2 Data Formats

- General: Good introduction to data formats. Will you discuss RNTuple , column-wise data formatting for faster, easier access, and connect with the challenges of analyzing huge data samples at the end user analysis level for timely production of conference

results and physics papers? The idea of these new formats, interfaces (evolution of TTree) and analysis frameworks (such as Coffea) is to deliver a reasonable turnaround and user friendliness so that running and re-running on analysis level ntuples does not become an unsurmountable bottleneck. I am sure there is information out there from where to extract quantitative statements about this challenge.

5.3.3 Visualization

- General: Will you discuss how ROOT and the Python ecosystem may complement each other and the work being done, planned within the US, as well as the challenges ahead (and eventually recommendations)?

5.3.4 Requirements for a sustainable end-user software ecosystem

- General: It looks like here you are starting to tackle what I was mentioning before. However, you go directly to support of personnel, documentation and training. Good comment about interoperability!

5.4 Analysis Models

5.4.1 Scale

- General: good introduction and comment about the importance of fast turnaround. In fact, this is one of the challenges to address.
- General: maybe you can elaborate in a more quantitative way here what the high-throughput challenges are in the context of the large datasets to be analyzed in the next 10-15 years, describe what is being or needs to be done, what technologies, solutions are being prototyped, etc.

5.4.2 Interfaces

- Same comment as for Scale but in the topic of interfaces

5.5 Dataset bookkeeping and formats

- Too early for me to comment

5.6 Collaborative software

- This is still a core dump of tools and references. Are you talking to the IRIS-HEP, HSF and the Coffea people? By the way, their documents should be consulted and referenced somewhere.

5.7 Training

- General: I think this is going in the right direction. I mean a description of the challenges in recruiting and retaining people with the needed (scarce and expensive) expertise, followed by ideas and recommendations on how to train and retain.

5.8 Thoughts

- These needs to be the “Recommendations” subsection

Bibliography

- General: needs to be much longer