ECFA Early Career Researchers Meeting

November 15th 2019

Working Group Report
Computing & Software

Erica Brondolin, Alison Elliot, Julián García Pardiñas, Katja Mankinen, Abhijit Mathad, Josh McFayden, **Jonas Rembser**, Sezen Sekmen, Pawel Sznajder

The Working Group

On November 11th, 16 of us met to discuss **Computing & Software**.

Very constructive discussion on large variety of topics, minutes in Google Doc.

Results of the discussion is compiled into these slides:

- Impression of the **briefing book**
- **Subjects** we think are important to be considered **for** the **strategy input**

Physics Briefing Book

We agreed with the **key points** the briefing book made about computing:

- **HL-LHC** poses immediate **computing challenges** which need to addressed
  - Strong focus on R&D in an environment that fosters innovation

- Impact on computing should be increasingly considered in detector design
  - Need for individuals to **link instrumentation and computing**

- Computing can benefit from **networks** that can also extend outside HEP
  - These structures need to be **strengthened**

Briefing book talks about "computing in HEP" as mainly the **central production processes** (simulation, reconstruction, data management...):

- We feel software and computing goes also closely with **physics analysis**

Physics Analysis Software

- Importance of not being repetitive and not **reinventing the wheel**
  - Leverage on software trends to **automate analyses**
  - encourage analysis description standards for analysis preservation
- allow researchers to work more **effectively** and move on **quickly**
- **Economize** time usage and allow more **time and personpower** to be directed to places where innovation and development is truly needed
- Collaborations should encourage **centralization of analysis tasks**
  - Too much duplicated work with plethora of analysis frameworks
- Build software / frameworks that are easy to understand and can be widely used, and can be used for multiple purposes.
  - Emphasize importance of using standard technology to **ensure transferability of software skills** to other areas

Encourage Expertise in Software

We feel it's important that **software** in HEP at for production and analysis level **should be contributed to by experts** to ensure most efficient use of resources and minimize time to physics.

- **Physicists** should be encouraged to become **experts** with **trainings** or by **self-learning**
- Other option are **computer scientists** that are paid and recognized maybe even outside the particle physics

Open Data

Open data is the idea that some data should be freely available to everyone to use and re-publish as they wish, without restrictions.

- Possibility to analyze outside the collaborations can bring **innovative ideas**
- Great tool for working with other fields (e.g. **machine learning**) and for **education** and **outreach**
- Can also be used by people within collaborations to **publish independently** about **new algorithms or tools**, **increasing visibility** of such developments

Possible counterarguments:

- **Non-experts** might use data to make **inaccurate scientific claims**
- Needs to ensure proper **recognition** of the work from the people that collected the data
- Risk that **development work within** experimental **collaborations** becomes **less attractive** and innovative as innovation would now happen from open data

Analysis Documentation and Preservation

Not only **Data preservation** and **open access** currently lack consistent policies in the experiments, but also **analysis preservation** and **reproducibility**.

- Experimental data and results should be available in a form that can easily reinterpreted
- The usage of **machine learning** poses new **challenges** to **reproducibility**

We acknowledged the importance of **clear documentation** of analysis software developments.

- Becomes increasingly important as **new tools** arise and start to be used by people
- Collaborations could make well documented analysis software a **requirement for** their **papers**
- Clear **documentation standards** inspired from **industry** can be adopted

Innovation in High Energy Physics Computing

- We strongly encourage the **support** of **software R&D projects** like DIANA-HEP or IRIS-HEP, as we are already benefiting from their work (e.g. uproot)
- Important to encourage **knowledge sharing** with industry and other computing communities
- We should not only encourage **collaboration with the industry** but also with **other fields** of **physics** or **science in general**
- Raise awareness in outside communities of the software and computing landscape in HEP, to **detect** possible **synergies**
- More efforts should be undertaken to **leverage already existing solutions** and learn from other approaches to similar problems
- There should be wider **awareness** in the HEP community **of the methods** used by **other fields** in physics
    - Astrophysics could be a good model to follow in particular

Ecological Computing Resources

- **Encourage the use of new technologies** not only to address immediate challenges, but also to reduce the ecological footprint of computing in HEP
  - For example GPUs and FPGAs
- Consider sustainability, environmental impact and **relation to local community** when establishing **computing facilities**
  - **Sharing** of resources
- However, this has to be **put in perspective**:
  - Computing only makes up a rather small fraction of the power consumption in HEP
  - more details in Environment WG report

Software for Remote Meetings

In an effort towards **reducing the amount of travelling**, the **software for remote meetings** should be improved and specialised for large-community meetings. Improving the remote experience can have a big impact.

- Get **inspiration from outside**: companies might have interesting solutions
- Functionalities to **remotely draft ideas** together on a **whiteboard**
- Better support for **large conferences**:
  - Focus on substituting the "**coffee break experience**"
    - Chatting platform
    - Could go as far as "chatroulette" to foster new encounters and exchanges between collaborators

Input on Human and Social Factors

Training

- The rise of heterogeneous computing requires **adequate training of the researchers** (both early and later career).
- This is of **utmost importance for software developers**, that should be either **highly-trained physicists** or **computer scientists** (possibly even from outside HEP).

Work recognition

- The computing/software activities should not be seen only as a means to get analyses but as a **proper research area**.
- An **increased number of devoted awards** inside the collaborations can highly reduce the invisibility of the contributions to software.
- **Encouraging more publications** on software/computing work helps to increase the recognition outside the collaborations.
- The **recognition of software** should be at a **comparable level** to the **recognition of hardware**.
- **Job descriptions** should explicitly include **software skills**, to increase the awareness on them.
- Increasing the number of **specific career opportunities** would have a big effect.

Career possibilities outside HEP

- **Researcher characteristics that are positively perceived** from the outside: ability to collaborate in large software infrastructures, experience with big data, problem-solving skills.
- Importance of larger and more solid **Alumni networks**.
- **Return from people** that have left the field and gained extra experience outside can be very beneficial, particularly for training.