



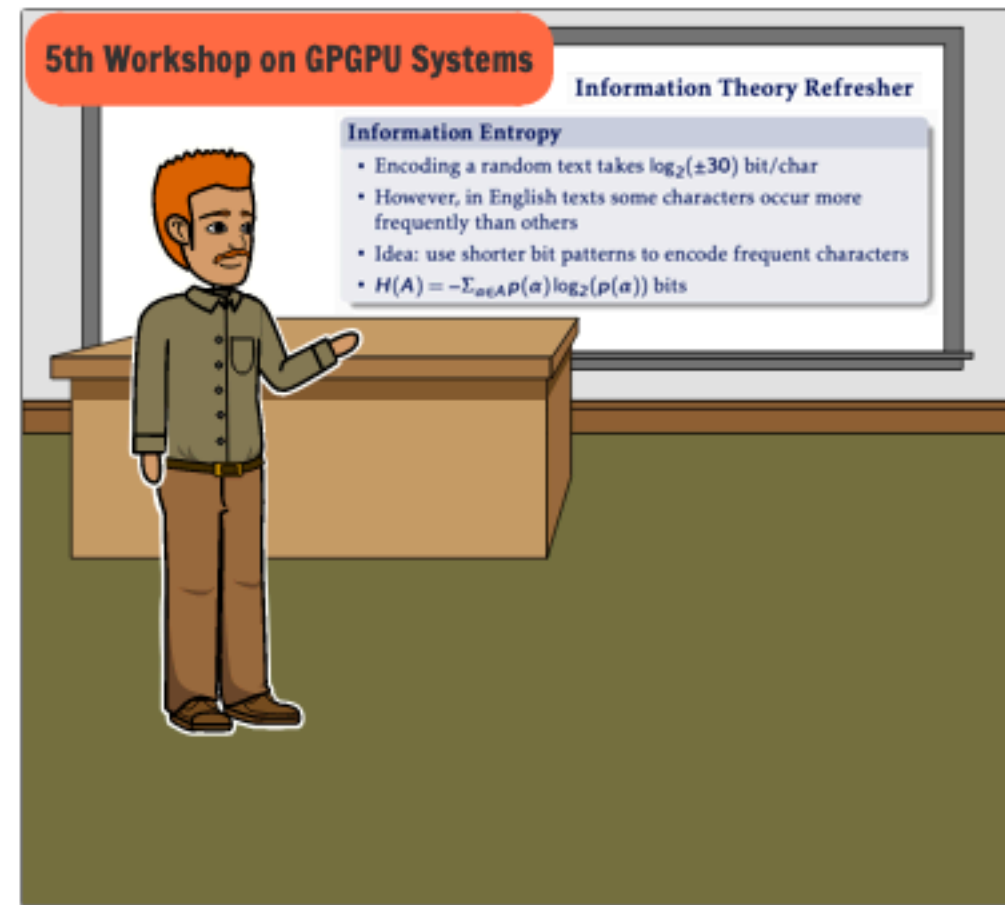
SIDIS
Software Institute for
Data-Intensive Sciences

THE SIDIS INITIATIVE

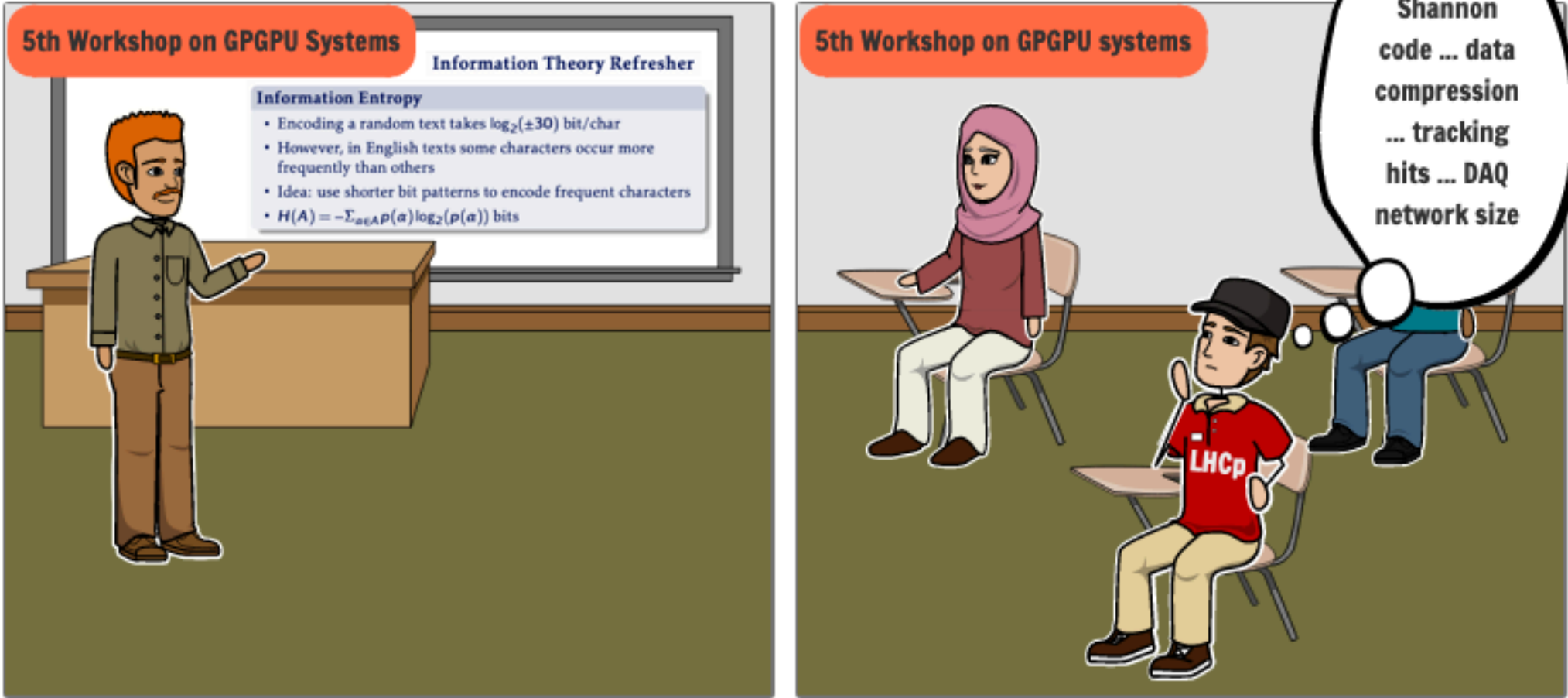
STEFAN ROISER, CERN

SWIFTHEP/EXCALIBURHEP WORKSHOP, 3 NOV 2021

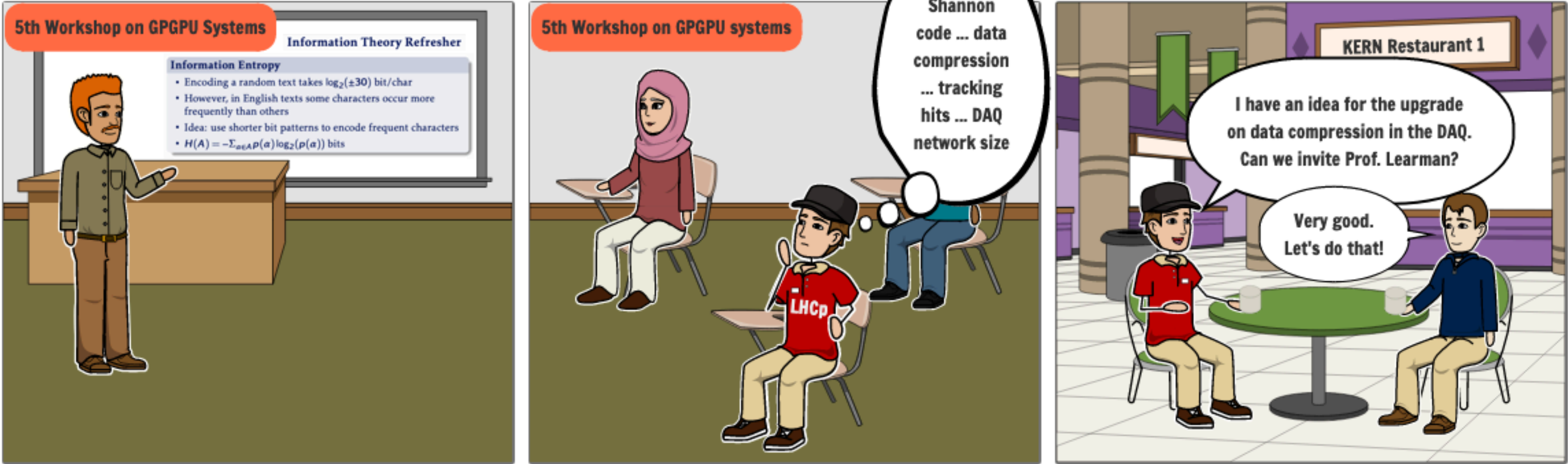
AN ALMOST TRUE STORY ABOUT THE LHCP EXPERIMENT AT KERN



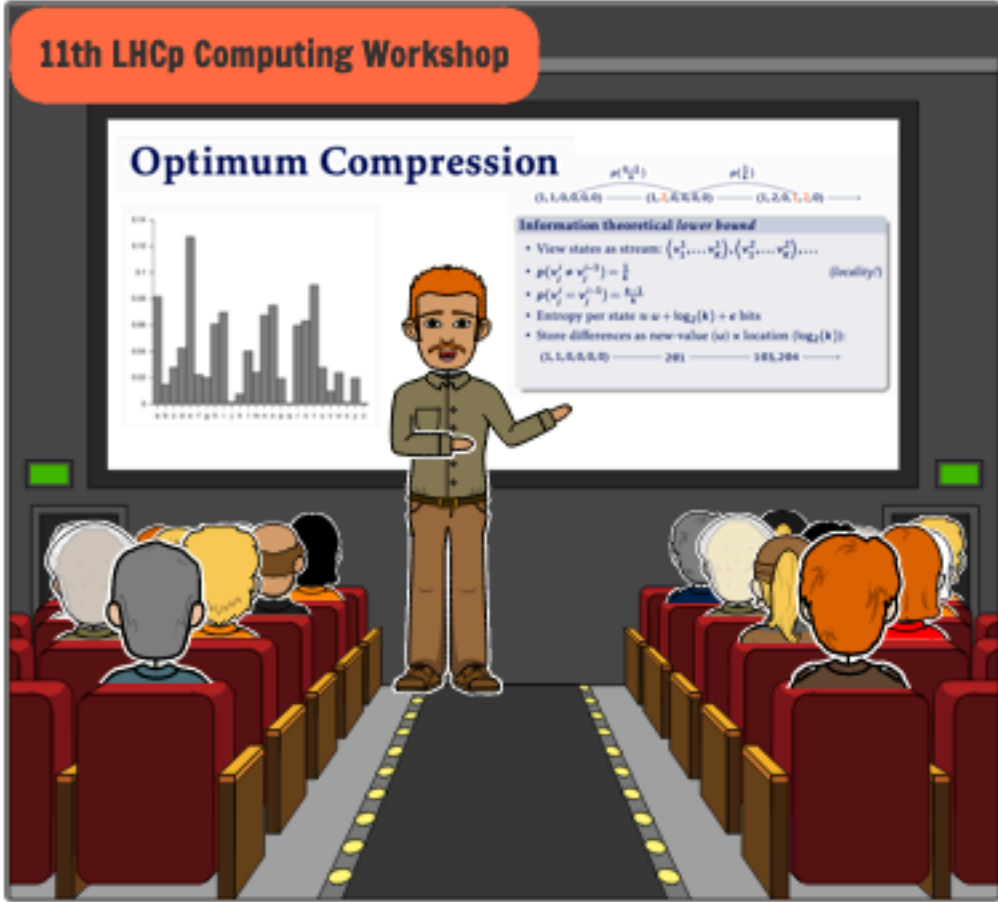
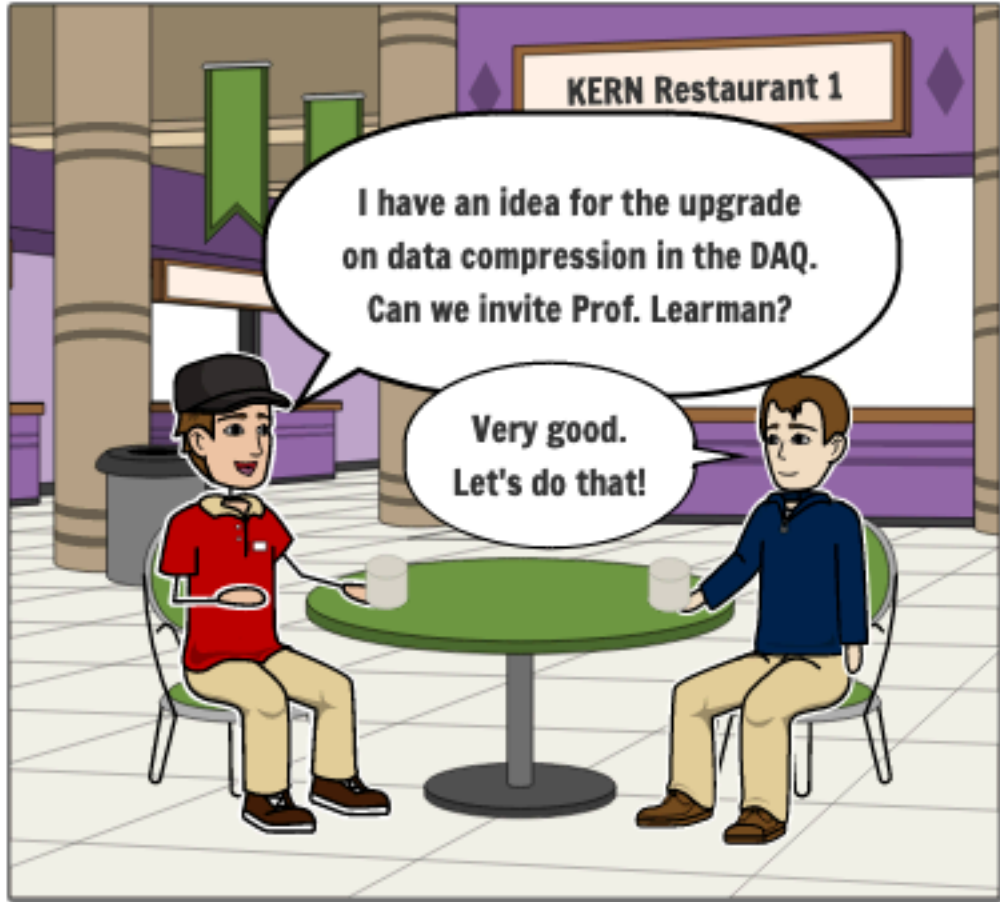
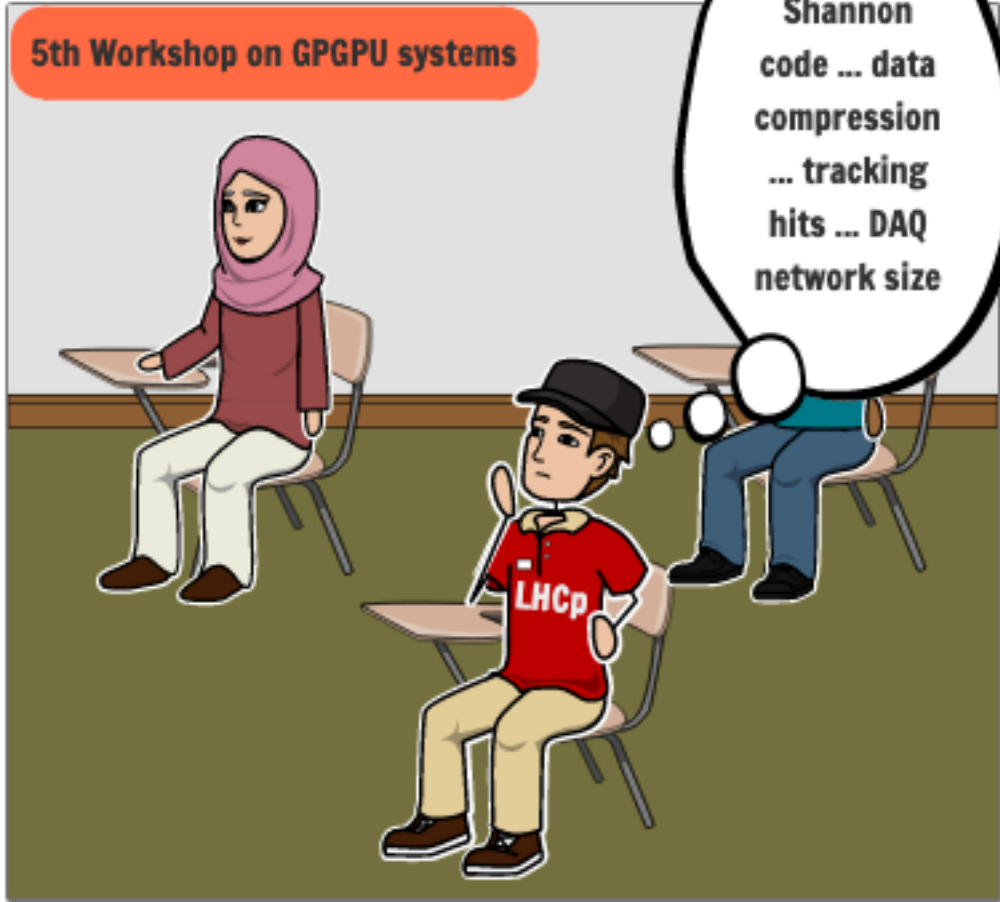
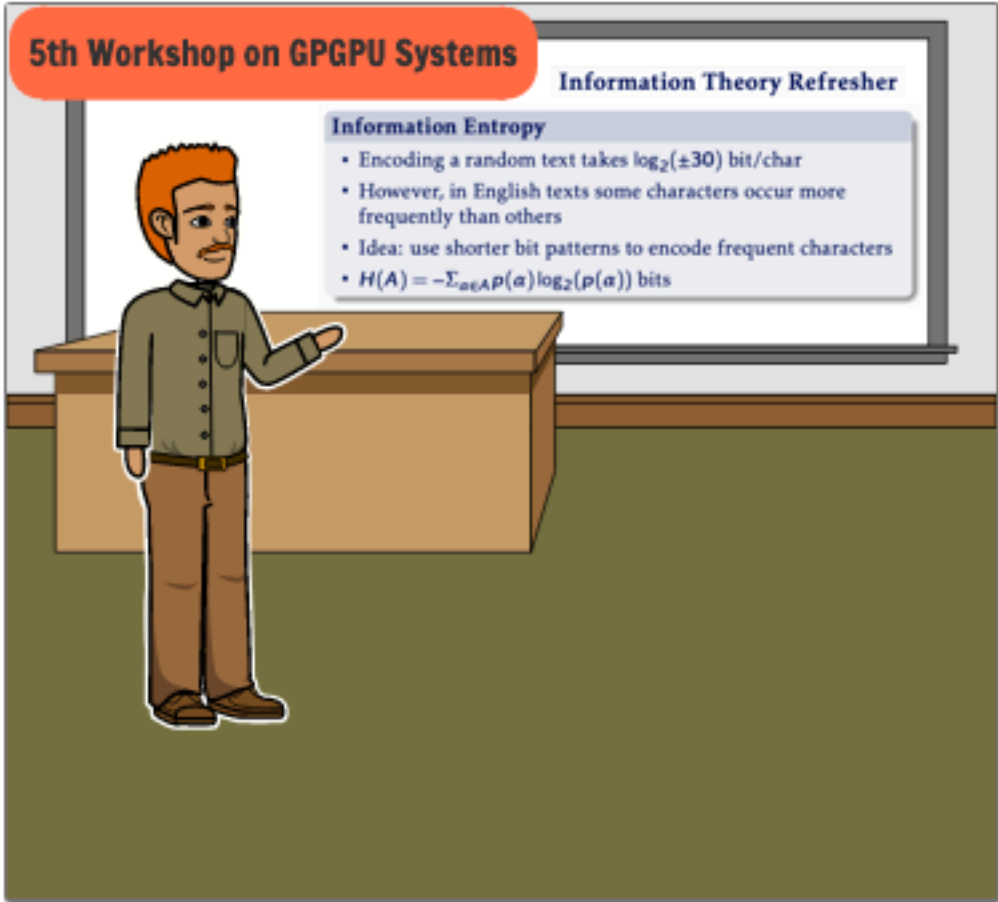
AN ALMOST TRUE STORY ABOUT THE LHCP EXPERIMENT AT KERN



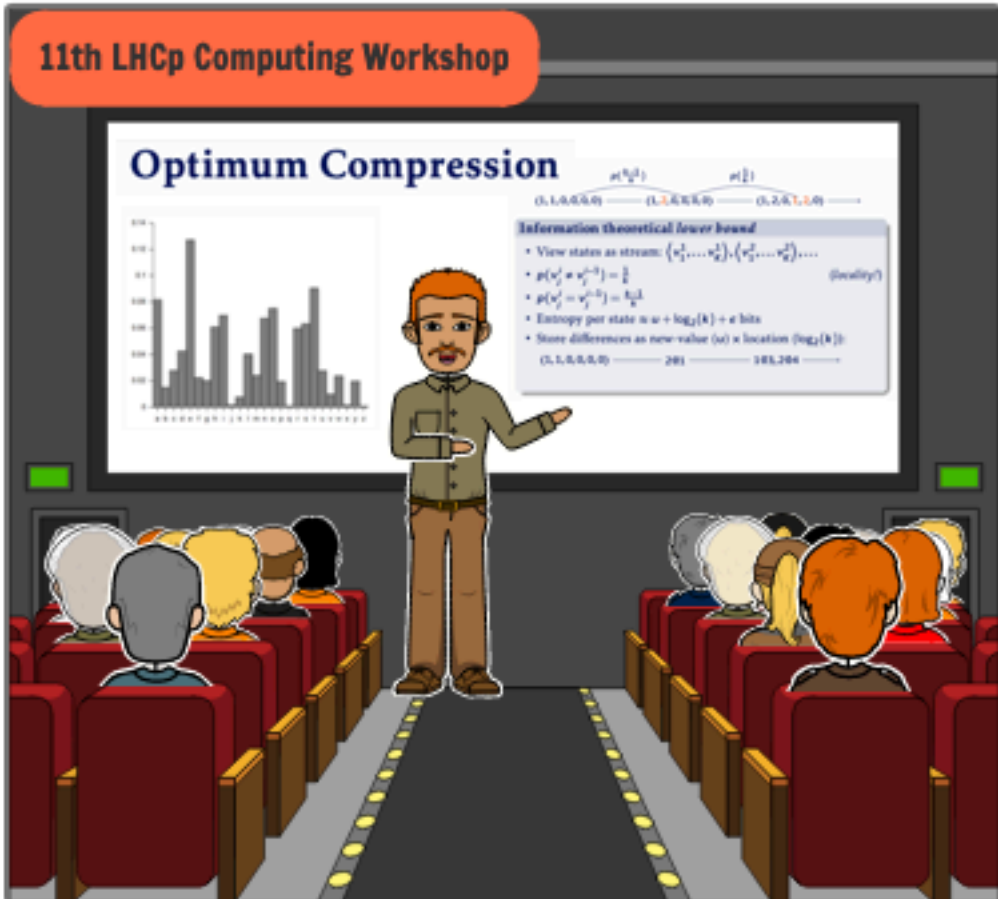
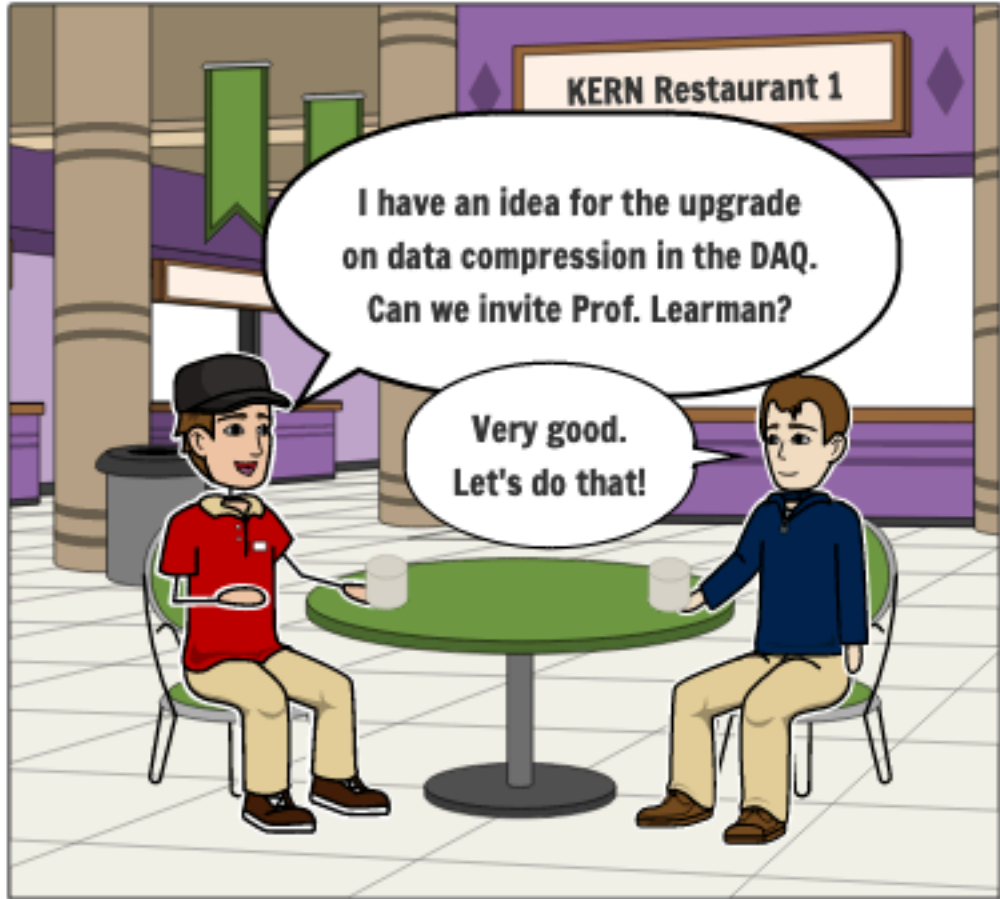
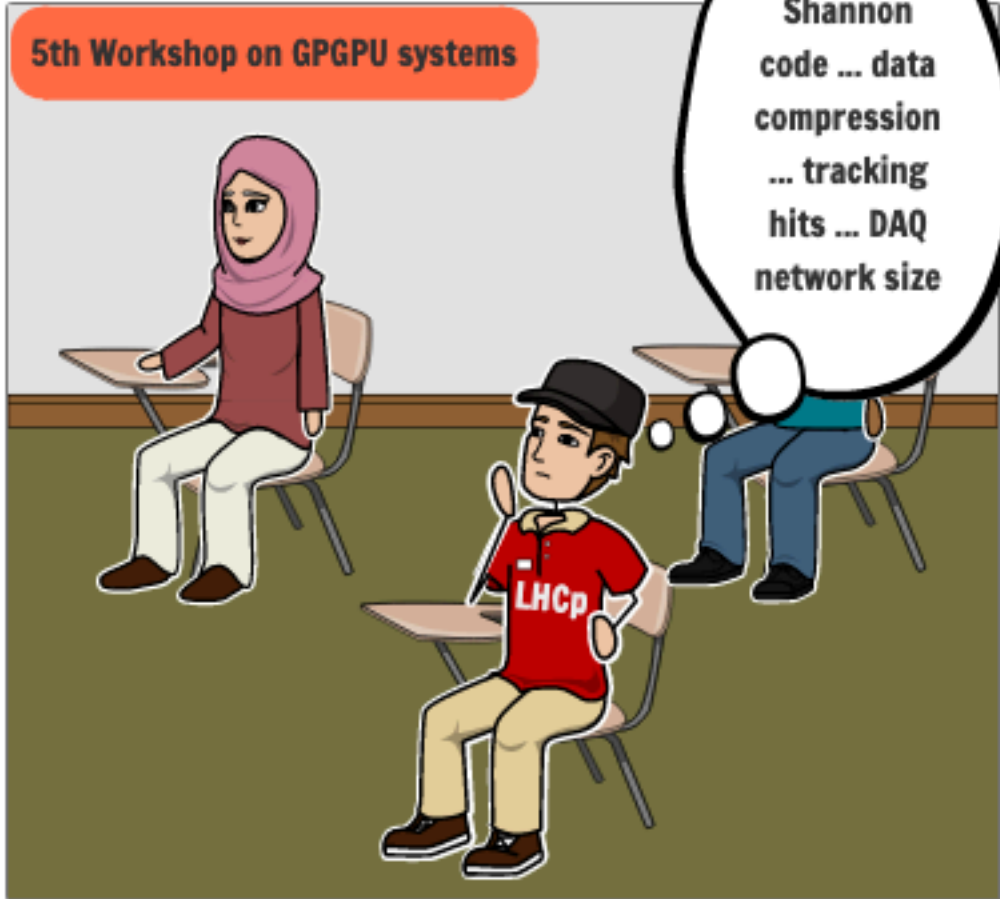
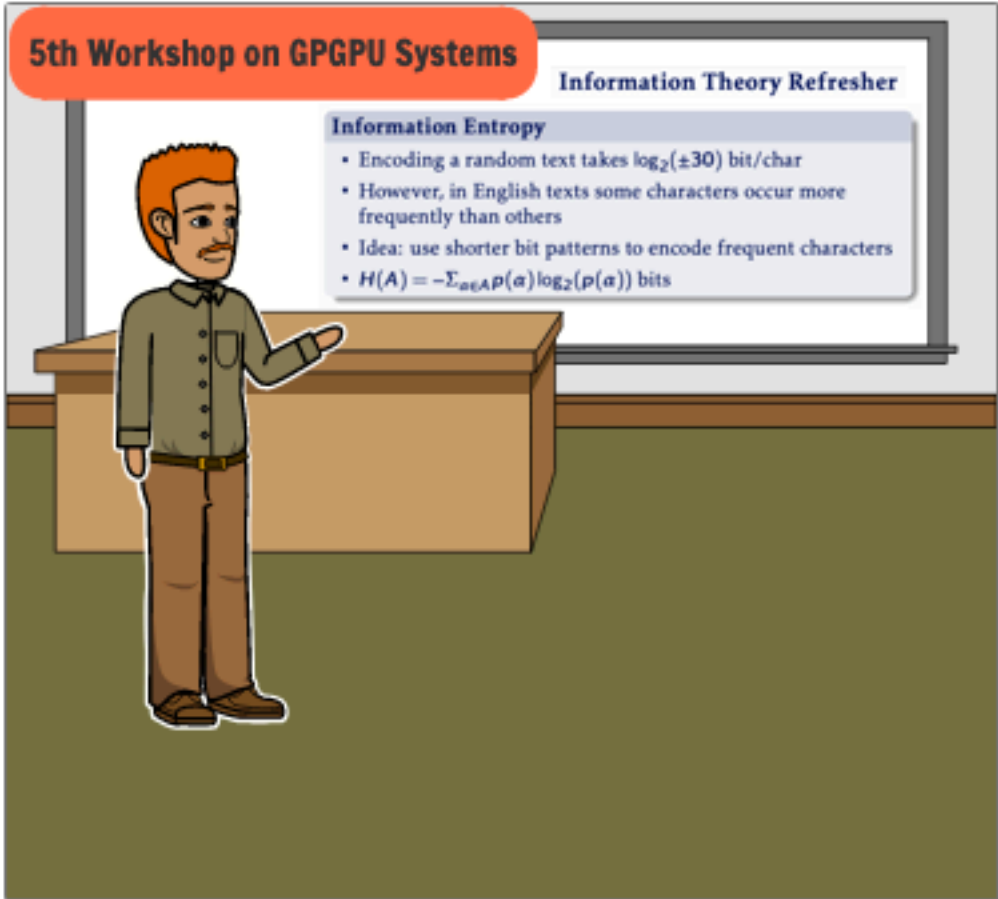
AN ALMOST TRUE STORY ABOUT THE LHCP EXPERIMENT AT KERN



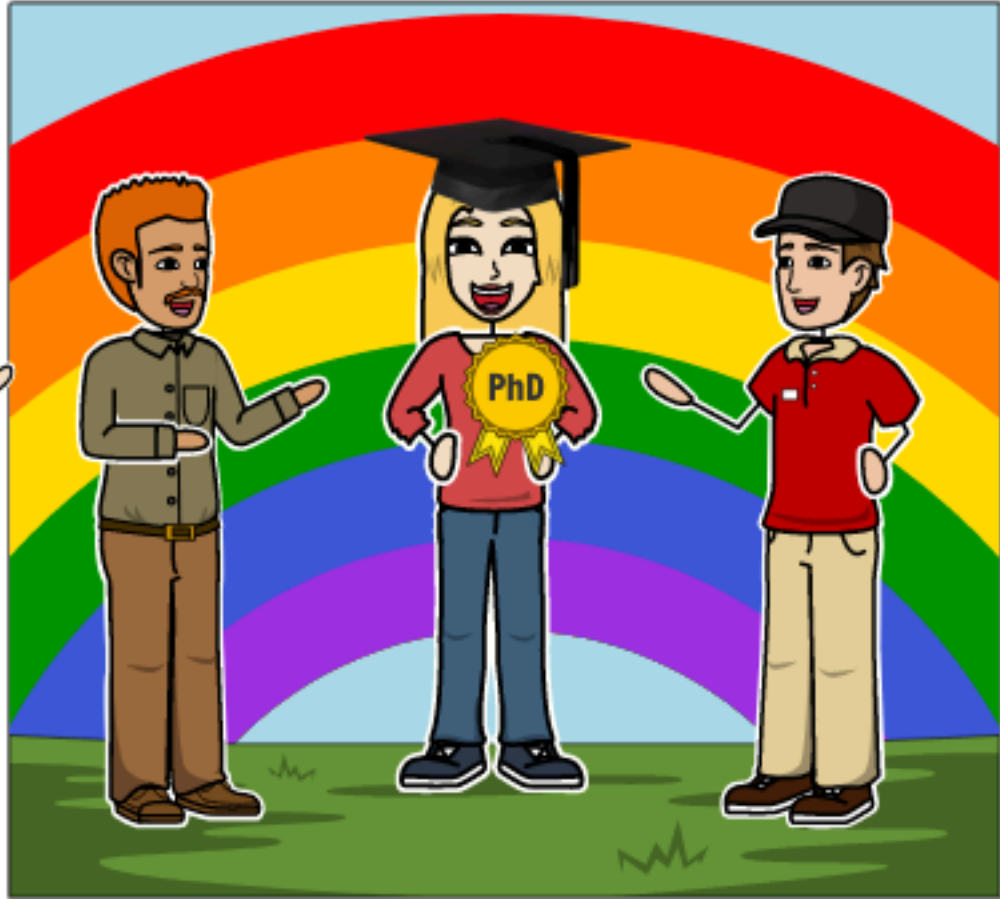
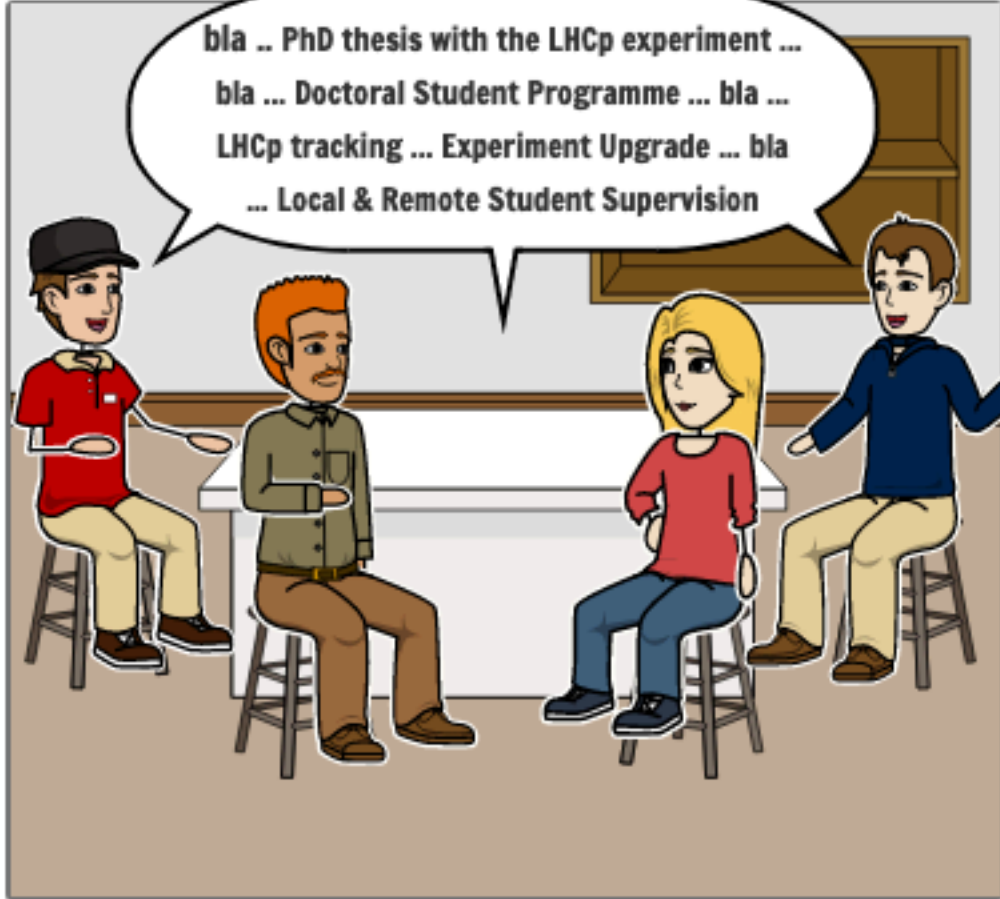
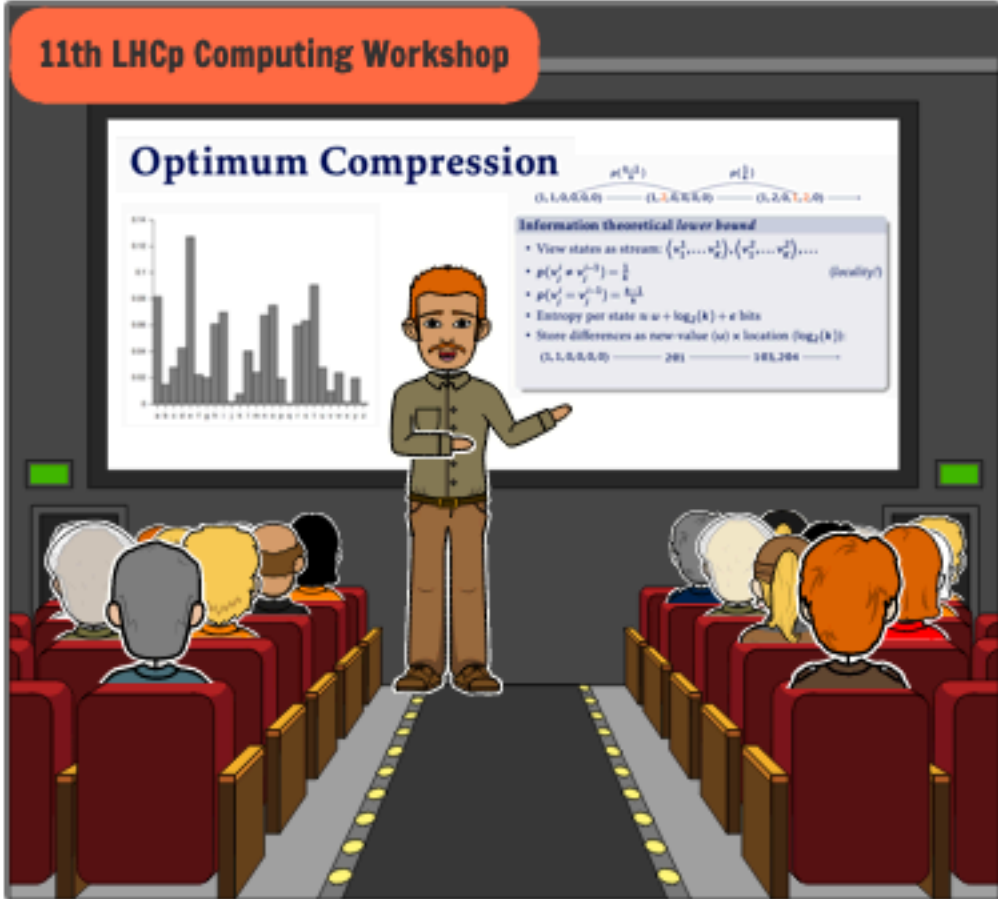
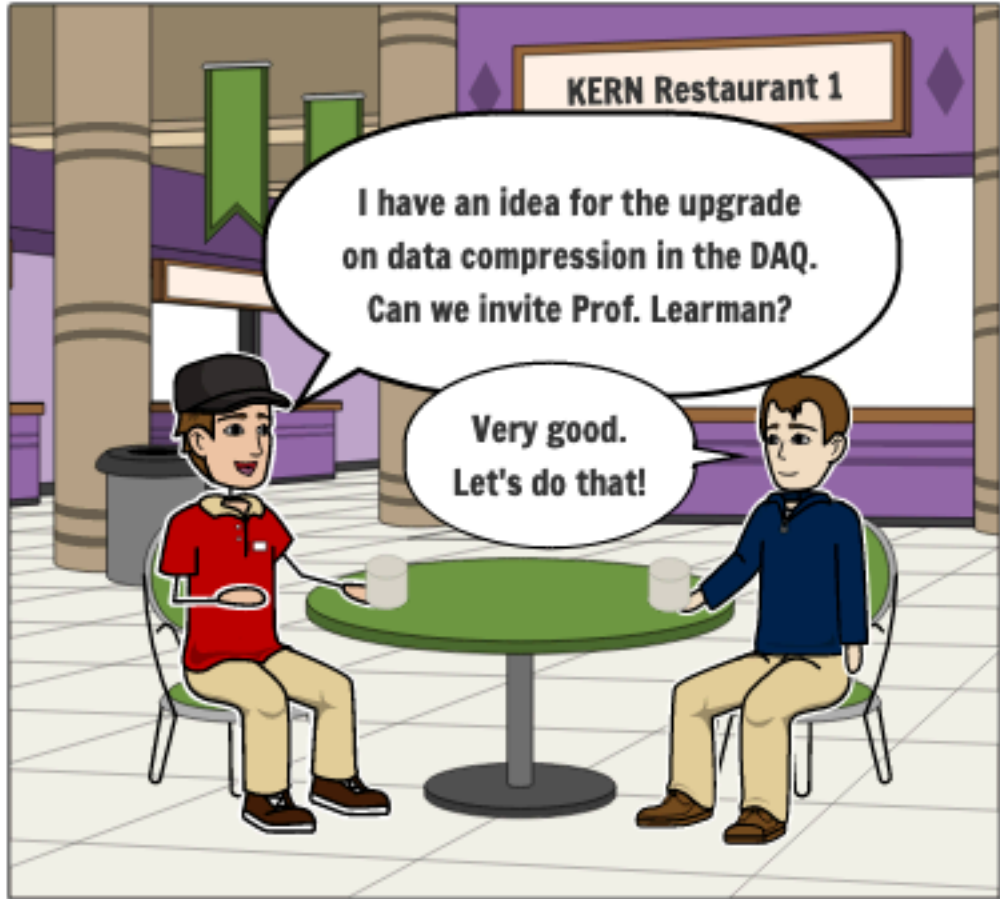
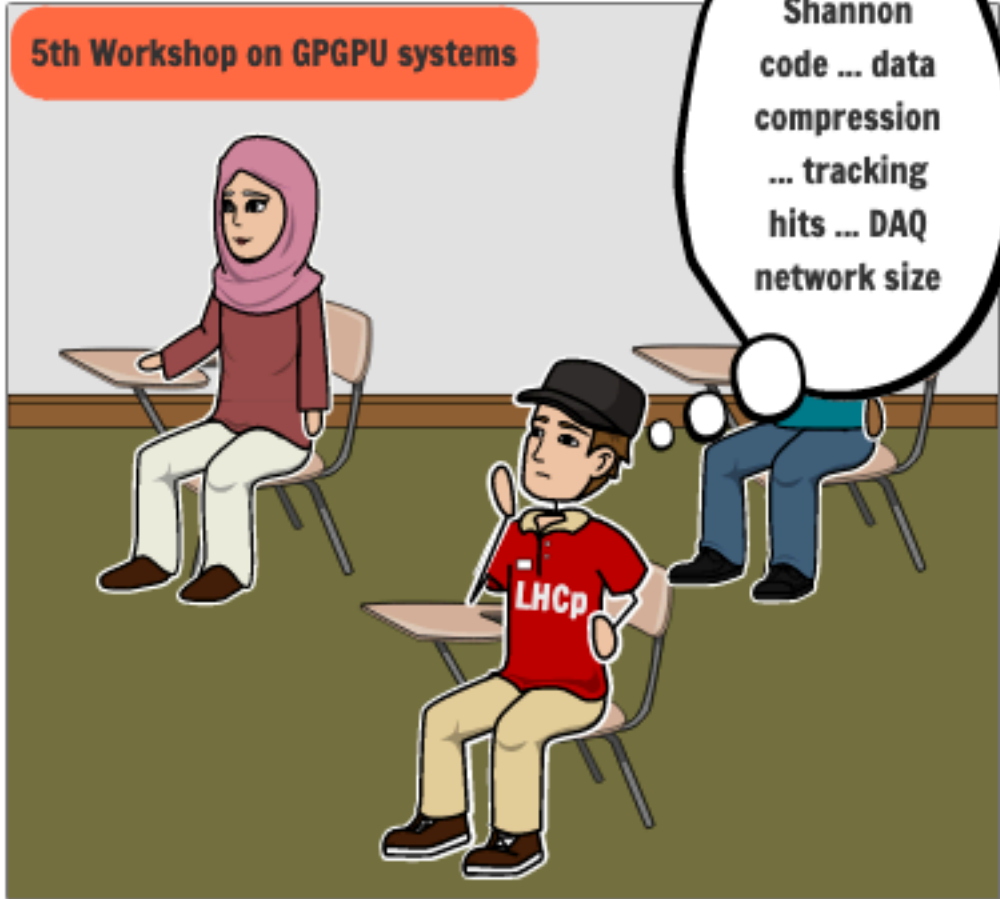
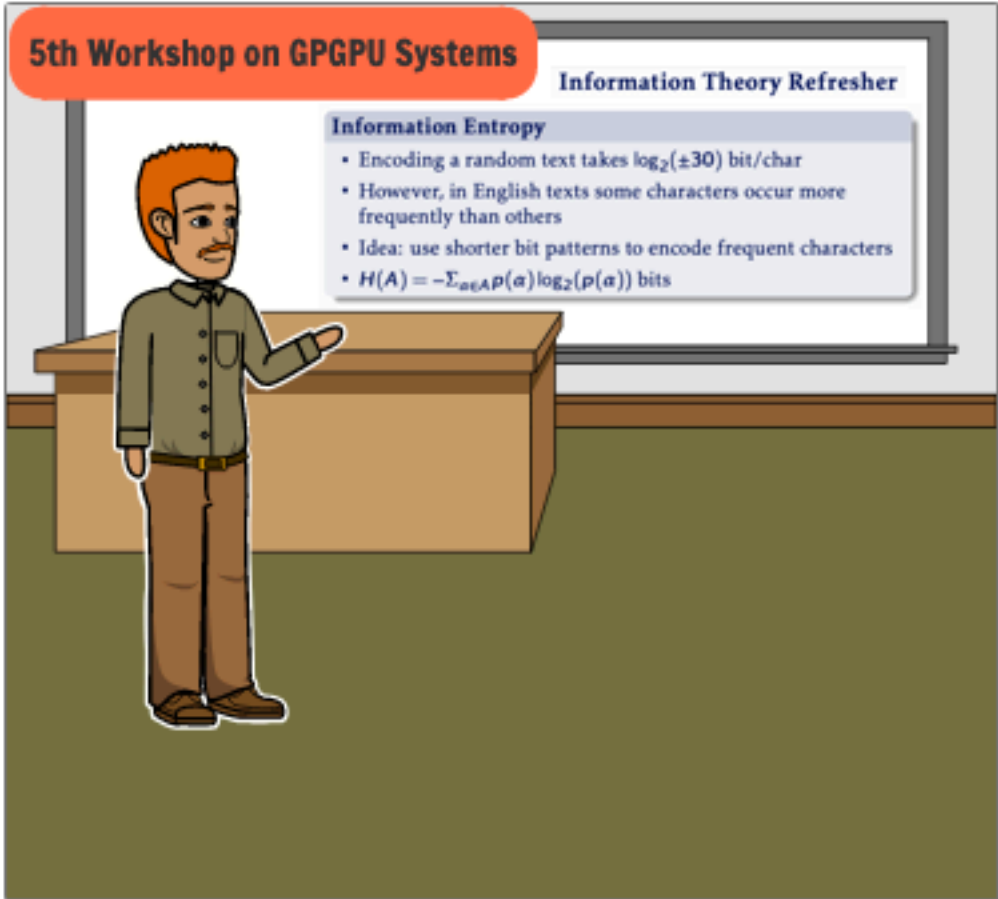
AN ALMOST TRUE STORY ABOUT THE LHCP EXPERIMENT AT KERN



AN ALMOST TRUE STORY ABOUT THE LHCP EXPERIMENT AT KERN



AN ALMOST TRUE STORY ABOUT THE LHCP EXPERIMENT AT KERN



SIDIS CONCEPT



SIDIS CONCEPT

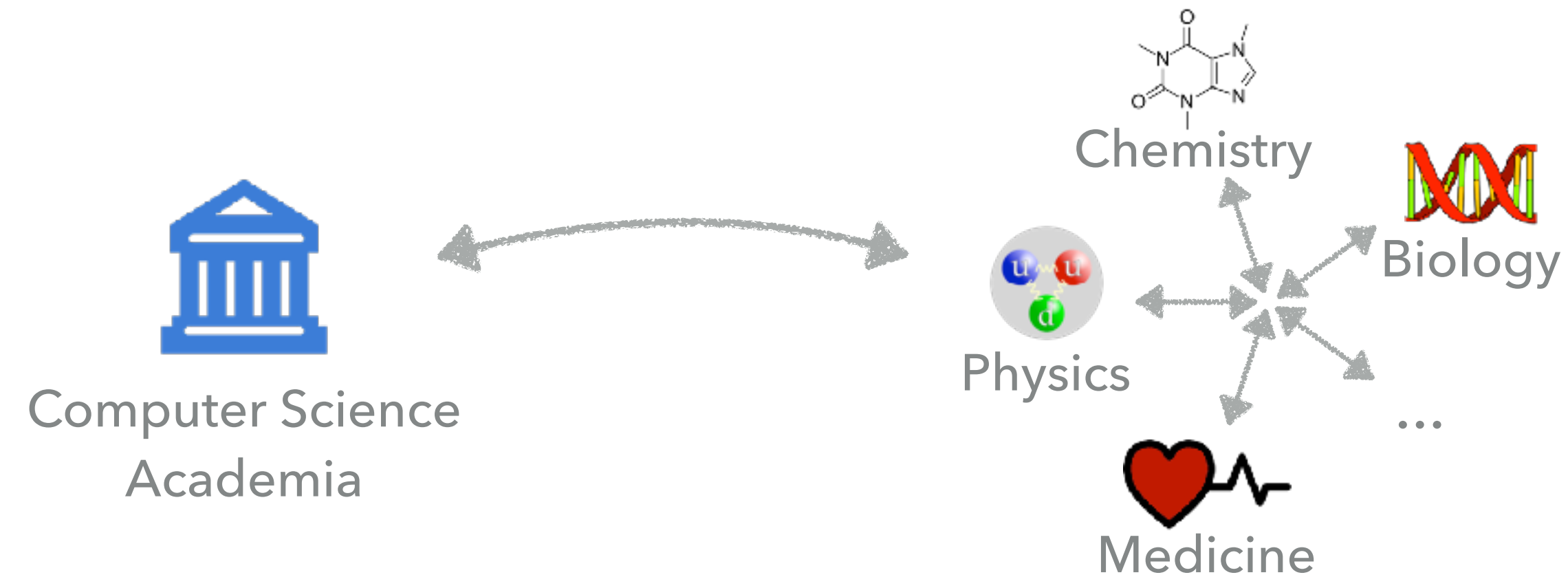
The “Software Institute for Data Intensive Sciences” aims to



- ▶ establish an environment for interactions between computer science and high energy physics software engineering
- ▶ enable training for high energy physics early career researchers
- ▶ lobby for software engineering in high energy physics

SIDIS CONCEPT

The “Software Institute for Data Intensive Sciences” aims to



- ▶ establish an environment for interactions between computer science and **data intensive sciences** software engineering
- ▶ enable training for **data intensive sciences** early career researchers
- ▶ lobby for software engineering in **data intensive sciences**

WHO IS INVOLVED

- ▶ CERN core team in IT-SC (Markus Schulz, SR) and EP-SFT (Graeme Stewart)
- ▶ Coordination together with members in several European Universities & Labs
- ▶ Strong collaborations with the HEP Software Foundation and CERN Openlab



The first trial to kickoff the activity unfortunately had to be postponed

CURRENT SIDIS ACTIVITIES

- ▶ Software Engineering Training
- ▶ Compute Accelerator Forum
- ▶ EU COST Action Funding

TRAINING: ALPAKA ABSTRACTION LAYER



▶ Alpaka is an abstraction layer for computing accelerator development

▶ Training held 29 June – 3 July 2020
<https://indico.cern.ch/event/912156/>

▶ ~ 30 participants

▶ [Recorded lessons](#) available on Youtube

Lesson 01: Introduction to alpaka



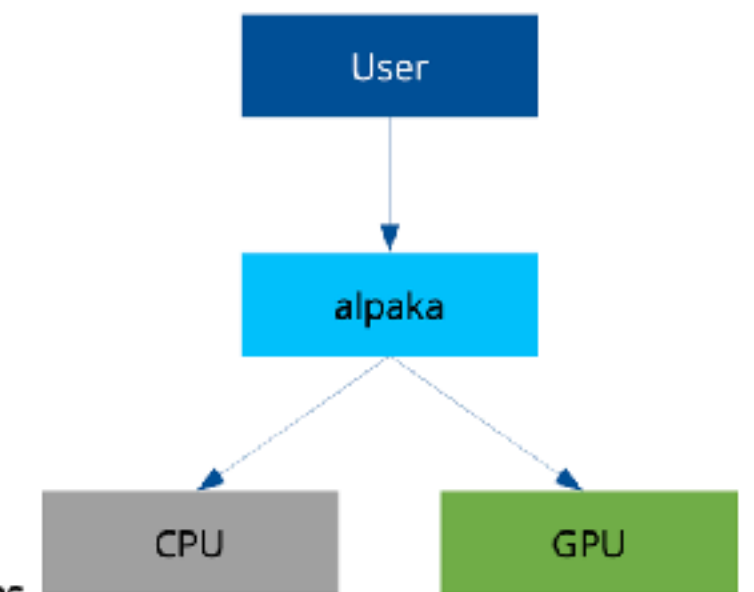
alpaka's purpose

Without alpaka

- Multiple hardware types commonly used (CPUs, GPUs, ...)
- Increasingly heterogeneous hardware configurations available
- Platforms not inter-operable → parallel programs not easily portable

alpaka: one API to rule them all

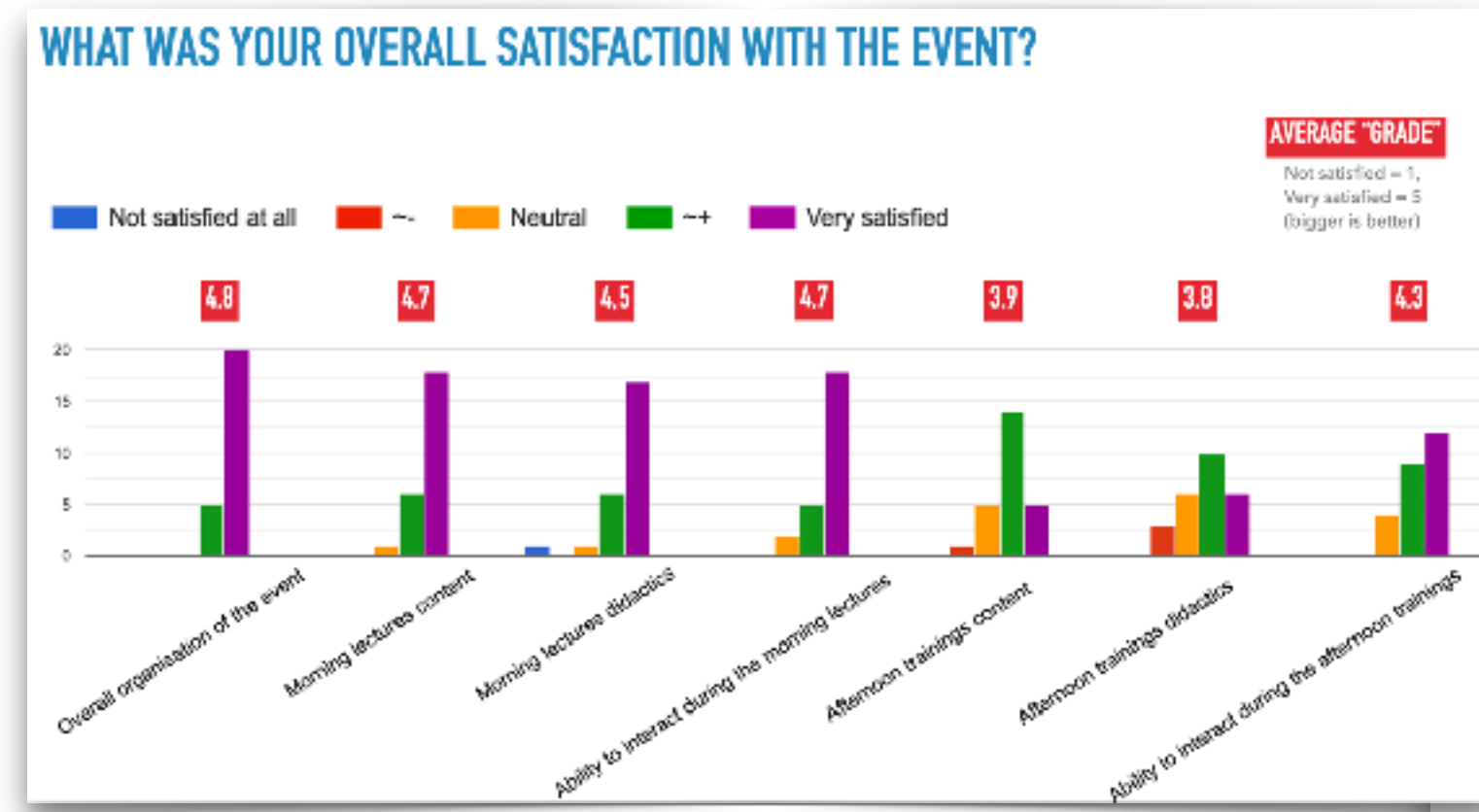
- Abstraction (not hiding!) of the underlying hardware & software platforms
- Code needs only minor adjustments to support different accelerators



TRAINING: C++

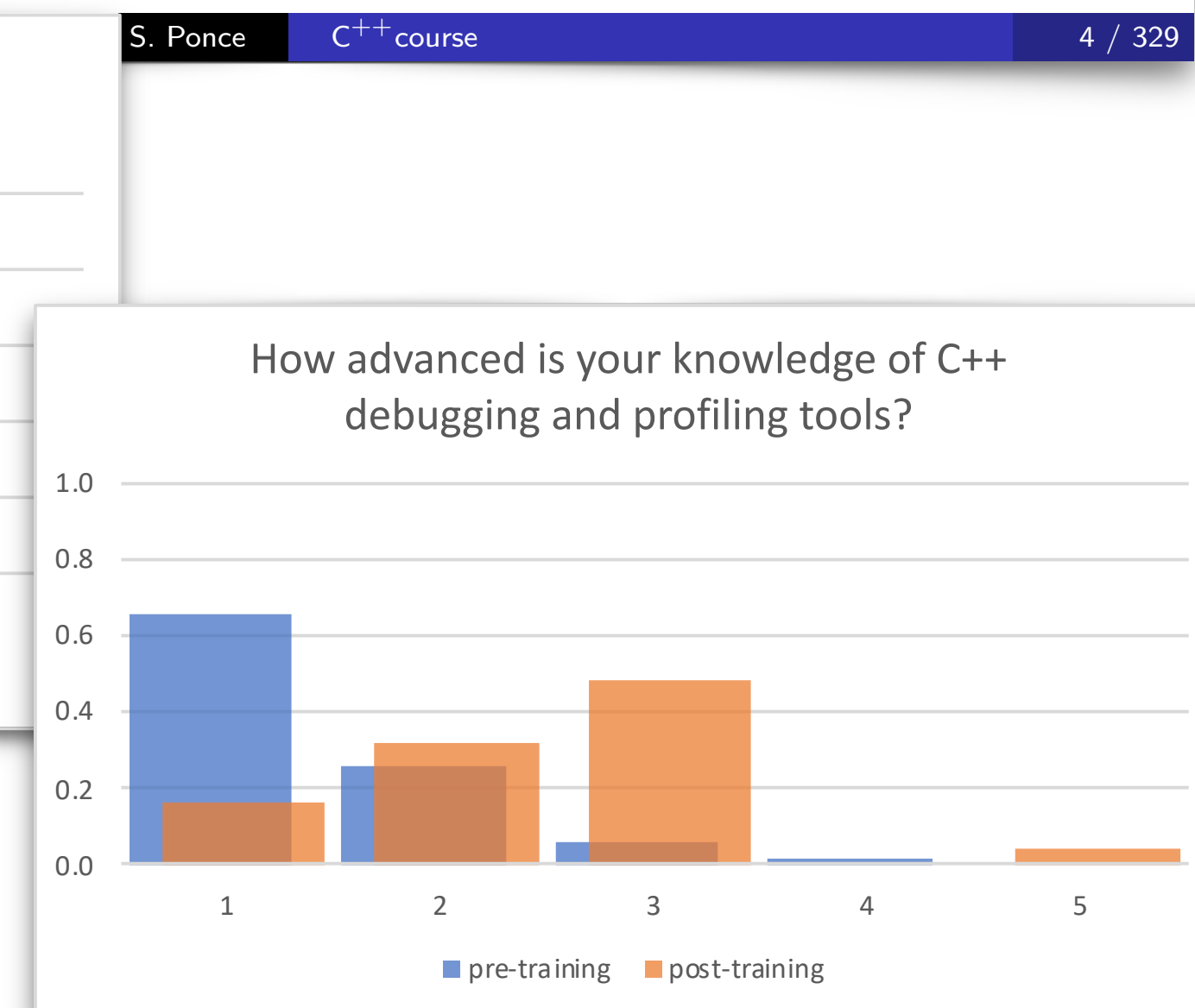
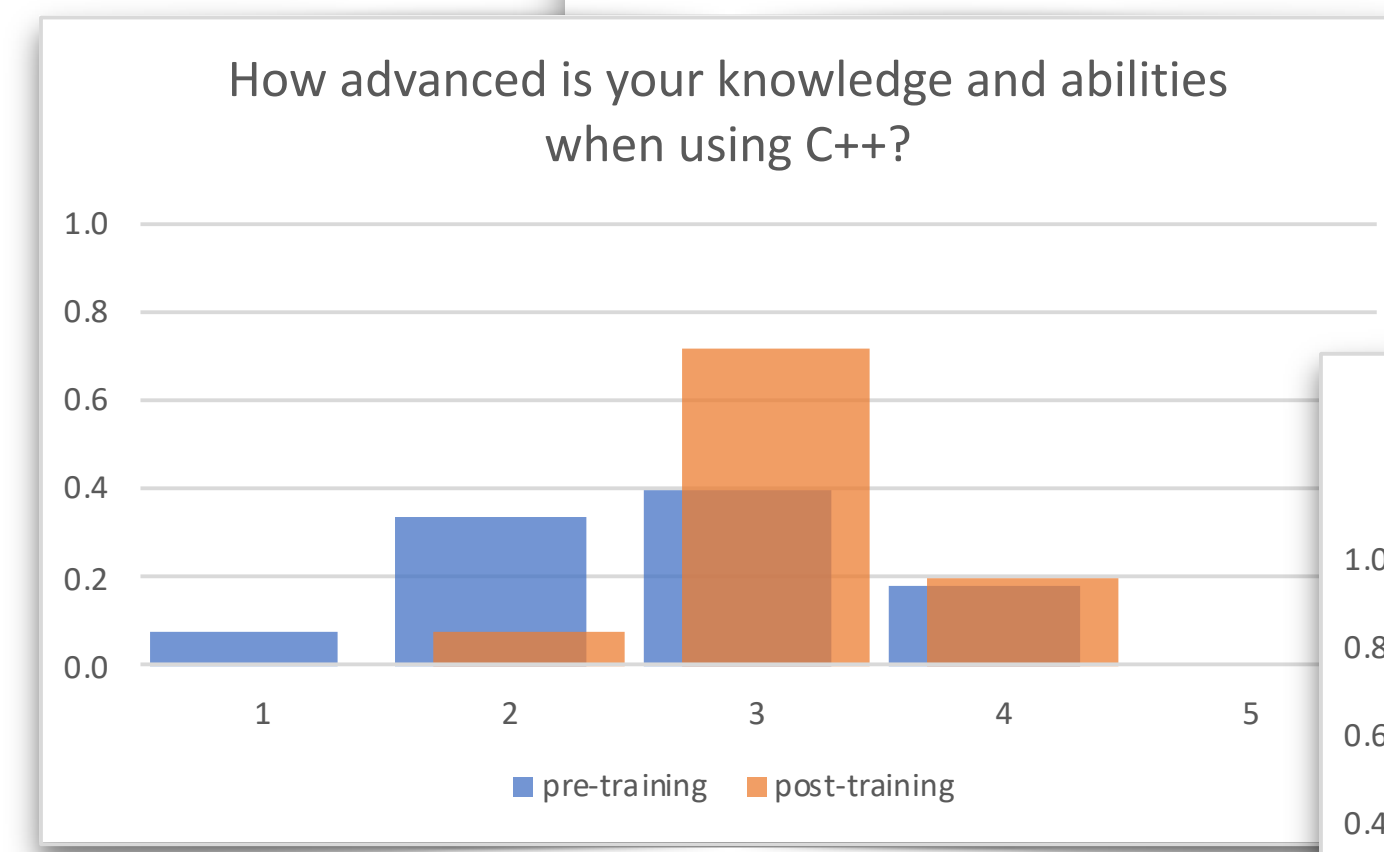
200 STUDENTS TRAINED IN 3 COURSES

- ▶ Started in Oct 2020
- ▶ 5 days morning lectures, 4 days afternoon hands-on trainings
- ▶ Based on & mainly taught by Sebastien Ponce (LHCb)
- ▶ Start from basic concepts to very advanced topics
- ▶ Aim to provide a generically applicable & good foundation for C++ programming
- ▶ Great example for a community wide effort
- ▶ 25 mentors tutoring in afternoon trainings, affiliated to 11 universities and labs (UK / Stewart Martin-Haugh)



Detailed outline

- History and goals**
 - History
 - Why we use it ?
- Language basics**
 - Core syntax and types
 - Arrays and Pointers
 - Compound data types
 - References
 - Functions
 - Operators
 - Scopes / namespaces
 - Control instructions
 - Headers and interfaces
 - Auto keyword
- Object orientation (OO)**
 - Objects and Classes
 - Inheritance
- Expert C++**
 - Constructors/destructors
 - Static members
 - Allocating objects
 - Advanced OO
 - Operators
 - Functors
 - Name Lookups
 - Core modern C++
 - Constness
 - Constant Expressions
 - Exceptions
 - Move semantics
 - Copy elision
 - Templates
 - The STL
 - More STL
 - Lambdas
 - pointers and RAII
 - Expert C++
- Variadic templates**
 - Perfect forwarding
 - SFINAE
- Useful tools**
 - C++ editor
 - Code management
 - The Compiling Chain
 - Debugging
 - Address Sanitizer
 - The Valgrind family
 - Static code analysis
- Concurrency**
 - Threads and async
 - Mutexes
- C++ and python**
 - Writing a module
 - Marrying C++ and C
 - The ctypes module
 - The cppy project



1st HEP C++

Course and Hands-on Training

<https://indico.cern.ch/e/cppfall20>
held from 12 to 16 October 2020

FOR INFO ON UPCOMING COURSE SUBSCRIBE TO "WAITING LIST" AT
[HTTPS://INDICO.CERN.CH/EVENT/1019089/REGISTRATIONS/71169/](https://indico.cern.ch/event/1019089/registrations/71169/)



1 Lecturer (S Ponce), 14 mentors, 20.5 hours, 50 students, 180 people

The event was organised and supported by faculty staff at



2nd HEP C++ Course and Hands-on Training

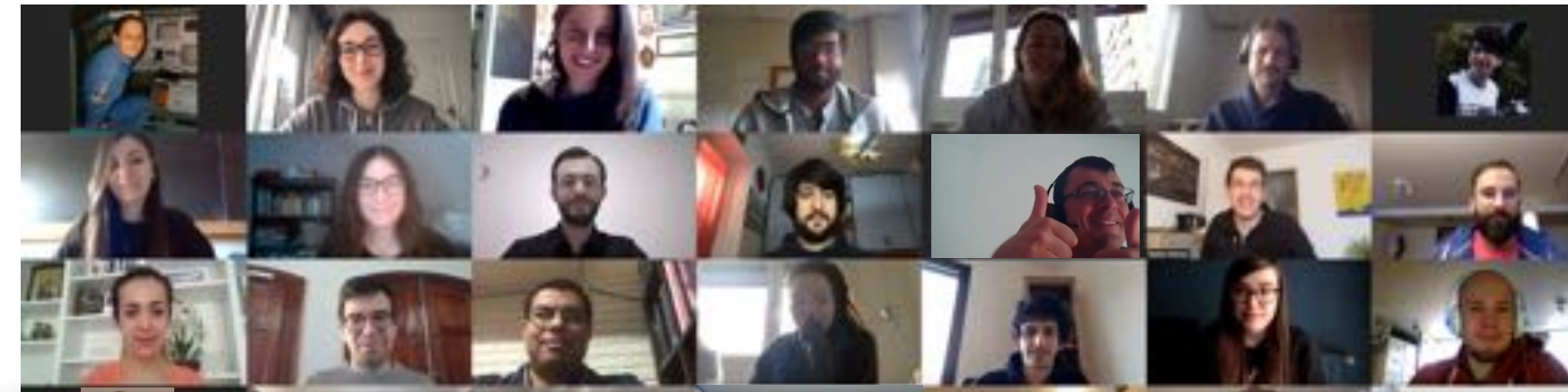
<https://indico.cern.ch/e/cppjan21>
held from 18 – 22 January 2021

Co-organised by [HSE](#) and [SIDIS](#)

1 Lecturer
14 Mentors
20.5 Hours
75 Students
120 people on the waiting list



The event was organised and supported by faculty staff at:



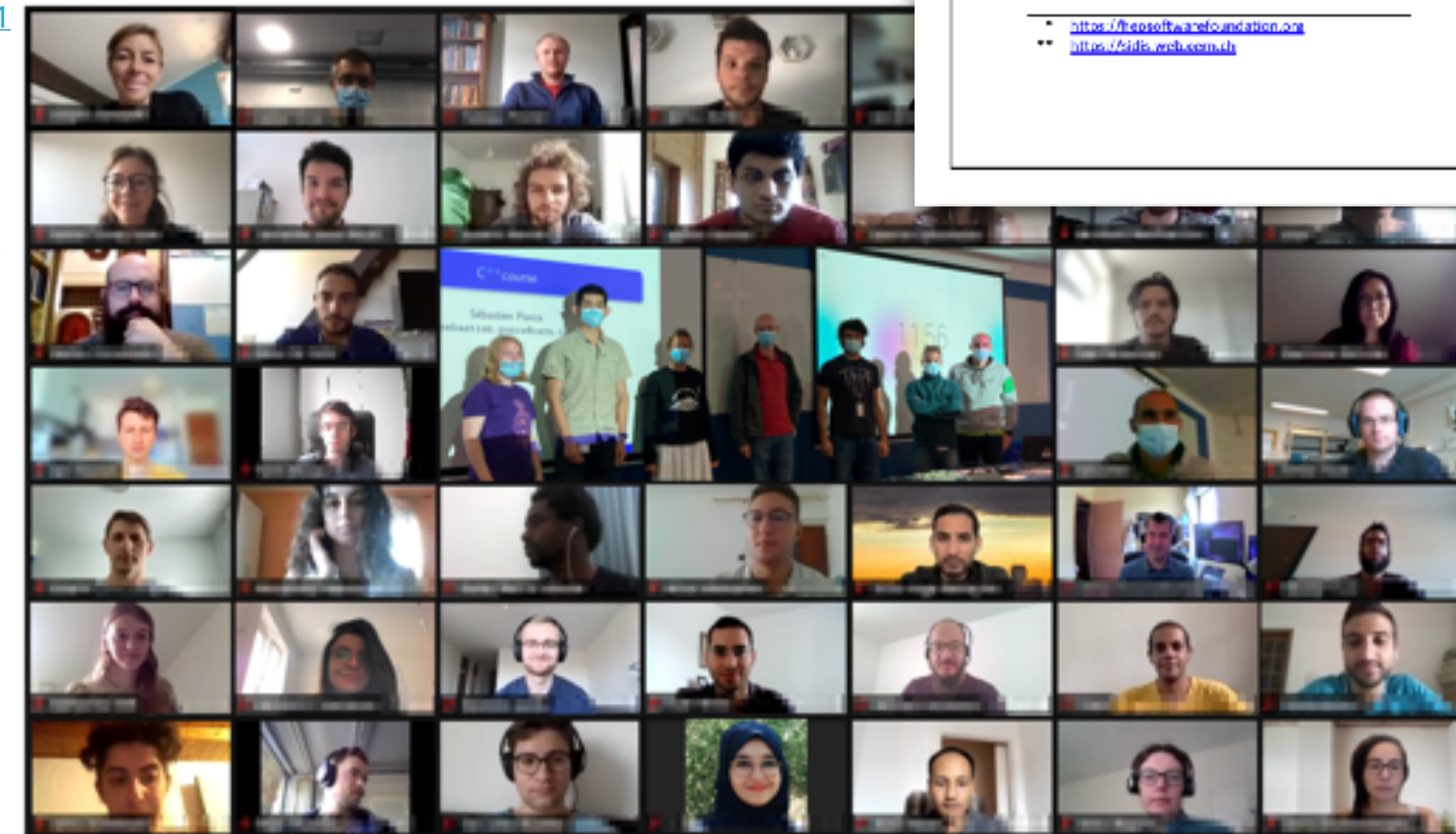
3rd HEP C++ Course and Hands

<https://indico.cern.ch/e/cppaug21>
held from 30 Aug. - 3 Sep. 2021

Co-organised by [HSE](#) and [SIDIS](#)

2 Lecturers
16 Mentors
20.5 Hours
75 Students
90 people on the waiting list

The event was organised and supported by faculty staff at:



CERTIFICATE OF PARTICIPATION

Amy Farrah Fowler

has attended the first HEP C++ Course and Hands-on Training from 12 - 16 October 2020

The course was organised into 12.5 hours of lectures and 9 hours of training sessions.

The lessons were taught by Dr. Sebastian Ponce (CERN) and the training sessions supported by mentors from within the high energy physics community.

The course content ranged from basic to advanced concepts of C++ software engineering and included up to the C++ 17 ISO standard.

The event was co-organised by the training working group of the HEP Software Foundation* and the Software Institute for Data Intensive Sciences**

More information about the event can be found at <https://indico.cern.ch/e/cppfall20>

(Signatures of Dr. Graeme A Stewart, Dr. Sebastian Ponce, and Dr. Stefan Roiser)

Dr. Graeme A Stewart
CERN / HSE

Dr. Sebastian Ponce
CERN / Course Lecturer

Dr. Stefan Roiser
CERN / SIDIS

* <https://hepswf.cern.ch>
** <https://sidis.wsb.cern.ch>

▶ Recently established "Jour fixe" to advance the course content

▶ Every 3rd Wed / month, 15:00 CE(S)T, <https://indico.cern.ch/category/14413/>

▶ Changes to the course in the pipeline: split into basic & advanced, integrate C++20, ...

▶ Working towards a self-sustained activity, with students becoming mentors

COMPUTE ACCELERATOR FORUM

- ▶ Monthly forum with technical presentations on hardware and software aspects on compute accelerators and heterogeneous computing in general.
- ▶ Every 2nd Wed / month 16:30 CE(S)T
 - ▶ Started in Oct 2020
 - ▶ Usually ~ 30 to 60 participants and lively discussions
- ▶ Agenda: <https://indico.cern.ch/category/12741/>
- ▶ Four announcements subscribe to compute-accelerator-forum-announce@cern.ch

Co-organizer: Ben Morgan

Marco Barbone

December 2021	08 Dec	Compute Accelerator Forum - Beams GPU & VecMem
November 2021		
	10 Nov	Compute Accelerator Forum - FPGA
October 2021		
	13 Oct	Compute Accelerator Forum - GTG21 + CERN IT GPU Infra
September 2021		
	08 Sep	Compute Accelerator Forum - Kokkos
July 2021		
	07 Jul	Compute Accelerator Forum - FPGA
June 2021		
	09 Jun	Compute Accelerator Forum - JUWELS and Belle II
May 2021		
	12 May	Compute Accelerator Forum - Alpaka & Llana
April 2021		
	14 Apr	Compute Accelerator Forum - FPGAs
March 2021		
	10 Mar	Compute Accelerator Forum - Detector Geometry
February 2021		
	10 Feb	Compute Accelerator Forum - libcu++, Infrastructure updates
January 2021		
	13 Jan	Compute Accelerator Forum - SYCL, HLS4ML
December 2020		
	09 Dec	Compute Accelerator Forum - NVIDIA Nsight
November 2020		
	11 Nov	Compute Accelerator Forum - Hardware Abstraction, ALICE & LHCb
October 2020		
	01 Oct	Compute Accelerator Forum - GPU Infrastructure

EU COST ACTION – FUNDING PROPOSAL



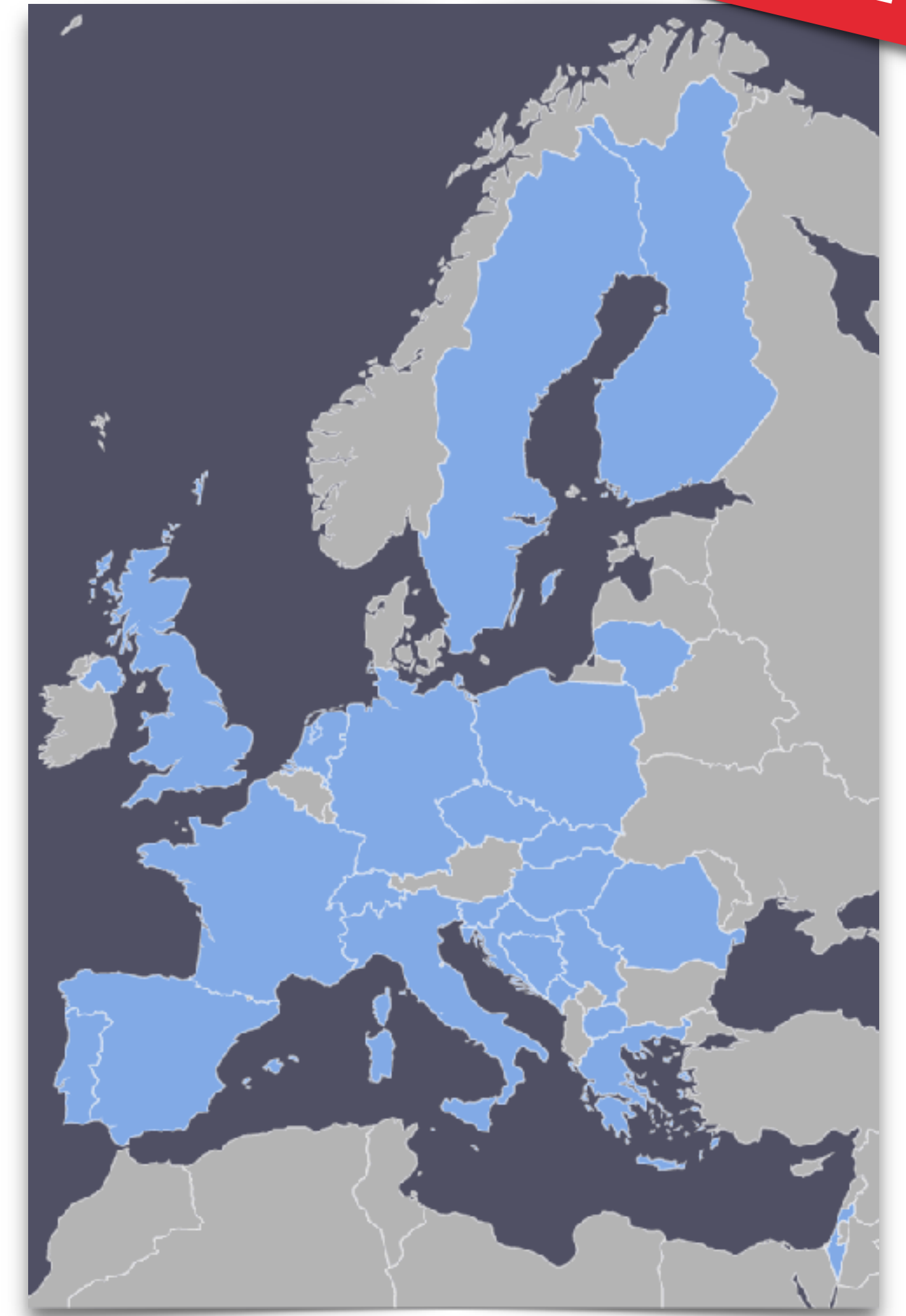
- ▶ COST funds the bottom-up building of pan-European inter-disciplinary research networks by funding
 - ▶ meetings, workshops, schools, training, remote placements
 - ▶ COST is a 4 years funding with a 500 k€ total budget
- ▶ Bring the SIDIS concept onto a larger European level
 - ▶ Building a community around programming for computing accelerators and heterogeneous platforms (data structures, HW agnostic programming, abstraction layers)
 - ▶ Collect and establish training for the necessary tools and languages
 - ▶ Build a catalog of available training (HSF, CERN, SIDIS, CSC, Bertinoro school, ...)
 - ▶ Develop new training when necessary
 - ▶ Raise awareness for the importance of software engineering in natural science

EU COST ACTION – FUNDING PROPOSAL



SUBMITTED 29 OCT '21

- ▶ 30 proponents from 24 European countries
- ▶ Computer science institutes, natural and life sciences (HEP, molecular biology, medicine, ...), HPCs, institutes bridging CS and NS, SMEs
- ▶ UK represented by Caterina Doglioni (U Manchester) and Neil Chue Hong (U Edinburgh/UK-RSE)
- ▶ Decision about funding mid-May 2022
- ▶ NB: COST has an “open funding model”, i.e. joining the initiative also after approval shall be easy to do



LINKS

▶ SIDIS

- ▶ Homepage: <https://sidis.web.cern.ch/>
- ▶ Mailing list: proto-sidis-concept@cern.ch *

* direct subscription: <https://e-groups.cern.ch/e-groups/EgroupsSubscription.do?egroupName=proto-sidis-concept>