UNIVERSITY OF COPENHAGEN

EUROPEAN SPALLATION SOURCE

Workshop on Perspectives and Applications of (Machine &) Deep Learning for Accelerated Scientific Discovery in Physics - Wrap-up (part 1)

Heloisa N. Bordallo

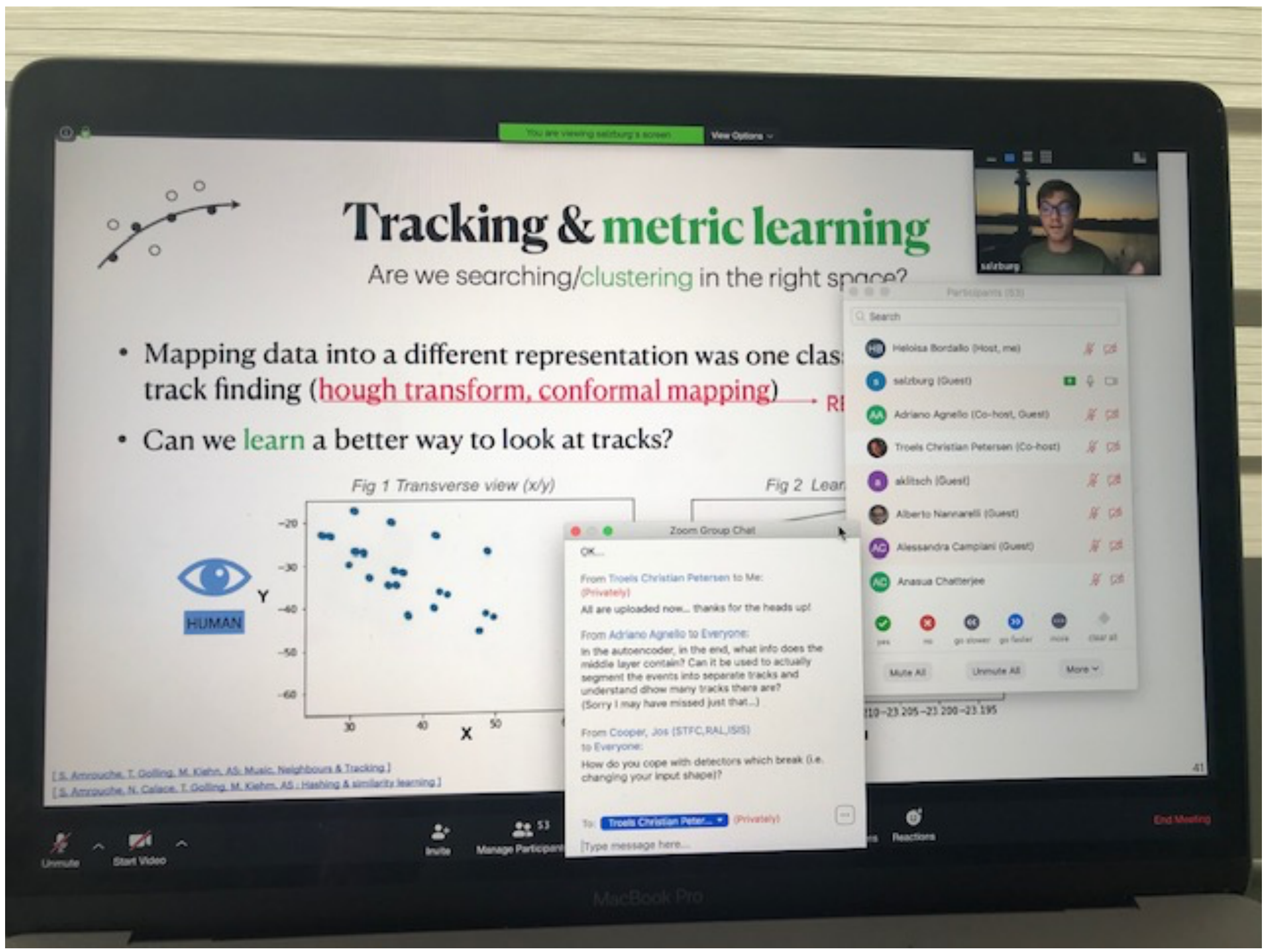
Assoc. Prof. in X-rays and Neutron Science - Niels Bohr Institute - University of Copenhagen

Scientific Advisor of the MIRACLES Spectrometer project at the ESS

bordallo@nbi.ku.dk

Niels Bohr Institute - University of Copenhagen - May 15, 2020

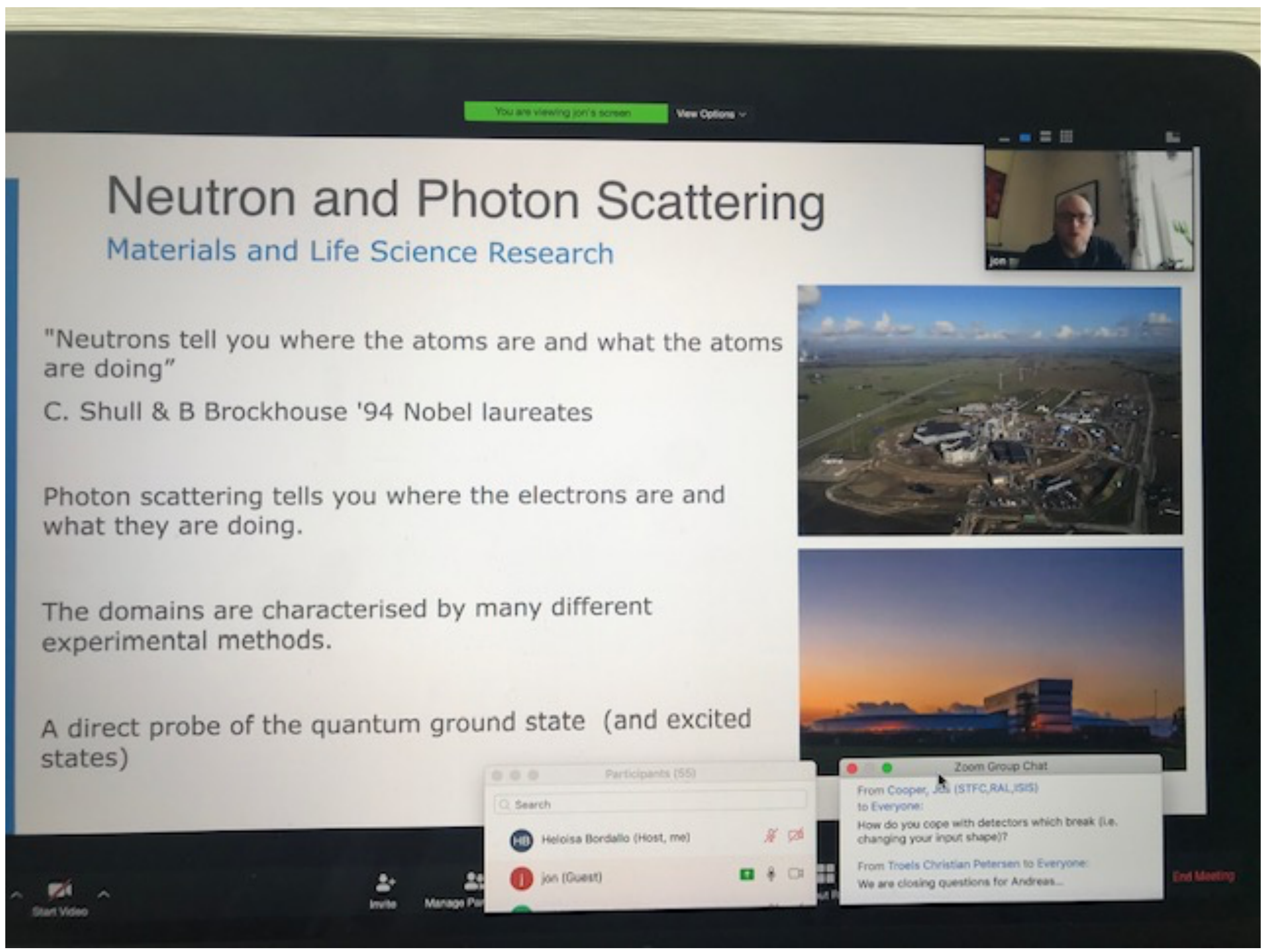
Machine learning and particle physics by A. Salzburger (CERN)



- ☑ Data analysis the most prominent field in ML.
- ☑ No out of the box solution.
- ☑ NN trained on simulated data: How good is the data?
- ☑ CPU time is expensive.
- ☑ DL assisted by ML: simulate boring events that could be interesting.

EUROPEAN SPALLATION SOURCE

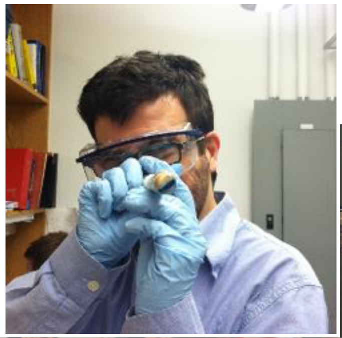
Data Challenges for accelerating scientific discovery at Neutron facilities by J. Taylor (ESS)



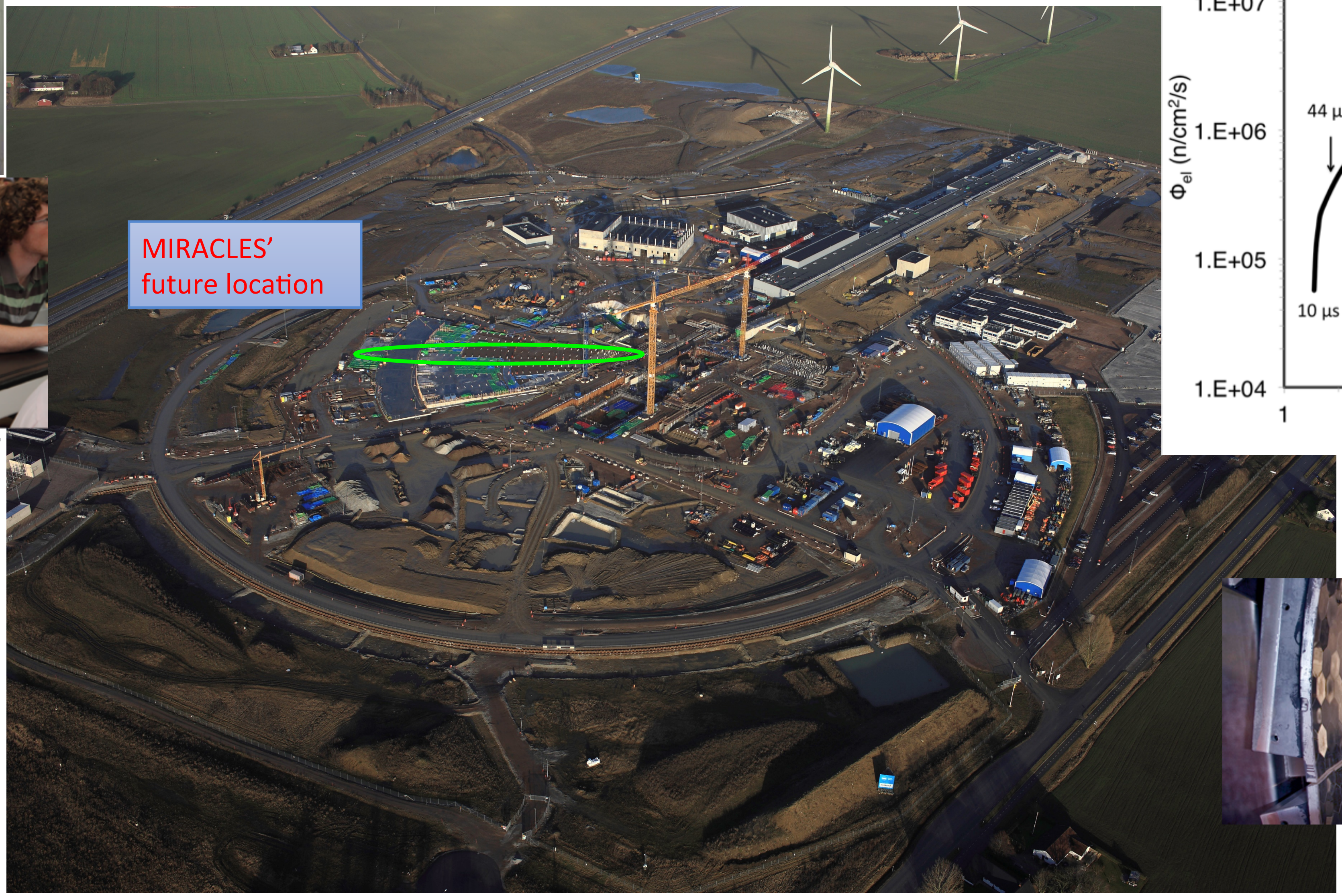
- Challenge 1: Different types of data and different Physics.
- Challenge 2: Velocity of data input and output.
- Challenge 3: Data processing becomes a limiting factor.
- Challenge 4: Lack of labelled training data, experimental background, analytical understanding of the process.
- First successes: Data mining, atomistic codes and ground work at STFC.

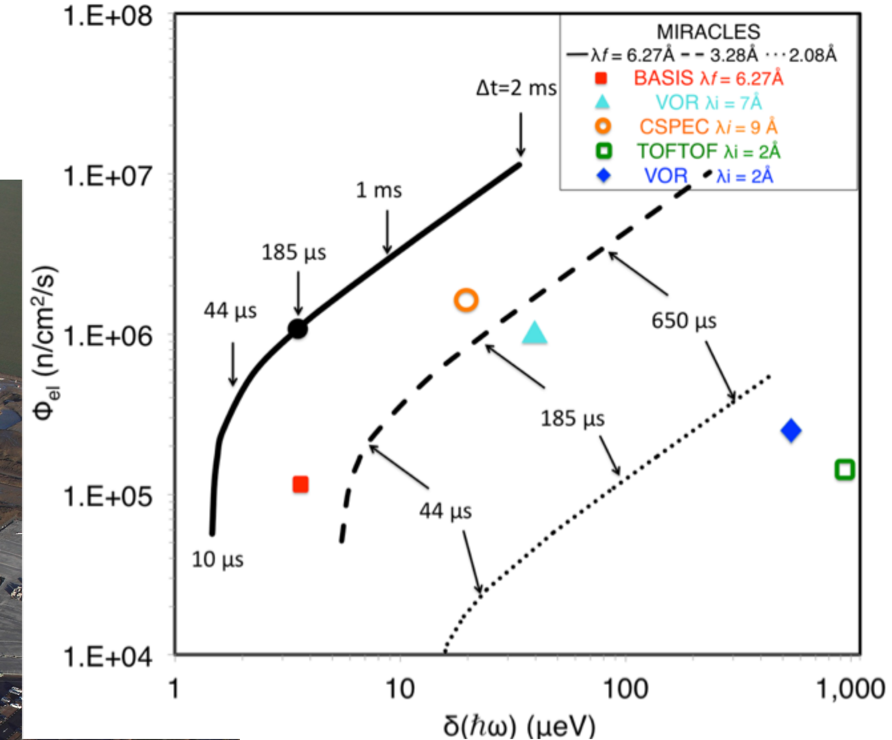
EUROPEAN SPALLATION SOURCE

MIRACLES DO happen!*



Ruep E. Lechner
✝02/04/2019





Luna, P *et al* (2019) Physica B **564**, 64–69.
Tsapatsaris, N. *et al* (2016) Rev. Sci. Instrum. **87**, 085118.
Tsapatsaris N. *et al* (2015) EPJ Web of Conferences, **83**, 03015.

*… with some help … MIRACLES will cost 13.525 M€.

EUROPEAN SPALLATION SOURCE

… but we still analysing data as we if we were in the 70s'

Future directions in quasi-elastic neutron scattering

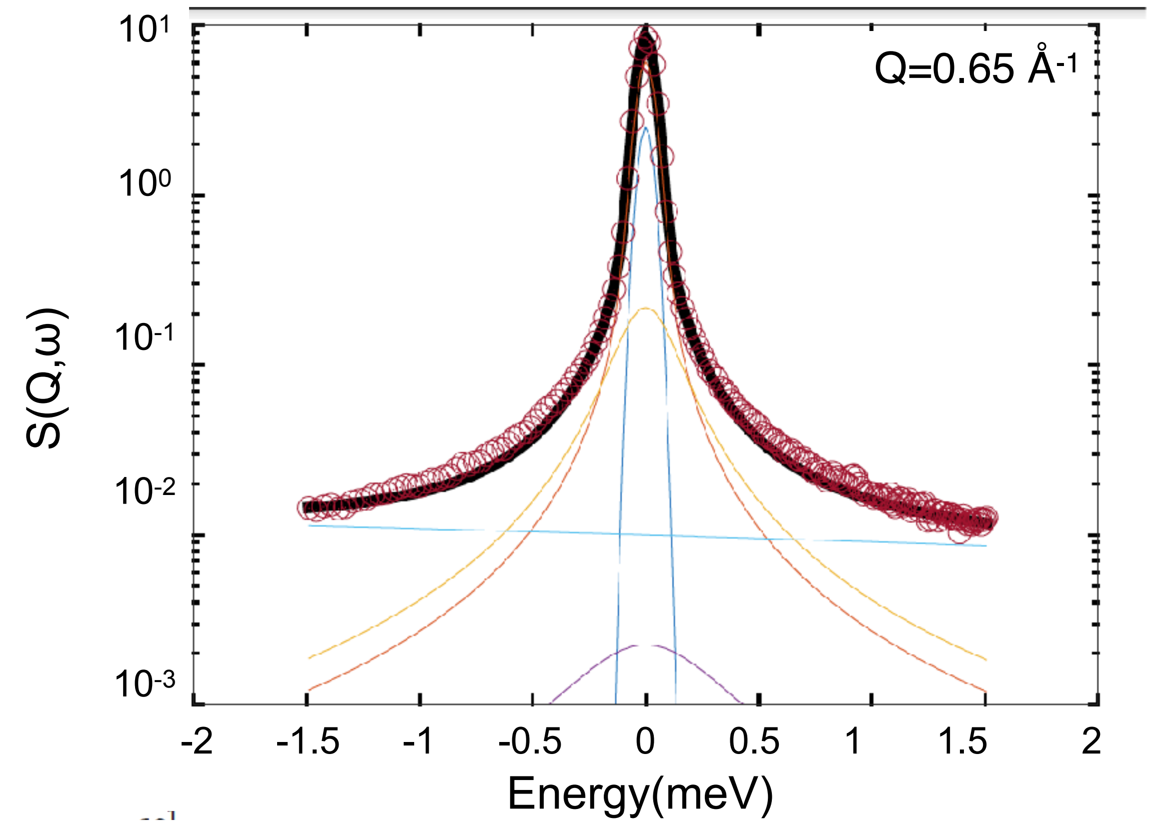
A.J. Dianoux
Institut Laue–Langevin, Grenoble Cedex, France

Physica B 182 (1992) 389–402
North-Holland

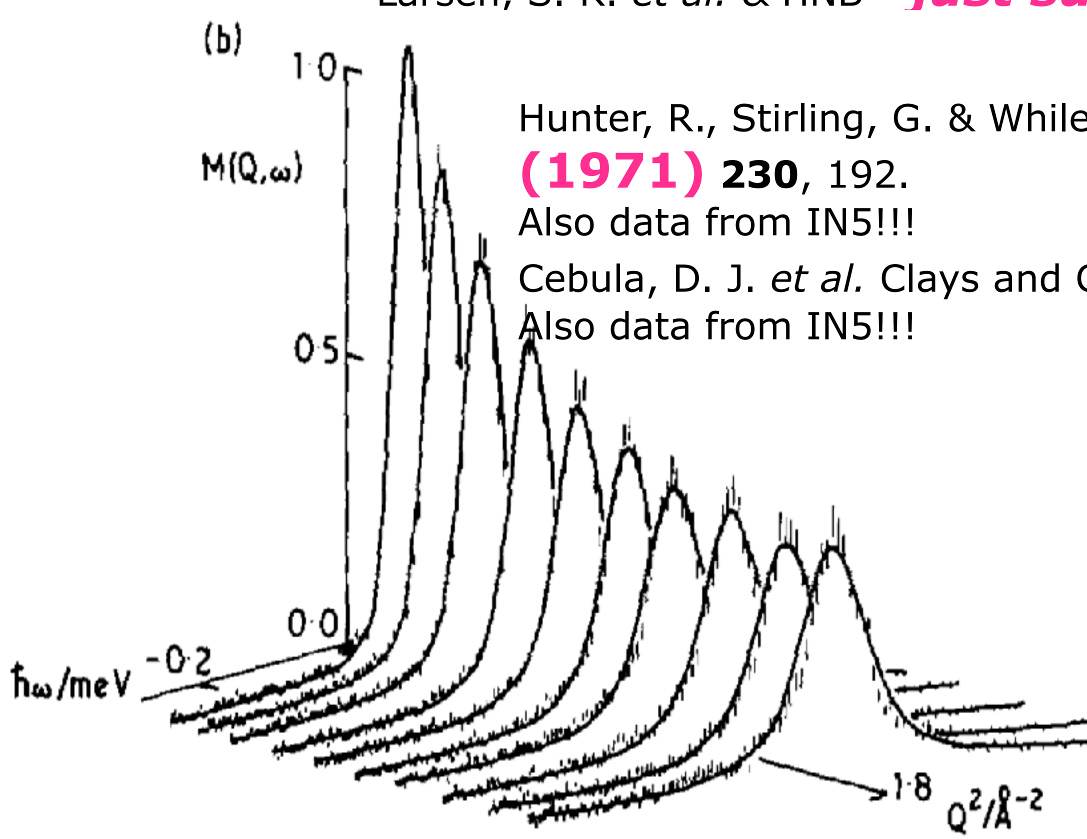
4. Conclusions

We can return to the main purpose of this paper which could be rephrased as: what is the future of QENS? For me this is not a scientific question, but more a political one (if one considers as political the funding of neutron facilities).
I have shown, using some selected examples, that the information brought by QENS is more *direct*, and sometimes more *detailed* than that obtained by the other spectroscopic techniques
.....

However, the full capabilities of QENS can only be realized if one has access to high intensity neutron sources, either continuous (reactor) or pulsed (spallation source). Furthermore, one needs to have an ongoing development of sophisticated spectrometers.



Larsen, S. R. *et al.* & HNB - ***just submitted*** - Data from IN5.



Hunter, R., Stirling, G. & While, J. Nature Physical Science 230, **(1971)** **230**, 192.
Also data from IN5!!!

Cebula, D. J. *et al.* Clays and Clay Miner **(1981)** **29**, 241.
Also data from IN5!!!

EUROPEAN SPALLATION SOURCE

… but we still analysing data as we if we were in the 70s'

Future directions in quasi-elastic neutron scattering

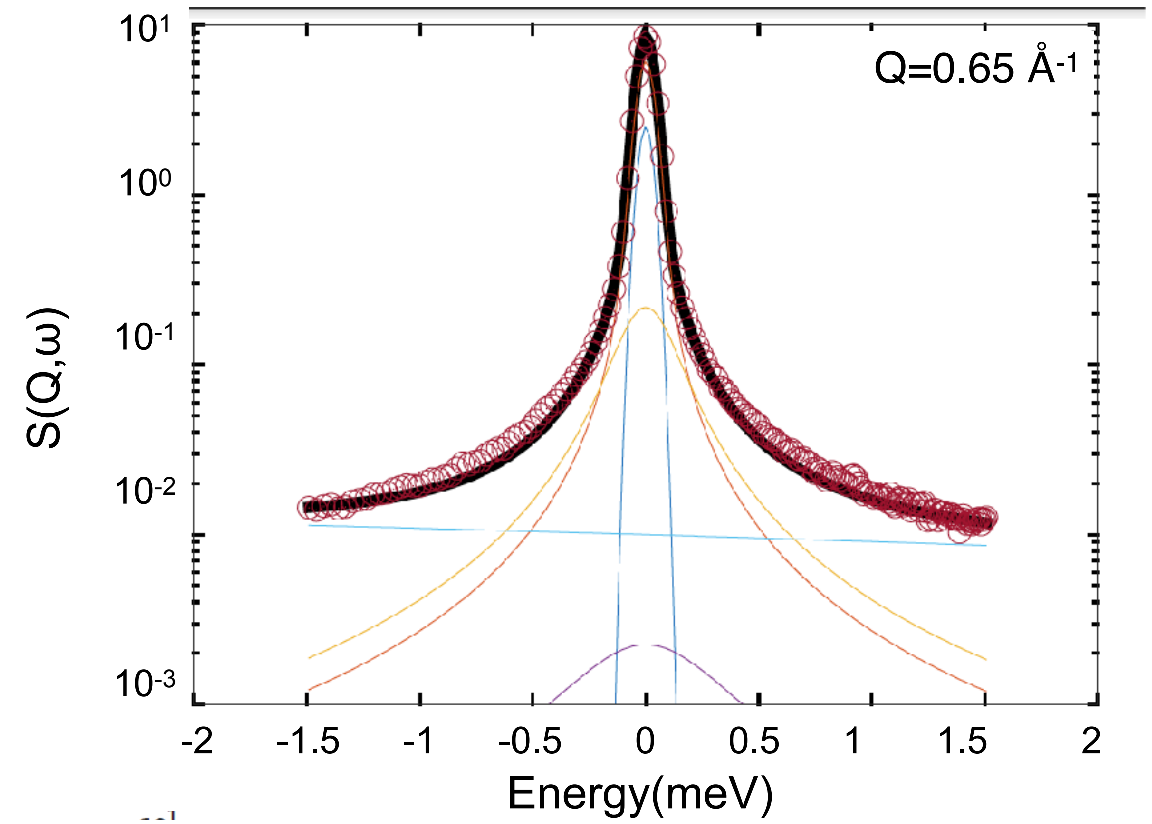
A.J. Dianoux
Institut Laue–Langevin, Grenoble Cedex, France

Physica B 182 (1992) 389–402
North-Holland

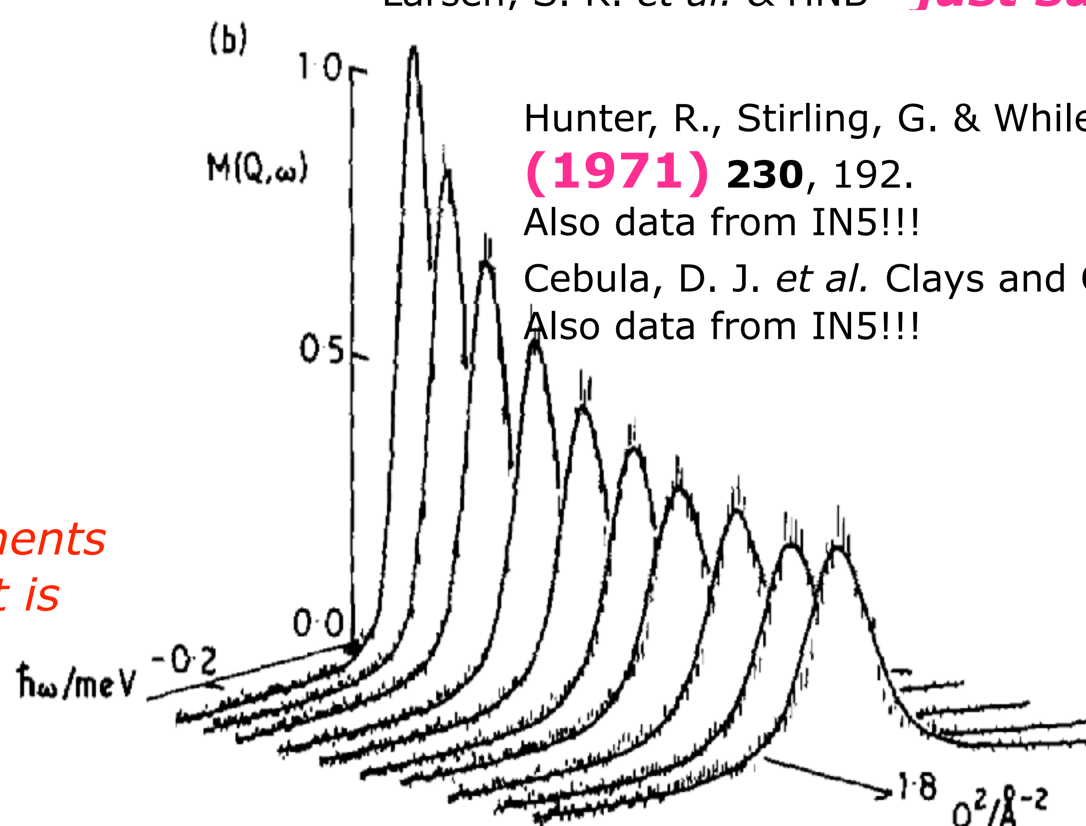
4. Conclusions

We can return to the main purpose of this paper which could be rephrased as: what is the future of QENS? For me this is not a scientific question, but more a political one (if one considers as political the funding of neutron facilities). I have shown, using some selected examples, that the information brought by QENS is more *direct*, and sometimes more *detailed* than that obtained by the other spectroscopic techniques …..

However, the full capabilities of QENS can only be realized if one has access to high intensity neutron sources, either continuous (reactor) or pulsed (spallation source). Furthermore, one needs to have an ongoing development of sophisticated spectrometers.



Larsen, S. R. *et al.* & HNB - ***just submitted*** - Data from IN5.



Hunter, R., Stirling, G. & While, J. Nature Physical Science 230, **(1971)** **230**, 192.
Also data from IN5!!!

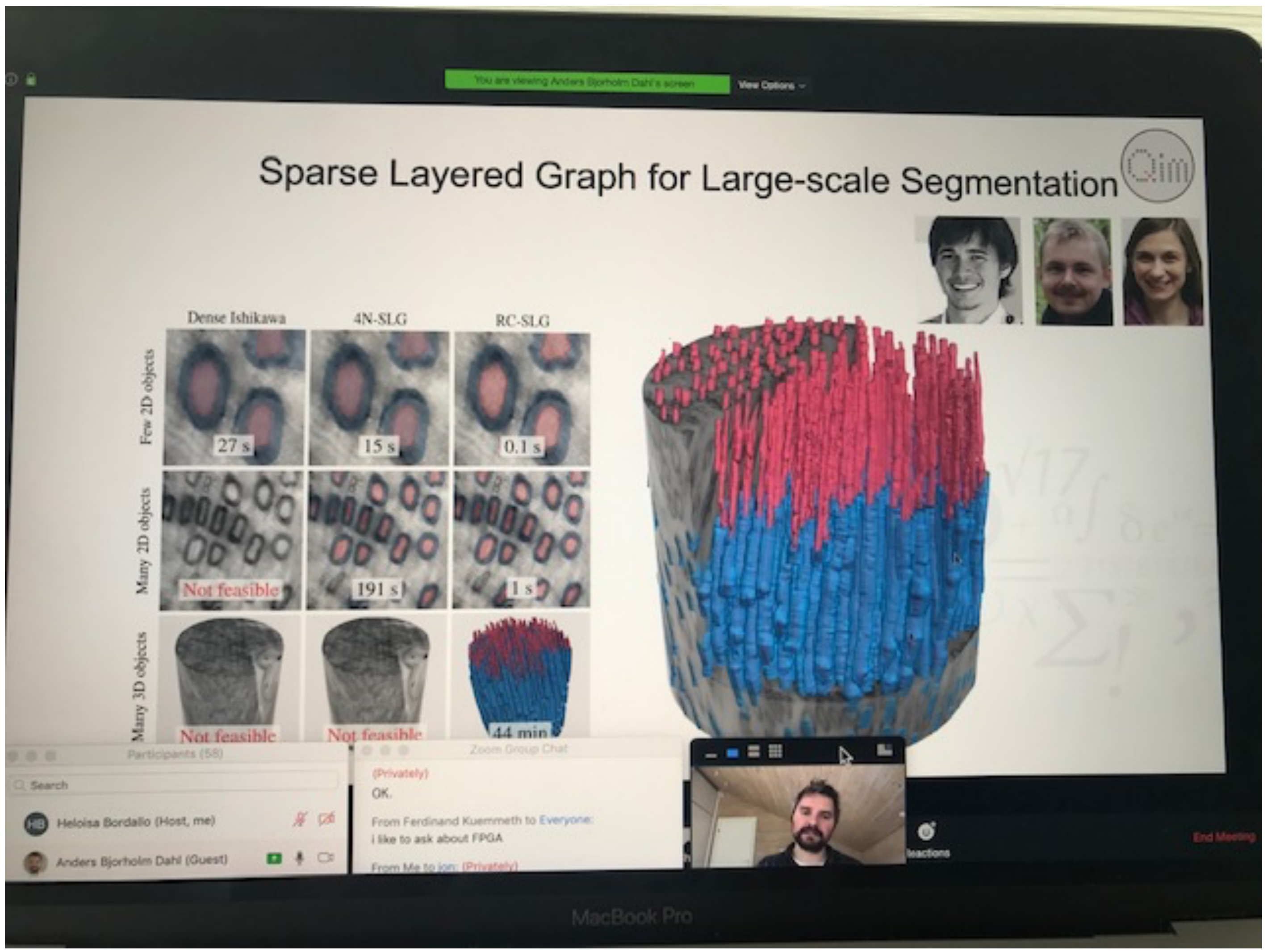
Cebula, D. J. *et al.* Clays and Clay Miner **(1981)** **29**, 241.
Also data from IN5!!!

… Remember … at ESS instruments performance will surpass what is already available…
And what about analysis?

Aldridge, L.P., Larsen, S.R. & HNB Physica B (2019) **561** 75.

EUROPEAN SPALLATION SOURCE

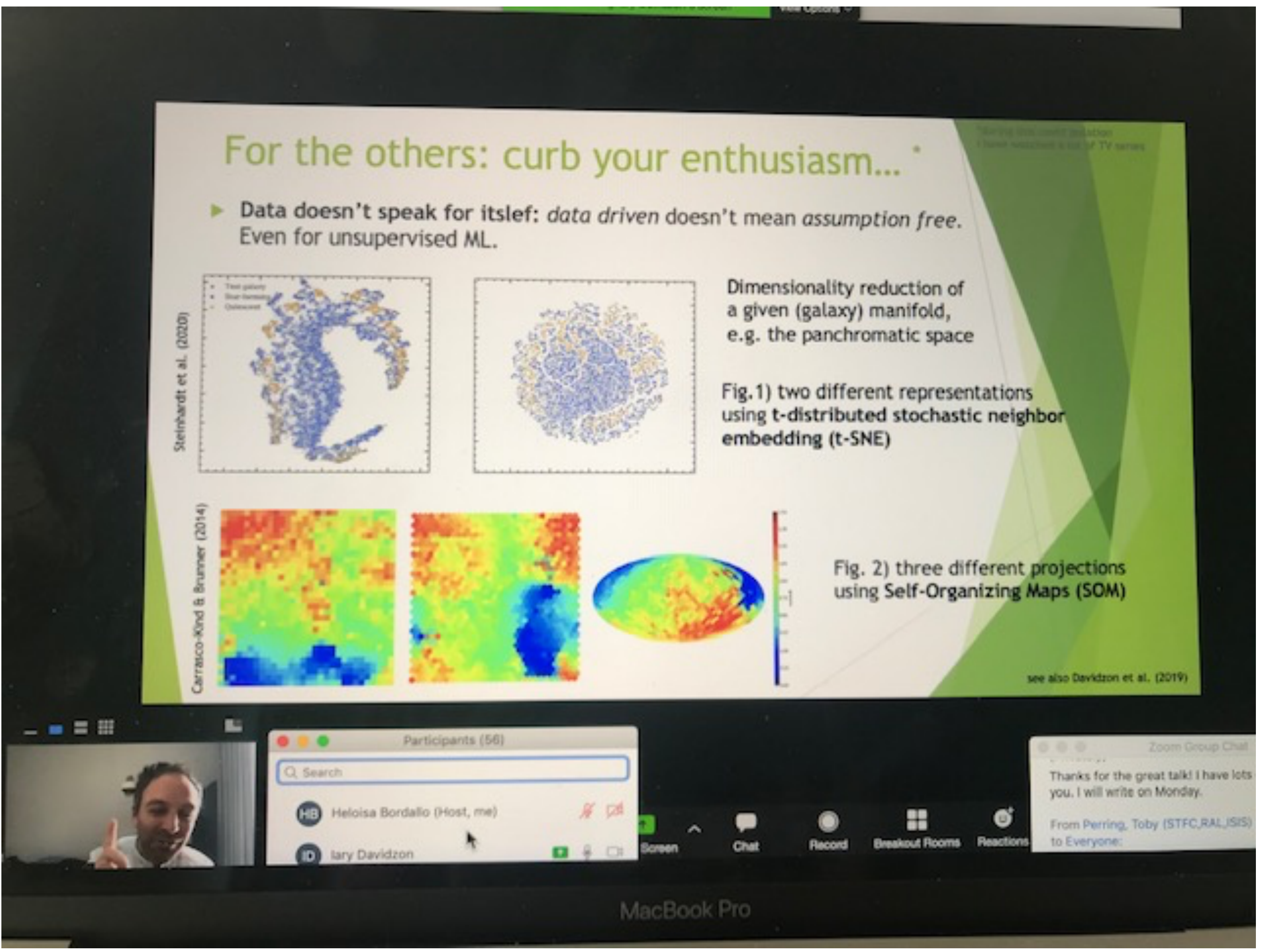
Collaboration on Analysis of 3D Image Data in the QIM Center by A. Dahl (DTU)



- ☑ Image analysis for 3D tomography not established like medical image analysis; but users knows the sample!
- ☑ AI/ML can help with segmentation and higher throughput.
- ☑ Needs: Methods that require little training data (each sample is unique); quick visualisation (data vs result); annotation and correction.

EUROPEAN SPALLATION SOURCE

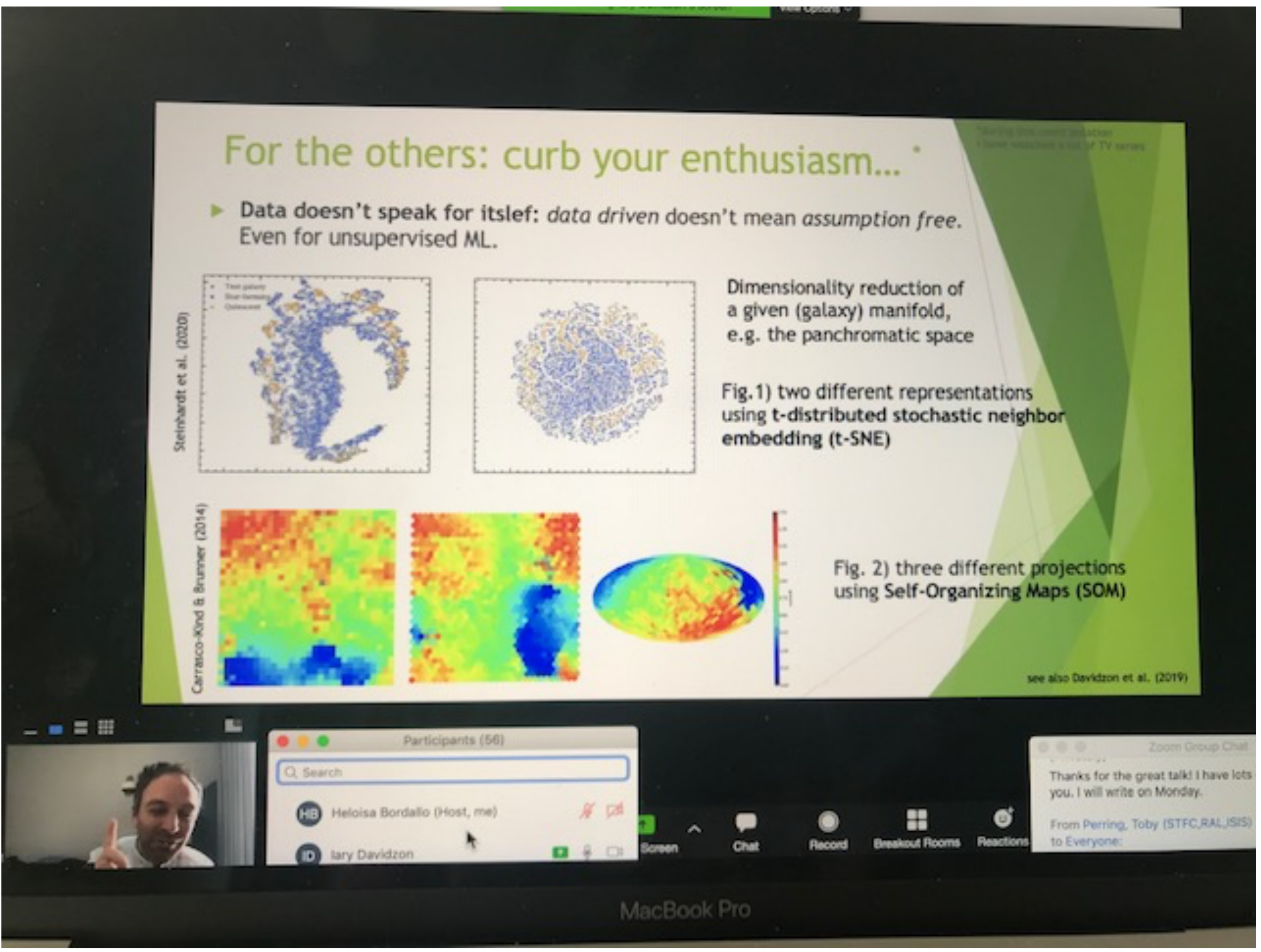
A data-driven approach in Astrophysics by I. Davidzon (NBI)



- ☑ Help with big data from future surveys.
- ☑ Methods that are well tested in industry may not be suitable for astrophysics.
- ☑ Treatment of uncertainties in the ML algorithm itself; one solution among many.

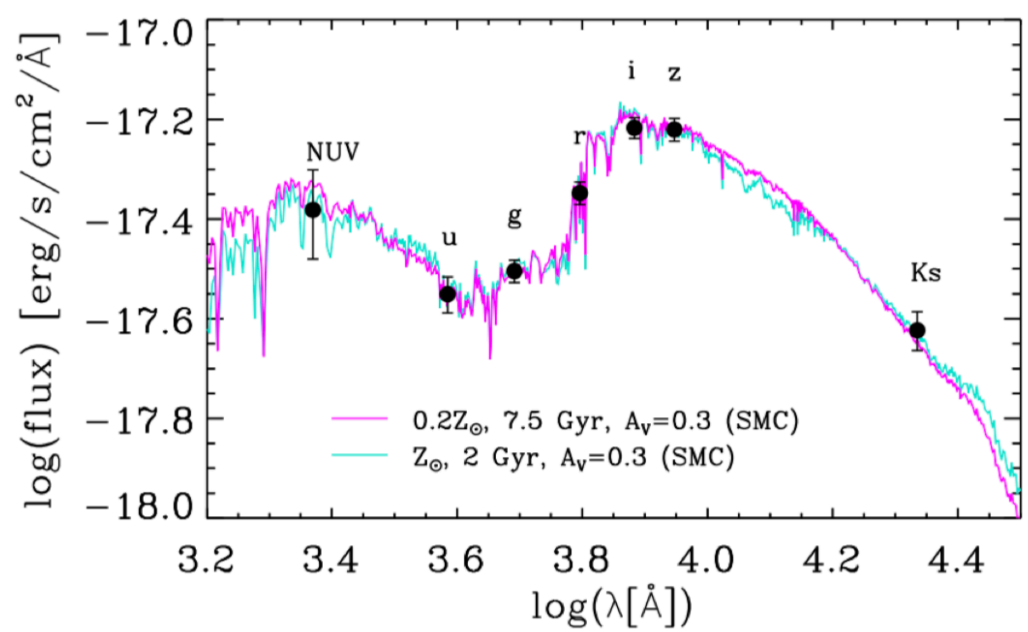
EUROPEAN SPALLATION SOURCE

A data-driven approach in Astrophysics by I. Davidzon (NBI)



- ☑ Help with big data from future surveys.
- ☑ Methods that are well tested in industry may not be suitable for astrophysics.
- ☑ Treatment of uncertainties in the ML algorithm itself; one solution among many.

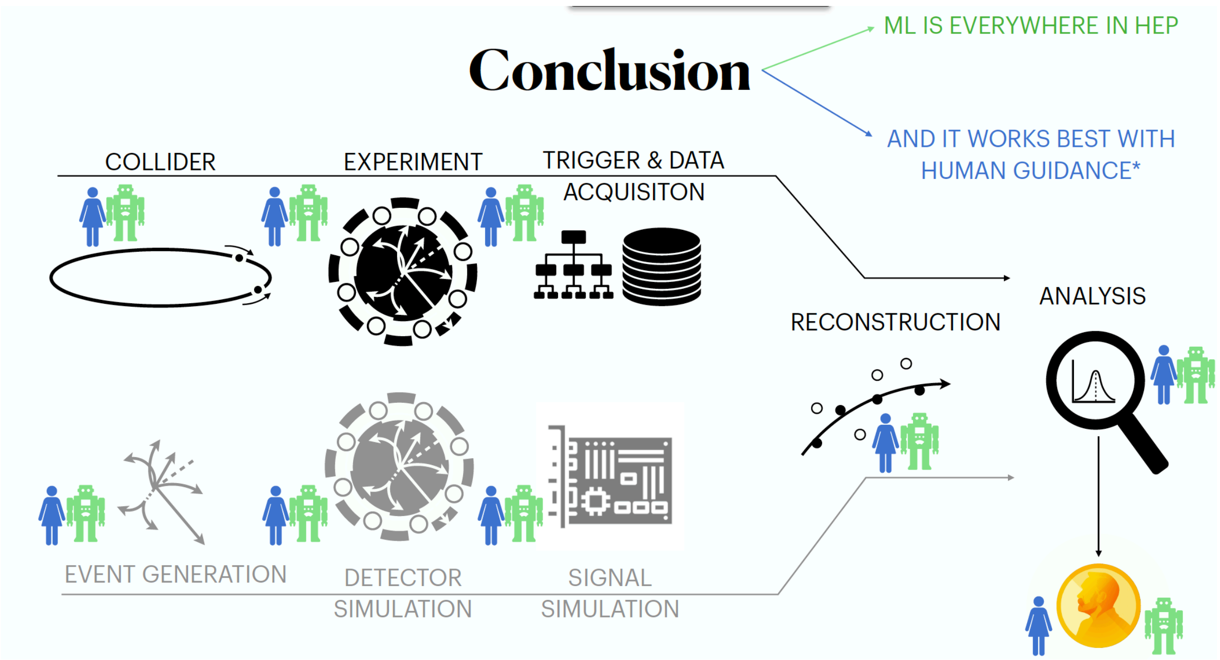
☑ Light coming from a Galaxy!

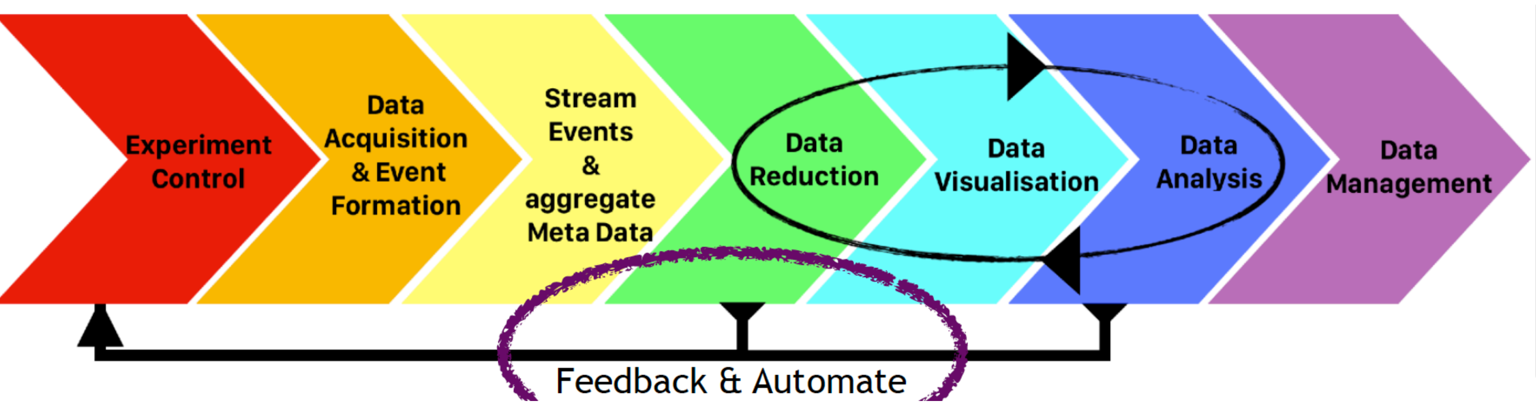


classical method: fitting galaxy models to photometric data points

EUROPEAN SPALLATION SOURCE

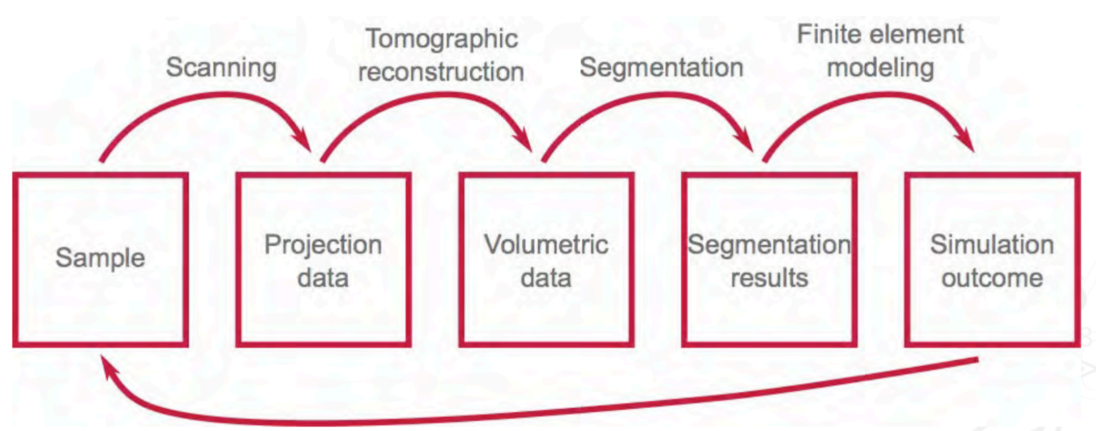
More commonalities than we may think: Lots of room for collaborations





- ☑ By all means learn from CERN experiences, as the imaging community can learn from medical imaging.
- ☑ Understand what is behind a solution: sharing algorithms, best data structure and practical experience.

1. Data collection, reduction, management
2. Extracting information from signal (measurements with error bars)
3. Interpretation (e.g. model fitting) to produce codified knowledge
4. From knowledge to “wisdom”: causality, big picture, future perspective

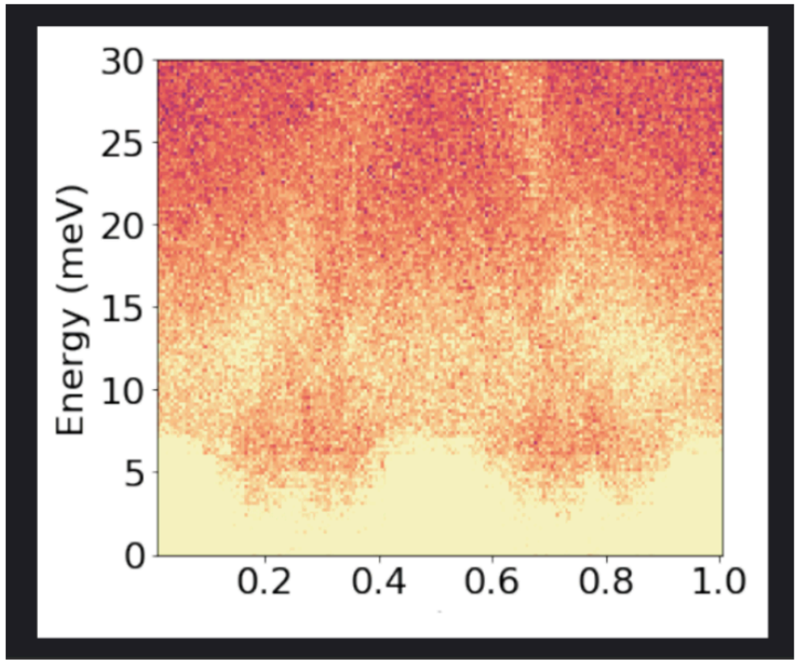


EUROPEAN SPALLATION SOURCE

More commonalities than we may think: Data deluge and data heterogeneity

- ☑ Data storage and CPU are expensive.
- ☑ Data popularity, data monitoring (is the experiment working, is my data useful?), data does not speak for itself (assumptions even in even unsupervised ML), measurement vs experiment…
- ☑ Difficulties in selecting relevant method, quantifying structures become the most time-consuming part (of imaging…).

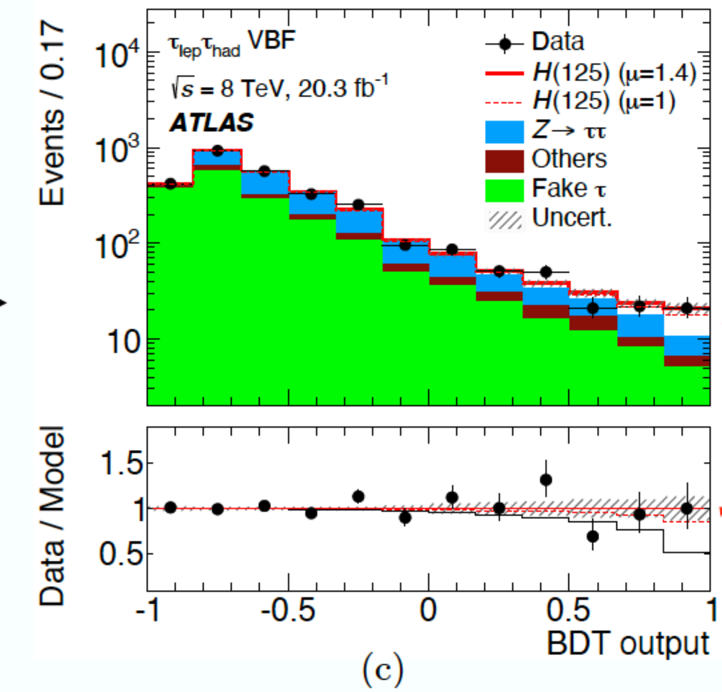
More commonalities than we may think: Dealing with noise



Useful data + noise + background(s)

- ☑ Do you know what you are looking for?

BDT →



Evidence for Higgs Boson

- ☑ A distinctive feature of Astronomy **(and Scattering)** is the complexity & uncertainty in the data: Common ML approaches ?

EUROPEAN SPALLATION SOURCE

More commonalities than we may think:

- ☑ We need out of the box solutions!
- ☑ For students it is a great opportunity to make a difference in Physics!
- ☑ The ultimate goal:

More commonalities than we may think:

- ☑ We need out of the box solutions!
- ☑ For students it is a great opportunity to make a difference in Physics!
- ☑ The ultimate goal:

EUROPEAN
SPALLATION
SOURCE

Conclusion: ML can revolutionise the way science is done

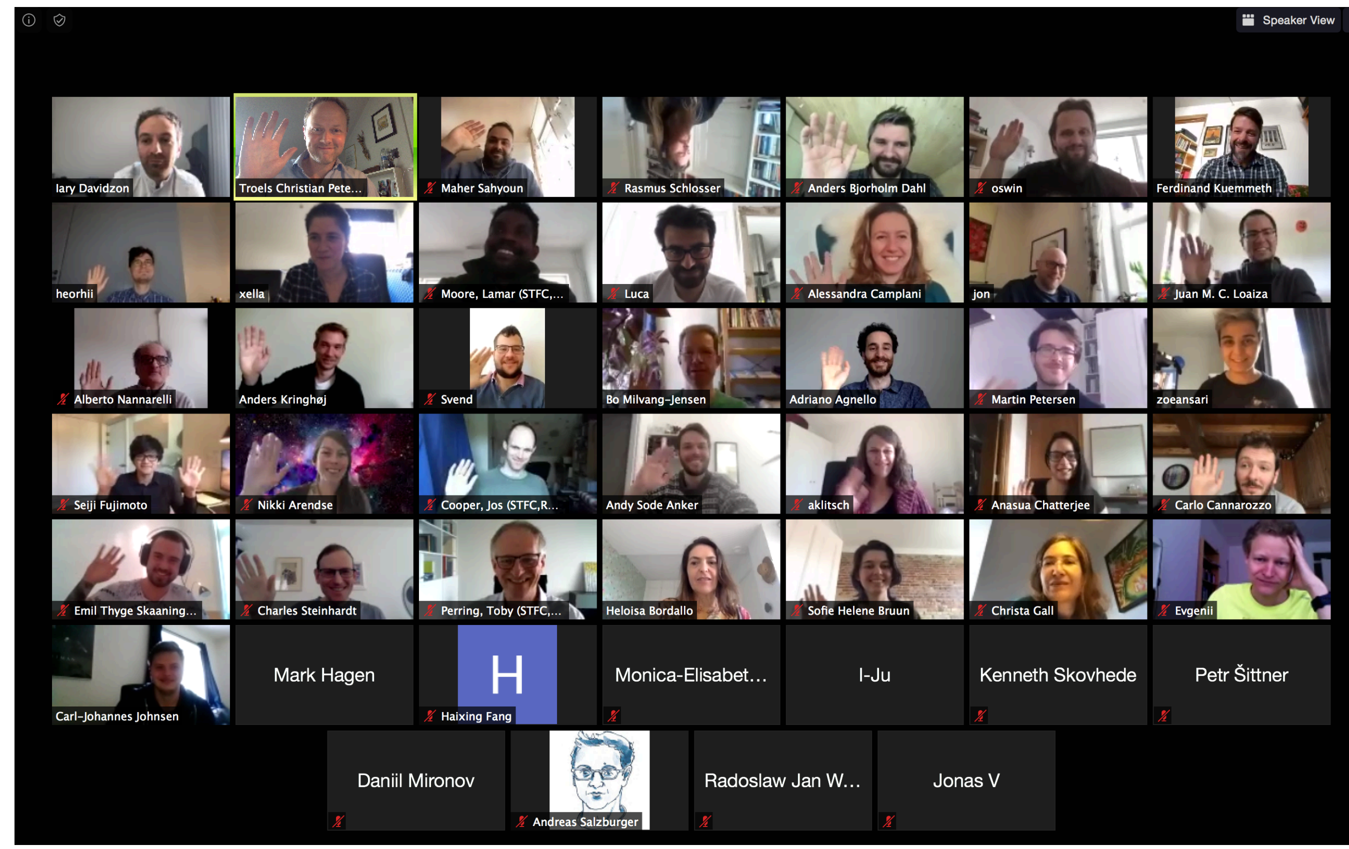
- Correlation vs Causation
  - We want to discover mechanisms
  - And understand why
- ML for automation
  - assisting users set up and configure
  - data driven automation for workflows
  - increased throughput - commander data
    - is it good or is it bad
  - automated discovery

Conclusion: ML can revolutionise the way science is done

- Correlation vs Causation
  - We want to discover mechanisms
  - And understand why
- ML for automation
  - assisting users set up and configure
  - data driven automation for workflows
  - increased throughput - commander data
    - is it good or is it bad
  - automated discoverv
- Easy to underestimate complexity need to dive in and explore
- Collaboration between ML experts is essential
  - Build collaboration and communication channels
  - Build a set of useful tools

Workshop on Perspectives and Applications of Deep Learning for Accelerated Scientific Discovery in the next generation X-ray and Neutron Sources - UCPH - 2019*

Workshop on Perspectives and Applications of Deep Learning for Accelerated Scientific Discovery in Physics - UCPH - 2020



*H. N. Bordallo, C. Lioma, J. Taylor and D. N. Argyriou (2020) IUCrJ **7**, 1.

EUROPEAN SPALLATION SOURCE