Integrating High Resolution Solar Physics

Report on WP5:*Towards a European Solar Data Centre*

SOLARNET General Assembly
23.01.2020

N. Bello González & WP5 participants

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

Themis, SST,
VTT, GREGOR

- European ground-based solar obs. facilities
- Strong efforts by the European solar physics community to build EST (ESFRI) in the Canary Islands

European participation in the DKIST/US project — Access to DKIST observations and data

European Open Science Cloud

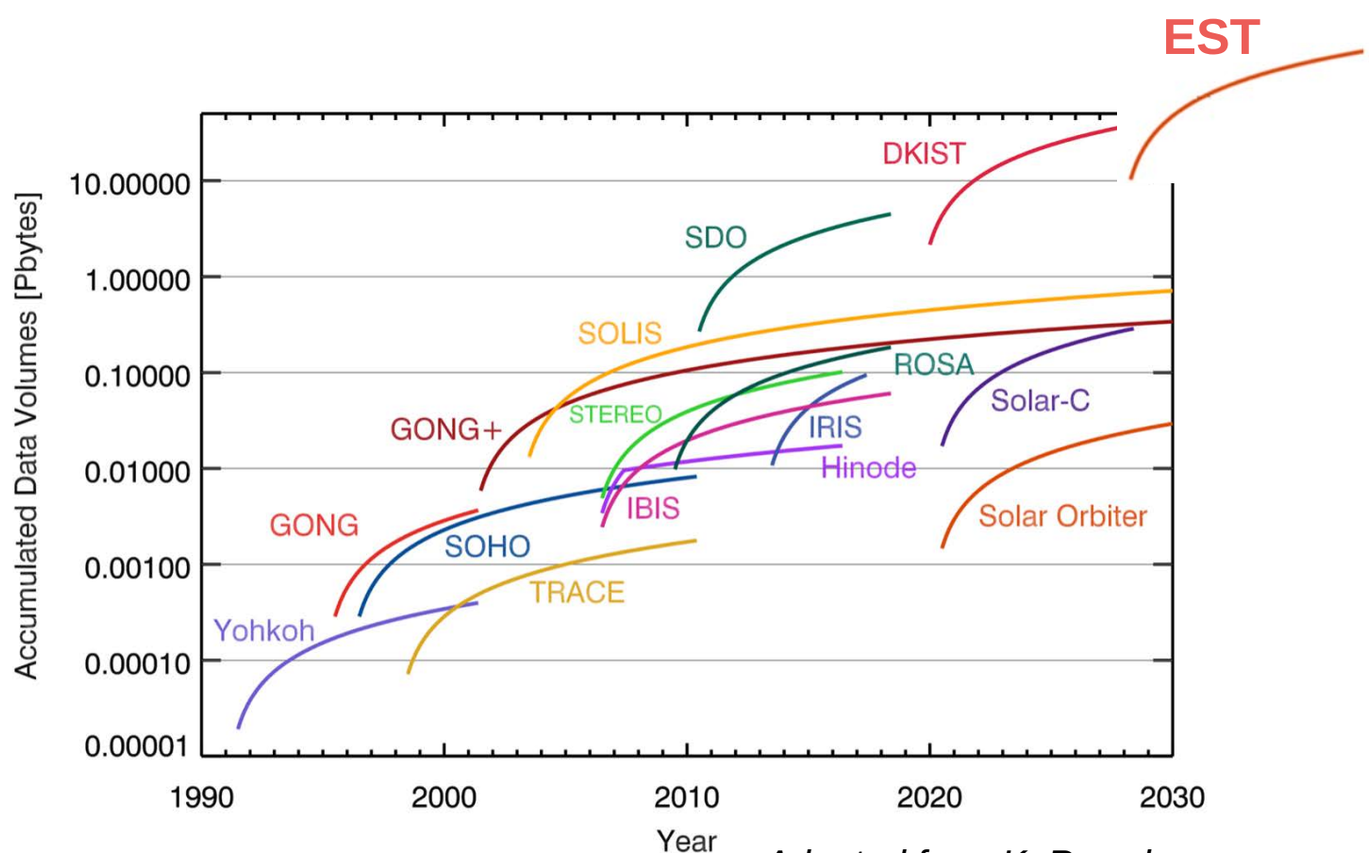
The EOSC is envisaged to be Europe's virtual environment for all researchers to store, manage, analyse and re-use data for research, innovation and educational purposes

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

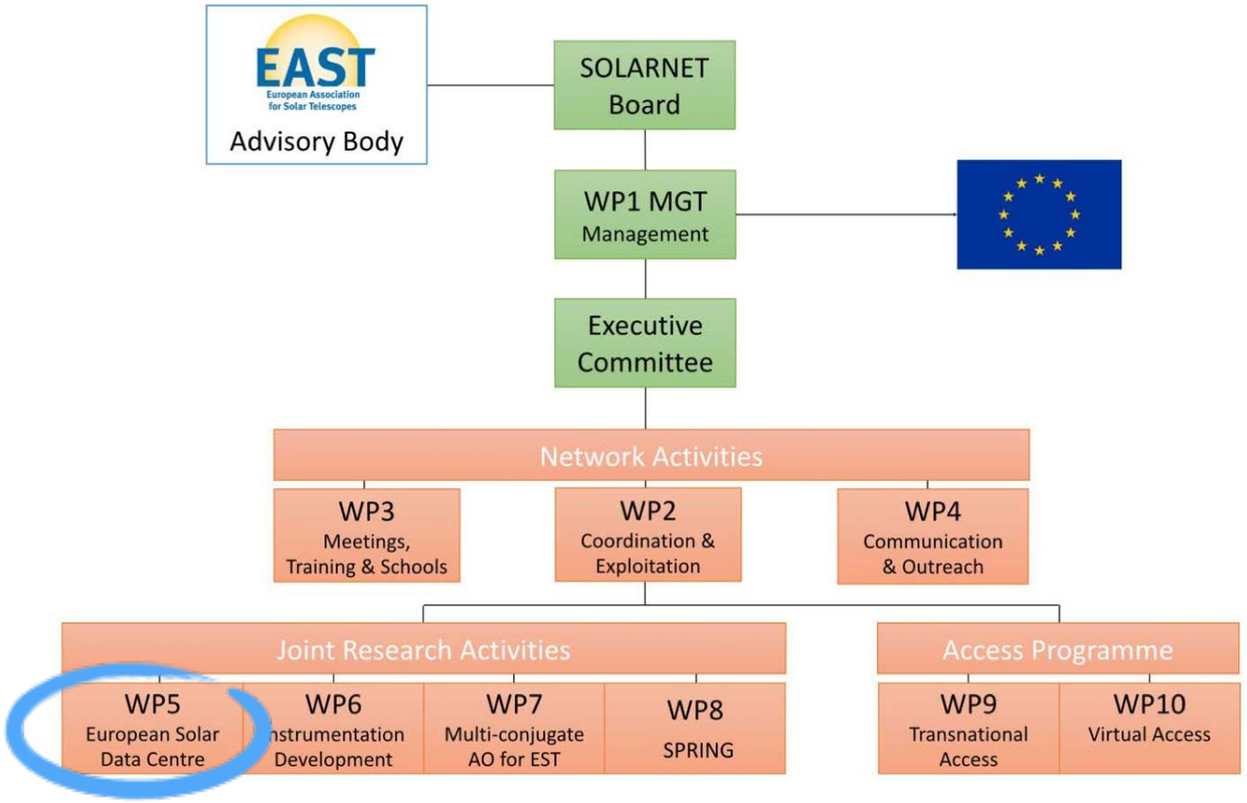
Ground-based solar observations

The number of data sets available has grown, and will continue to grow as an increasing amount of data flow from **ground-based** observations are made and will be available

http://sdc.uio.no/open/solarnet-20.3/WP20.3%20Deliverable%20D20.4_v1.2.pdf



Adapted from K. Reardon



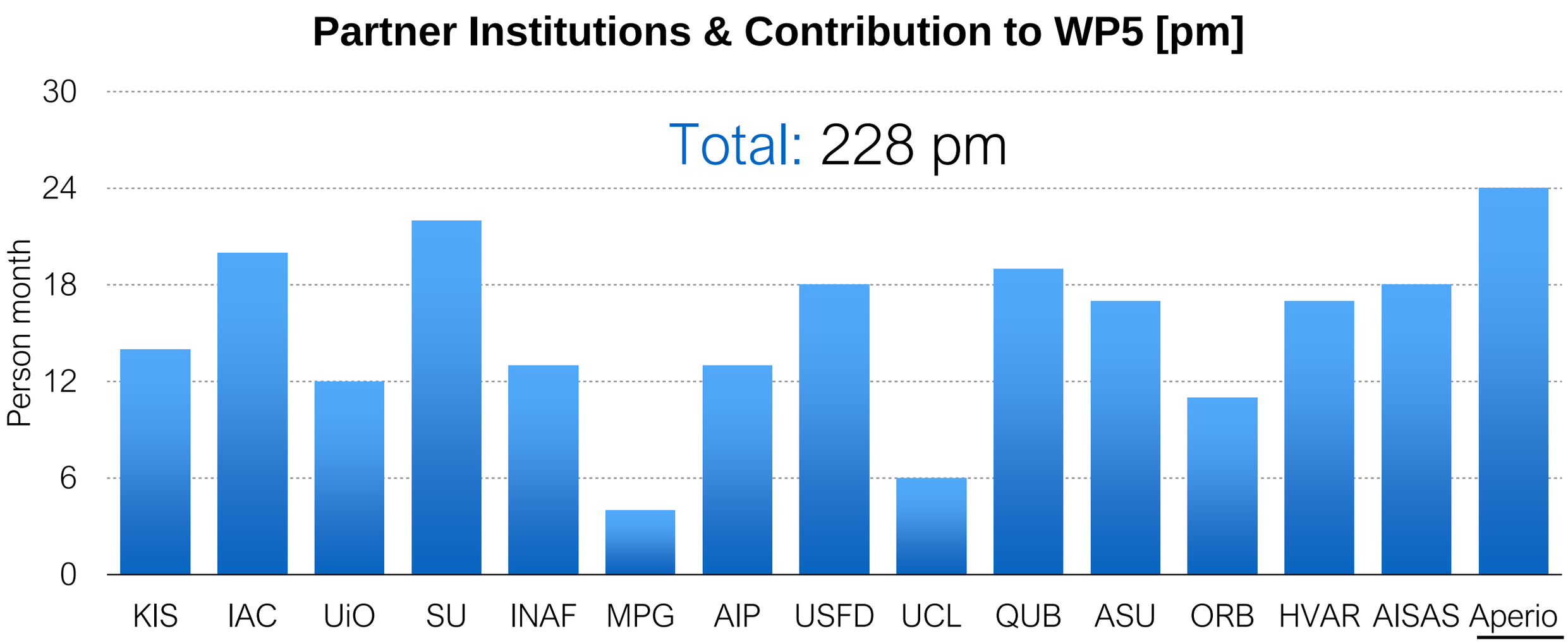
- WP5: Data centre activities — Towards a European Solar Data Centre

Objectives. Given the number of solar infrastructures operated and under development by the European institutions, WP5 focuses on data centre activities of common interest to the Solar Physics community for the curation and dissemination of solar data, to boost its use and re-use and thus optimise its scientific exploitation in line with the spirit/essence of the European Open Science Cloud (EOSC) project.

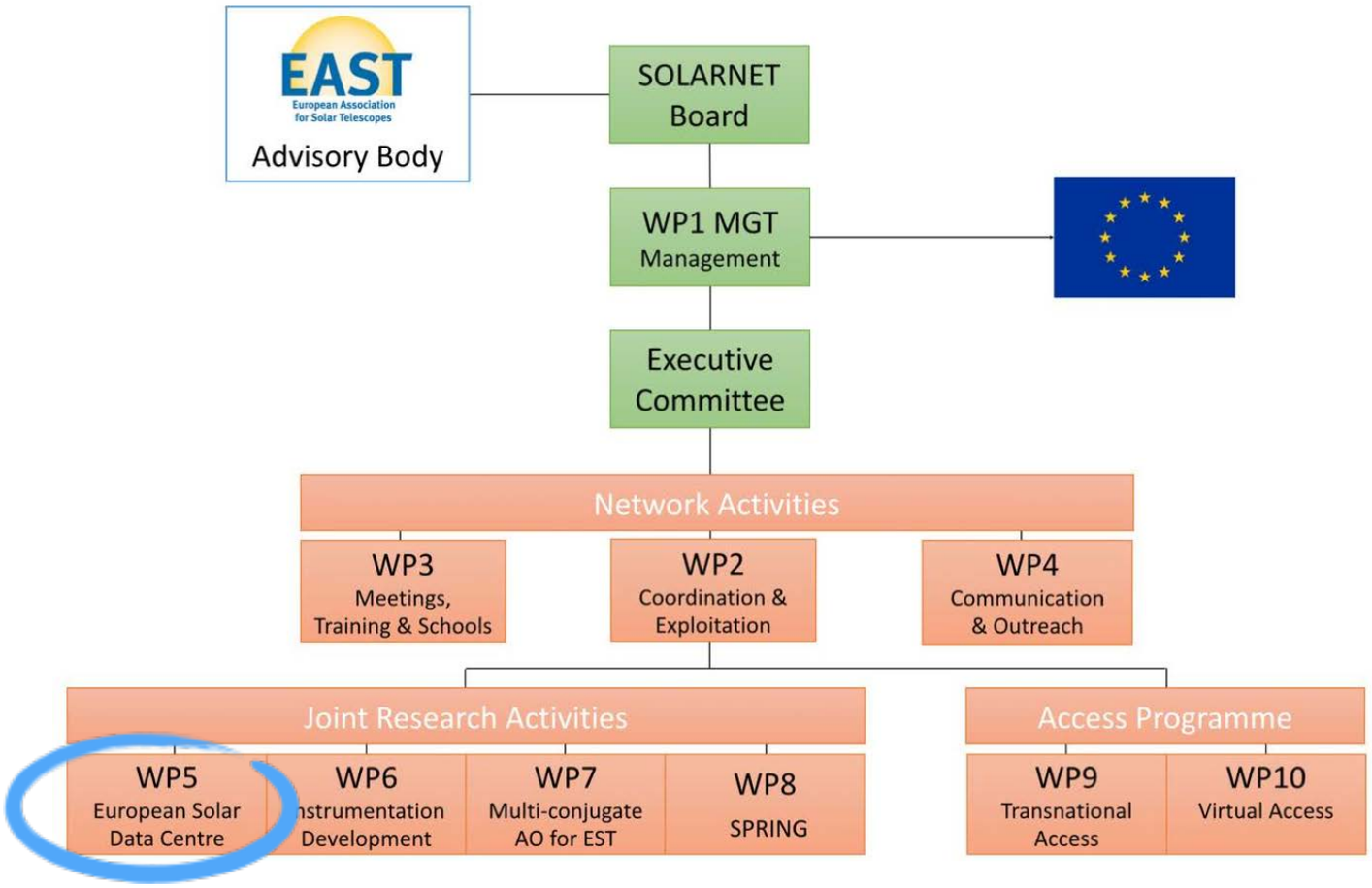
These common efforts can be understood as the first steps towards building up a *European Solar Data Centre*, envisaged for EST.

- Participants: 15 institutes involved (Coordinator: NBG/KIS)
- Funding: 1MEur

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.



Most institutes distribute their WP5 PMs within the SOLARNET running period (4y)



Running activities — 4 main lines of action

1. Data calibration
2. Data archiving
3. Data dissemination
4. High-level data products

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

1. Data calibration

Data pipelines	Aperio/KIS
Real-time image reconstruction	IAC, QUB
Image restoration for EST	SU
Combining *speckle* & MOMFBD	KIS
Slit-spectra reconstruction	MPS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

2. Data archiving

Curation of existing/archived data	INAF, SU, AISAS, HVAR, AIP, KIS, UiO
Solar database mining technologies	UiO, INAF, UCL
Machine learning for identification and classification of solar features	AIP, AISAS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

3. Data dissemination

SVO: From a prototype to an operational VO

ROB

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

4. High-level data products

Visualisation tool for multi-instrument and simulated data	QUB
Basic computational tools	USFD
High level data and analysis tools	USFD, ASU

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

WP5 Deliverables

Deliverable Number[14]	Deliverable Title	Lead beneficiary	Type[15]	Dissemination level[16]	Due Date (in months)[17]
D5.1	Automated data pipeline prototype	1 - KIS	Demonstrator	Public	48
D5.2	Improvements on image reconstruction techniques	4 - SU	Report	Public	48
D5.3	Release of software for reconstruction of slit-spectrograph data	10 - MPG	Other	Public	48
D5.4	Report on Data Curation	18 - HVAR	Report	Public	48
D5.5	Report on solar database mining	3 - UiO	Report	Public	48
D5.6	Assessment report of suitable CNN architectures	11 - AIP	Report	Public	12
D5.7	Feature identification toolkit with CNNs	11 - AIP	Other	Public	48
D5.8	Report on automatic identification of solar features in SCD data	19 - AISAS	Report	Public	36
D5.9	SVO Manual	17 - ORB	Report	Public	36
D5.10	SVO final report	17 - ORB	Report	Public	48
D5.11	Visualisation tool for high-resolution multi-instrument data and simulations	15 - QUB	Other	Public	48
D5.12	Release of Basic Computational Tools and high-level analysis tools	13 - USFD	Other	Public	18

Ready!

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

Institute	Status report
KIS	GRIS & LARS data have been curated and are available in www.sdc.leibniz-kis.de
IAC	Application of NN techniques allows to produce online reconstructed-like data suited for quick-looks
UiO	Stein Haugan is fully employed by ITA but will charge some months to the project. Another position has been announced (for WP10, 2 and 5 for a total of 36 PM) with application deadline October 15. Work has not commenced yet.
SU	In the "Contrast and power spectrum" task of 5.1.4 we will implement an improvement to MFBD-based image restoration methods, making the contrasts in restored data less dependent on the seeing quality and the partial correction already done by adaptive optics. I [M. Löfdahl] have so far concentrated on better characterizing the 85 control modes of the SST AO system and producing an extended set of modes that take into account the correction done to the first 230 Karhunen-Loeve modes that the control modes are based on.
INAF	Started to consider specs of the data archived by our institution
MPG	The work on the software package is ongoing.
AIP	Data curation, ongoing. Survey on suitable NNs performed — Deliverable submitted.
USFD	In progress. Up to date in the framework of EMESSUA code development (Task 3) the following tasks have been completed: FLCT analysis has been performed on an SST/CRISP observational data set at seven different heights (H-alpha, [-1.03A, -0.77A, -0.26A, Ha line centre, +0.26A, +0.77A, +1.03A). Numerical data of 3D magneto convection (MURAM) has been transferred to the USFD storage facility (4TB). Region of interest has been selected, Gamma and LAVD algorithms have been applied for velocity flow identification to be used as initial input for 3D velocity field reconstruction for validation of ML technique to be applied to the observational data (1). Initial work on identification of intensity vortex and magnetic surfaces of associated magnetic flux tubes has been successfully performed. The profiles of magnetic field, current, Lorentz force, density and temperature as a function of radius of magnetic flux tube at different heights (essential for energy calculations) were identified in the numerical magnetoconvection data (MURAM simulations). Algorithm automatically detecting chromospheric spicules and fibrils, coronal loops, fans and other curvilinear features in the solar atmosphere has been developed. It has proved especially successful for providing large-scale statistics on transverse wave motion of these structures. Code (MatLab, input H-alpha observations).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

Institute	Status report
USFD	5.4.2 T1: In progress, just started. 5.4.3 T1: MONAMI v1 ready and available at https://github.com/komabi/MONAMI-Manual. T2: ASDA v2 ready and the source code with updates and documantationa are available at https://github.com/PyDL/ASDA T3: In progress. Up to date in the framework of EMESSUA code development) the following tasks have been completed: FLCT analysis has been performed on an SST/CRISP observational data set at seven different heights (H-alpha, [-1.03A, -0.77A, -0.26A, Ha line centre, +0.26A, +0.77A, +1.03A). Numerical data of 3D magneto-convection (MURAM) has been transferred to the USFD storage facility (4TB). Region of interest has been selected, Gamma and LAVD algorithms have been applied for velocity flow identification to be used as initial input for 3D velocity field reconstruction for validation of ML technique to be applied to the observational data (1). Initial work on identification of intensity vortex and magnetic surfaces of associated magnetic flux tubes has been successfully performed. The profiles of magnetic field, current, Lorentz force, density and temperature as a function of radius of magnetic flux tube at different heights (essential for energy calculations) were identified in the numerical magneto-convection data (MURAM simulations). Algorithm automatically detecting chromospheric spicules and fibrils, coronal loops, fans and other curvilinear features in the solar atmosphere has been developed. It has proved especially successful for providing large-scale statistics on transverse wave motion of these structures. Code (MatLab, input H-alpha observations).
UCL	UCL has been examining the use of capabilities developed under the FP7 HELIO project to provide access to context information that will be needed to facilitate the exploitation of the highly complex observations that will be made by EST and DKIST. While detailed metadata search engines are required when the instruments of an observatory have a wide variety of operating modes and parameters, and a potentially restricted field of view, the same is not necessarily the case for observations that are more synoptic in nature. The HELIO Data Provider Access Service (DPAS) directly searches internet-accessible archives to find what observations are available. The ability to search, for example, for all available H-alpha or EUV observations and return an IDL save-set of the search results and a plot of the coverages of each instrument has been developed and tested; it is in the process of being deployed across the multiple instances of the DPAS in order to ensure a very robust search capability.
QUB	Barebones framework selected developed and tested, however it has been identified that an overhaul may be necessary to adapt the framework to be more object-orientated in order to allow multiple datasets to be visualised efficiently. Imaging data from Hinode/SP and GREGOR/GRIS-IFU can be read in.
ASU	Task 4: We have worked on the update of the KAPPA package to make it compatible with the latest version of the atomic-data database CHIANTI 9.0, released on March 2019. The new CHIANTI version has undergone major changes in the atomic database, structure of data, and in the software itself. These improvements are now being implemented to the new version of KAPPA which should be released next year together with its documentation and manual. Task 5: FLARIX code was further tested for common parameters of the electron beam heating. That included a new height grid to improve stability of the code and computation of Mg II lines. Documentation was extended and updated.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824135.

Institute	Status report
ORB	Added a sample set of data for CHROMIS and CRISP to the SVO. We've had internal discussions on the design of an operational SVO.
HVAR	Preparing correct metadata for FITS headers and exporting observing data into appropriate formats
AISAS	The main goal of this deliverable is to provide a tool for automatic identification of solar features in SCD data taken at the Lomnicky peak observatory.To do it, we proceed in two paralel ways:1) The pipeline for full photometric reduction of the SCD data is developed in order to provide data suitable for automatic detection of features. The staff members of AISAS (J. Rybak, S.J. González-Manrique, P. Gomory) are involved in this topic. Tests on enhancements of the observed structures using image processing tools started in collaboration with AIP (C. Denker).2) We (at AISAS: S.J. González-Manrique) started to investigate possible approaches in aplication of machine learning to automatic identification of solar features in close collaboration with AIP (C. Kuckein, A. Diercke). The first test are performed using full-disk H-alpha solar images from ChroTel. For that, we participatd on visual identification of solar features on about 1000 images. The result was presented at the Machine Learning in Heliophysics conference in Amsterdam (September 2019). The object detection of solar features in H-alpha works but the detection rate needs to be significantly improved (this part is common with AIP) . As a next step, will start to label solar features on images taken with SCD.
Aperio	The pipeline is under development in Python. Image reconstruction is partially implemented.