Introducing the European Fusion Community's (initial) approach to Open Data Access

FAIR for Fusion

Pär Strand

par.strand@chalmers.se

par.strand@chalmers.se

Outline
- Background/Introduction
- Preliminaries
- Fair 4 fusion components
- Summary/Getting involved/collaborating

*NB: this is a very new project - not yet started – and the slides reflect this….

NFRP-2018-11:
Open data access for fusion research

Coordination and support action

The scope covers the comprehensive assessment of open data requirements and issues within the fusion programme, the recommendation of the best technical approaches for providing easy access to data, and the development of support platforms and tools required to implement an open data policy adapted to the needs of the fusion research programme. This could be achieved by pooling the talent and knowledge from other big science programmes and organisations, such as those participating in the EIROforum. The duration of the action will be for 2 years.

The successful conclusion of this action would, in addition to developing the tools and platforms needed for an open data approach, raise the profile and awareness of open data within the fusion programme.

It would also lay the foundations for implementing an open data policy that is well adapted to the needs of the present and especially the future fusion energy research programme, particularly in the run up to the operation of ITER from the middle of the next decade.
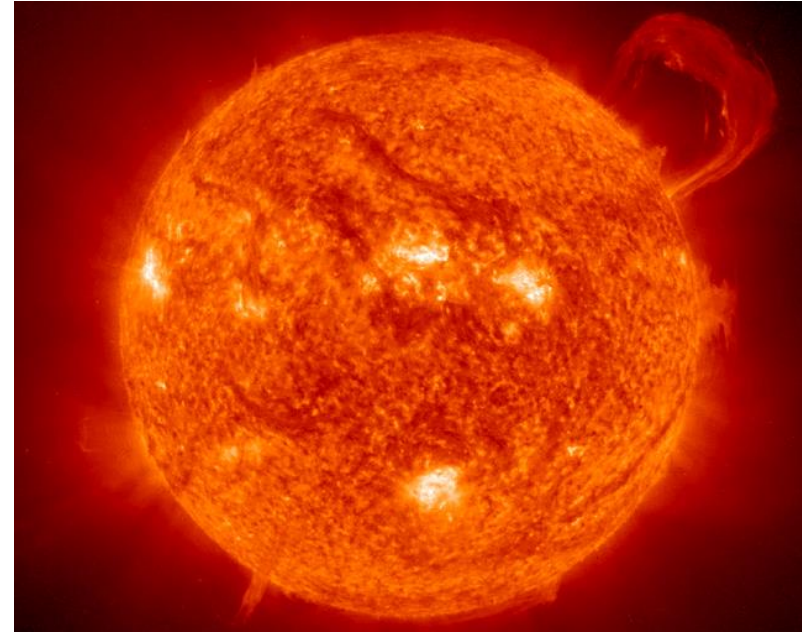
Fair for Fusion: 2 year project starts 1st of september

Background/Introduction

A little bit of a fusion primer…

Fusion

- Energy source for the sun and other stars

- Provides a potential source of base load energy production

- Been working on this for more than 50 years
  - Has turned out to be a somewhat difficult problem

Courtesy of NASA/JPL- Caltech

Fusion

- Energy source for the sun and other stars

- Provides a potential source of base load energy production

- Been working on this for more than 50 years

  - Has turned out to be a somewhat difficult problem
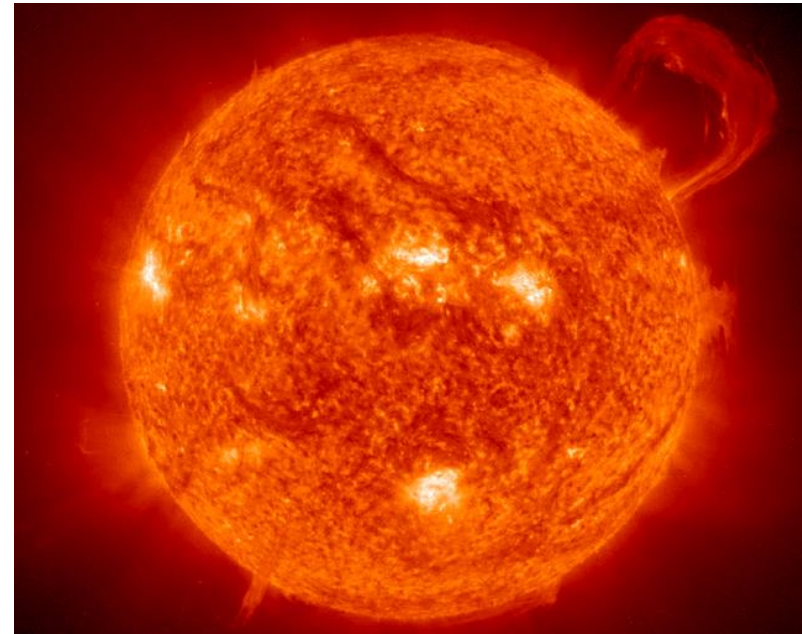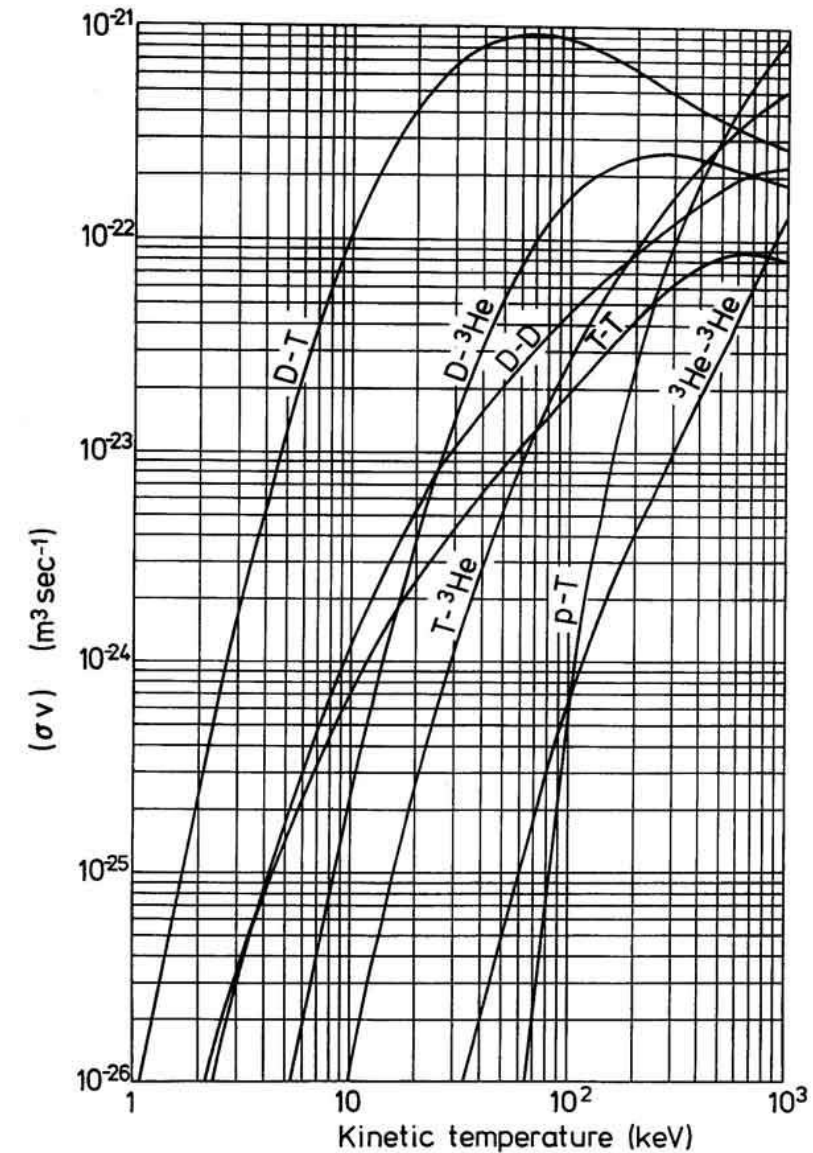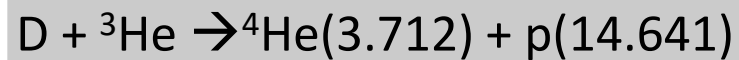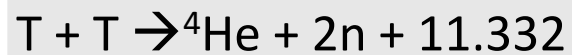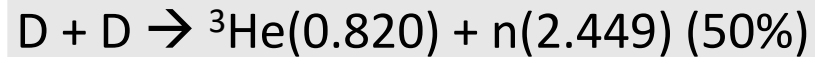

Courtesy of NASA/JPL- Caltech

"Every time you look up at the sky, every one of those points of light is a reminder that fusion power is extractable from hydrogen and other light elements, and it is an everyday reality throughout the Milky Way Galaxy."
--- Carl Sagan, Spitzer Lecture, October 1991

D + T → ⁴He (3.561) + n(14.029)

D + D → T(1.011) + p(3.022) (50%)

D + D → ³He(0.820) + n(2.449) (50%)

p + T → ³He + n − 0.764

T + T →⁴He + 2n + 11.332

D + ³He →⁴He(3.712) + p(14.641)
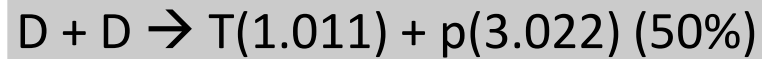
Sun: 10 Million degrees, Fusion: 100 Million degrees (in vicinity of materials)

$_3^6 Li + _0^1 n \rightarrow _2^4 He + _1^3 H + 4.8 MeV$ Allows for breeding tritium in the cooling blanket
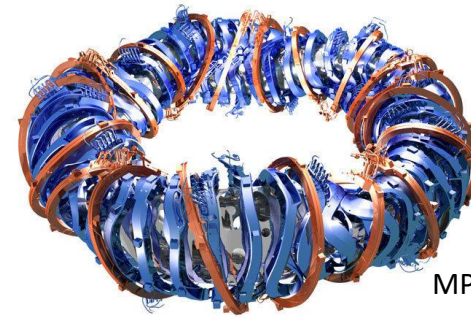
Fusion

- Two main lines of research
  - ~~Inertial confinement~~
    - Implosion of small pellets
    - NIF at LLNL. Laser MégaJoule (CEA)
  - Magnetic confinement
    - Two main type of configurations studied:
      - Stellarator – W7X
        - Recently inaugurated in Greifswald in Germany
        - Steady state device
      - Tokamak – ITER
        - Under construction in Cadarache in France
        - Inductive, pulsed device

LLNL

MPG -IPP

ITER:
5.3T, 15MA, 15keV

50MW in
500 MW out (Q=10)

< 1000s
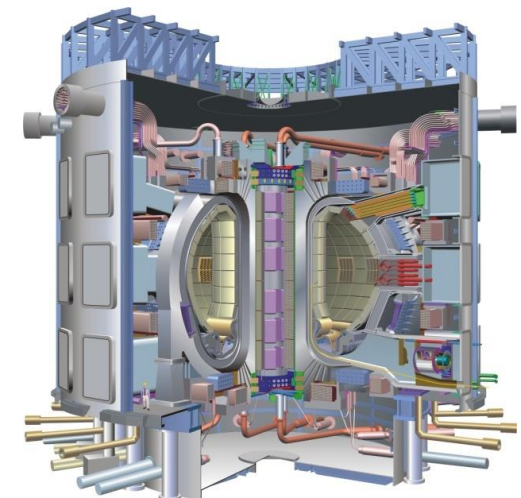
ITER organisation
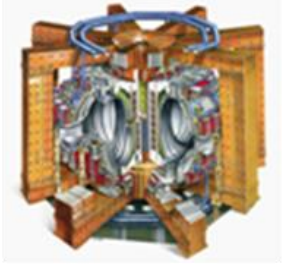
Fair4Fusion

Materials under neutron bombardment

JET

Fusion is feasible

Fusion is practical, attractive

Power Plant

Fusion is plausible

DEMO

Fusion is commercially exploited

JT-60SA

ITER

50 MW

500 MW
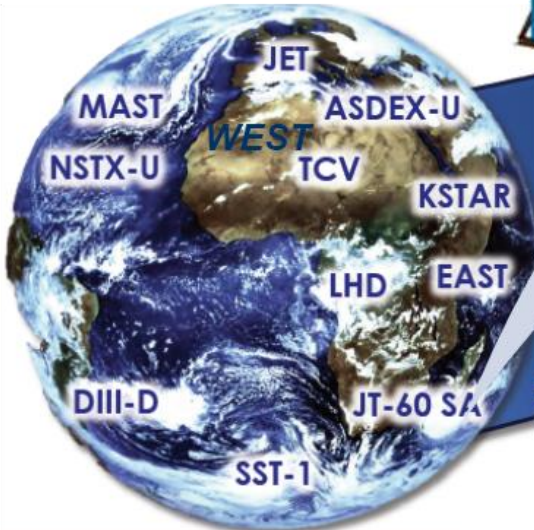
Fusion facilities around the world

MAST
JET
ASDEX-U
WEST
NSTX-U
TCV
KSTAR
LHD
EAST
DIII-D
JT-60 SA
SST-1

~2020
Operation

~2025

~2050

T. Donné, EUROfusion

Fusion

- A clean, sustainable, CO_2 free and virtually inexhaustible energy source.

- Inherently safe and free from proliferation issues.
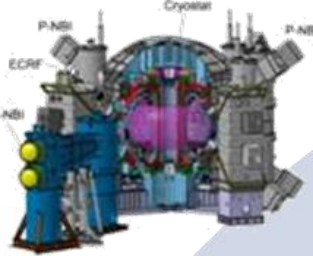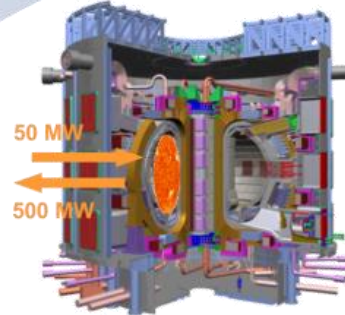
- A "new" energy source adding base load capacity (intermittent and seasonal) to the current energy portfolio!

- Backed by large scale international collaborations(EUROFusion, ITER, …).

ITPA confinement database

Aggregated data volumes are large but not excessive: 10-100s Tb/year for the european devices, ITER likely ~5Pb/yr

Eurofusion – the joint European research activity for fusion in Europe

- Main existing European facilities (JET, MAST, WEST, Asdex, TCV, Wendelstein 7X) are all represented in Fair4Fusion

- Data management and requirements are broadly similar BUT:
  - No standardized formats only a (small) subset of tools are shared over devices
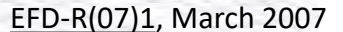 - Interpretation of data depends on mapping dynamic evolution of the plasma and are device specific
  - Only JET is purely community owned experiment (others are nationally funded with partial joint access)
  - Differences in access policies and access tools

Main interface to stakeholders for Fair4Fusion! The experimental data (and modelling in more long term) is what we will open up!

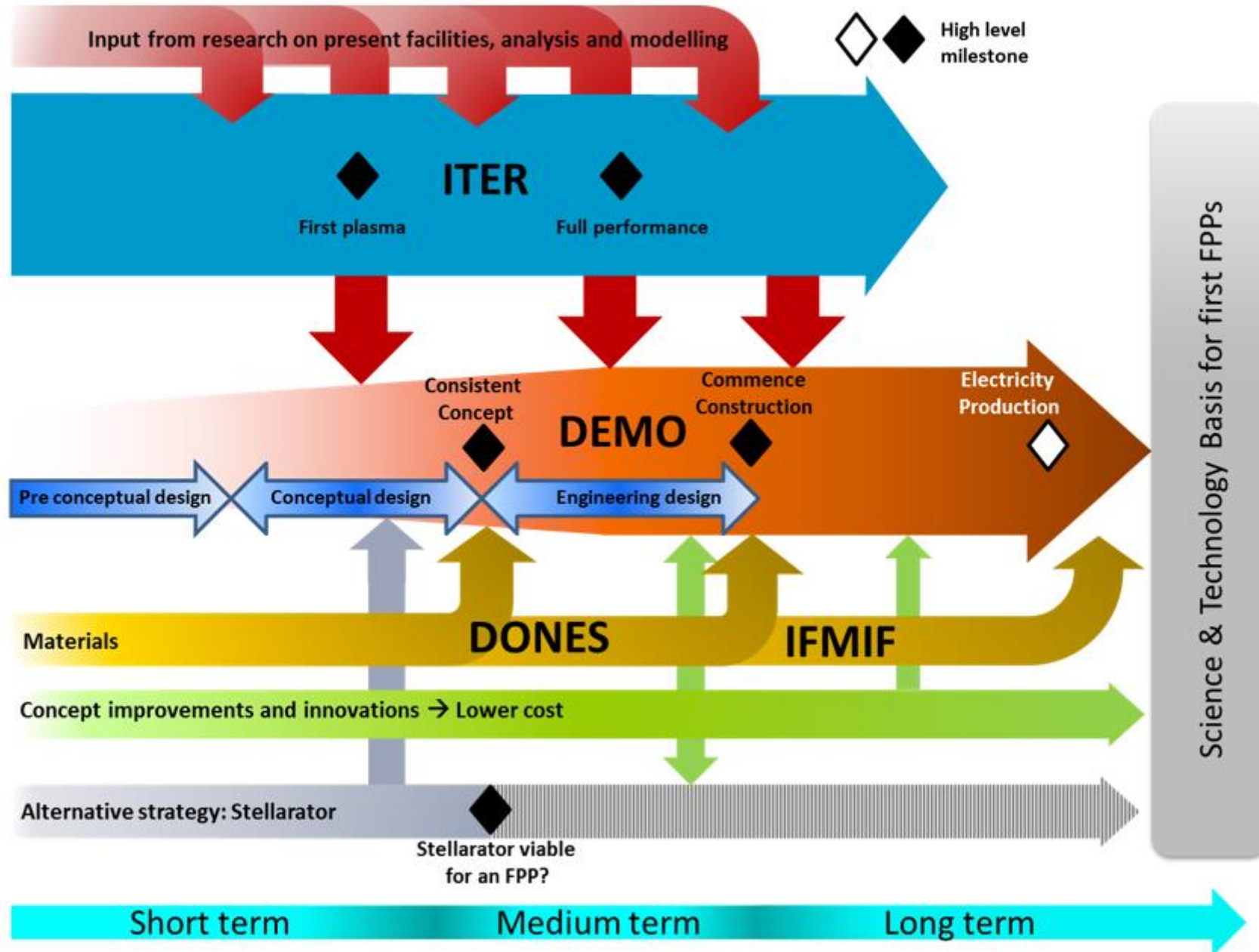Europe has developed a standardized data dictionary and data format for ITER (IMAS IDS). Increasing use also in EUROfusion.

Potential data sharing format!
- Not a formal standard (yet)

ITER will have requirements on open access at least between partners.
- Potential for F4F to provide some solutions

Roadmap for fusion energy

Preliminaries

A short summary of discussions so far in the fusion community

Preliminaries - EUROfusion Open Access WG

Target audience

The main users of the open data access and tools are the active fusion researchers within Europe (= EUROfusion).

- Promote "internal" collaboration

- Build tools facilitating data sharing between experiments and researchers on a higher level than what is now available.

On top of this:

- open data access for the public (grant condition for many research associations within EUROFusion and is expected to become a requirement for the continued EUROfusion engagement in Horizon Europe, detailed scope to be defined).

Tools to extend current capabilities at the different experiments in a non-intrusive way as possible

Data Access Policy - as open as possible - as closed as necessary

- Data is open and publicly accessible

- An embargo period is employed for new data (18 mos, 24 mos??)
  - Embargo is broken when data becomes published.

- Data is released under license agreement (EU PL; Creative Commons,..)
  - Exceptionally, data can be withheld due to specific licensing or security issues

- FAIR based implementation (Findable, Accessible, Interoperable and Reusable/reproducible*) - implications on technology and policies

- Single approach for community and public data but possibly with restrictions on exchanges for the public

- IMAS (ITER) format should be explored for joint data sharing

- Provenance should be a key part of the development

- Applies also to Modelling data - impact on future activities!

- User Interfaces
  - Supported by a EUROfusion wide Authentication & authorization infrastructure (CFP-ADMIN-AWP18-PMU-03)
  - "Portal access" for public and authenticated users
  - API – authenticated users only

*different and demanding set of technical requirements - exact scope need to be defined.

Next steps

- Prepare a first draft version of a Data management plan (WG activity)
  - Iterate with EUROfusion members through GA and consultation bodies
  - Deadline for final agreed version end 2019 (?),

- Initiate and Implement FAIR4FUSION
  - Final blueprint end 2021
  - Demonstrators and technology recommendation during 2019 and 2020
  - Resource envelope for full implementation needed mid term for Horizon Europe proposal for EUROfusion.

Fair 4 Fusion components

A short overview of WPs …

Members and work packages

WP	LEAD	PPM
1. Management with Communication & Outreach	CHALMERS	31
2. Policy and use Case Definition	MPG	24
3. Blueprint Architecture for Fusion Open Data	PSNC	35
4. Data Foundations for Open Access of Fusion Data	UKAEA	42
5. Open Data Demonstrator Development	NCSR-D	61

Chalmers, UKAEA (CCFE), NCSR-D "Demokritos", CEA, PSNC, MPG (IPP), EPFL (SPC)
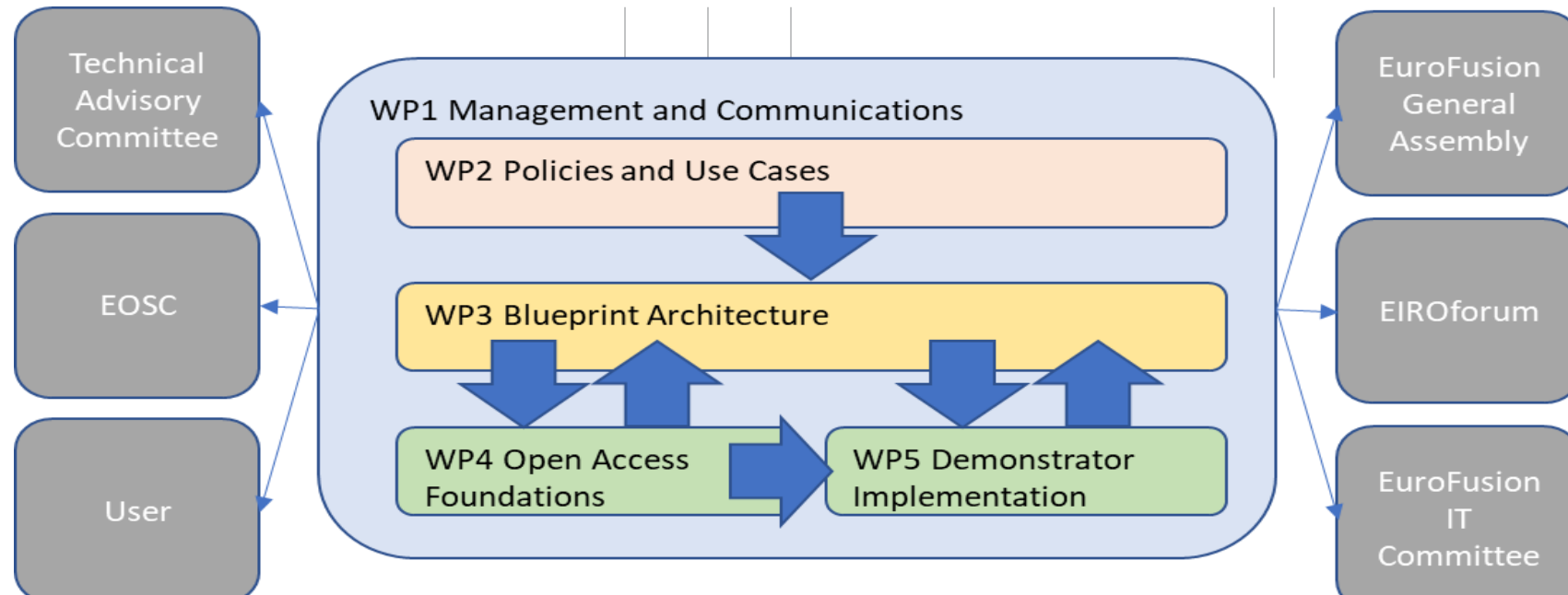
Management, communication and outreach

Specific objectives:

- To maintain the consortium plan and budget to reflect changing needs within the project;

- To provide effective communication channels between partners and between the project and the EC;

- To oversee adherence to a project wide quality plan and risk management policy;

- To facilitate and support the project's governance bodies;

- To manage and coordinate synergies and collaboration with related organisations.

- To organise demonstrations and training sessions as well as to plan and support dissemination activities for the consortium

- To establish the project's online presence.
  - The consortium intends to set up and maintain the project website for the duration of the project plus - at least - two more years. The website will include a public area through which available information will be disseminated (e.g. public deliverables, news, etc.), as well as a private area for distributing restricted information amongst consortium members and supporting project management activities.

- The consortium will also utilise social media for information dissemination (e.g. Facebook, LinkedIn, Twitter) to the public. As far as it concerns the research community, the consortium will create a ResearchGate account, while it will consider using YouTube as a means for distributing multimedia material such as videos regarding the developed open data demonstrator.

Policy and use Case Definition

Specific objectives:

- Define access policies for different candidate end users (include any legal or ethical restrictions);
  - Access mechanism for fusion users, for fusion related engineering communities and general public
- Policies for access to different 'levels' of data - raw, calibrated, science, modelling, engineering
- Define license under which data is released;
- Define policies on when and where data can be replicated;
- Define support mechanism;
- Understand policies in relation to national policies on open data, commercialisation of results and international treaties (including Euratom and GDPR). Link to EF Open Access WG. This task will consider feedback and recommendations coming from this group.
- User acceptance will be in coordination with WP5 (demonstrator), while feedback from the wider fusion community will be collected as part of events/workshops etc.

Blueprint Architecture for Fusion Open Data

Specific Objectives

- Fusion Open Data Platform requirements specification

- Current technology survey, summarizing the current status of the technologies and standards used inside Fusion community and experiments, and looking into latest status and maturity of the other big science programmes (e.g. EIROforum members, etc) and ICT technologies focused activities(including EOSC ecosystem)

- Propose the detailed architecture for the prototype evaluation → WP5

- Propose long term Architecture Blueprint for Fusion Open Data Framework → EUROfusion Horizon Europe

Data Foundations for Open Access of Fusion Data

The work package looks at the well-known fundamentals for open data, and the larger open science initiative, in particular concentrating on

- metadata,

- persistent identifiers,

- authorisation,

- access harmonisation and

- provenance.

We will define solutions based on existing (or emerging) standards and recommendations both internally from the fusion community and from external projects and organisations such as EOS, EIROforum and RDA Europe.

Implementation plans will be provided for the blueprint and (in part) tested in the demonstrator project.

Open Data Demonstrator Development

Specific Objectives:

- Customise the selected platforms and technologies to address the requirements set out by the Fusion research community;

- Implement authentication and authorisation mechanisms in order for the platform to be able to interface with the data providing sites;

- Deploy and test the selected technologies and platforms;

This work package will deliver a customised, integrated and tested container-based platform for the efficient, development-friendly management and processing of fusion data. The platform will include mature, widely-used, open-source and in-house software depending on the policy and data management requirements which will be dictated by WP2, 3 and 4.
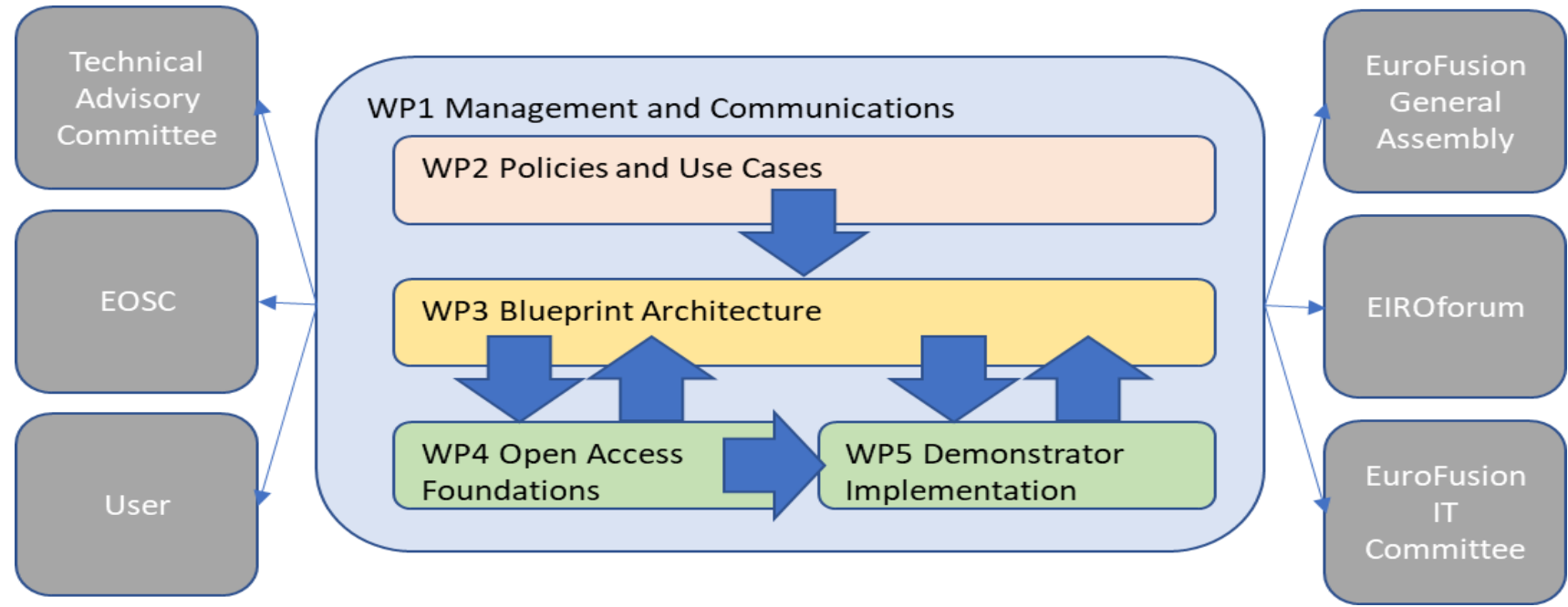
Summary

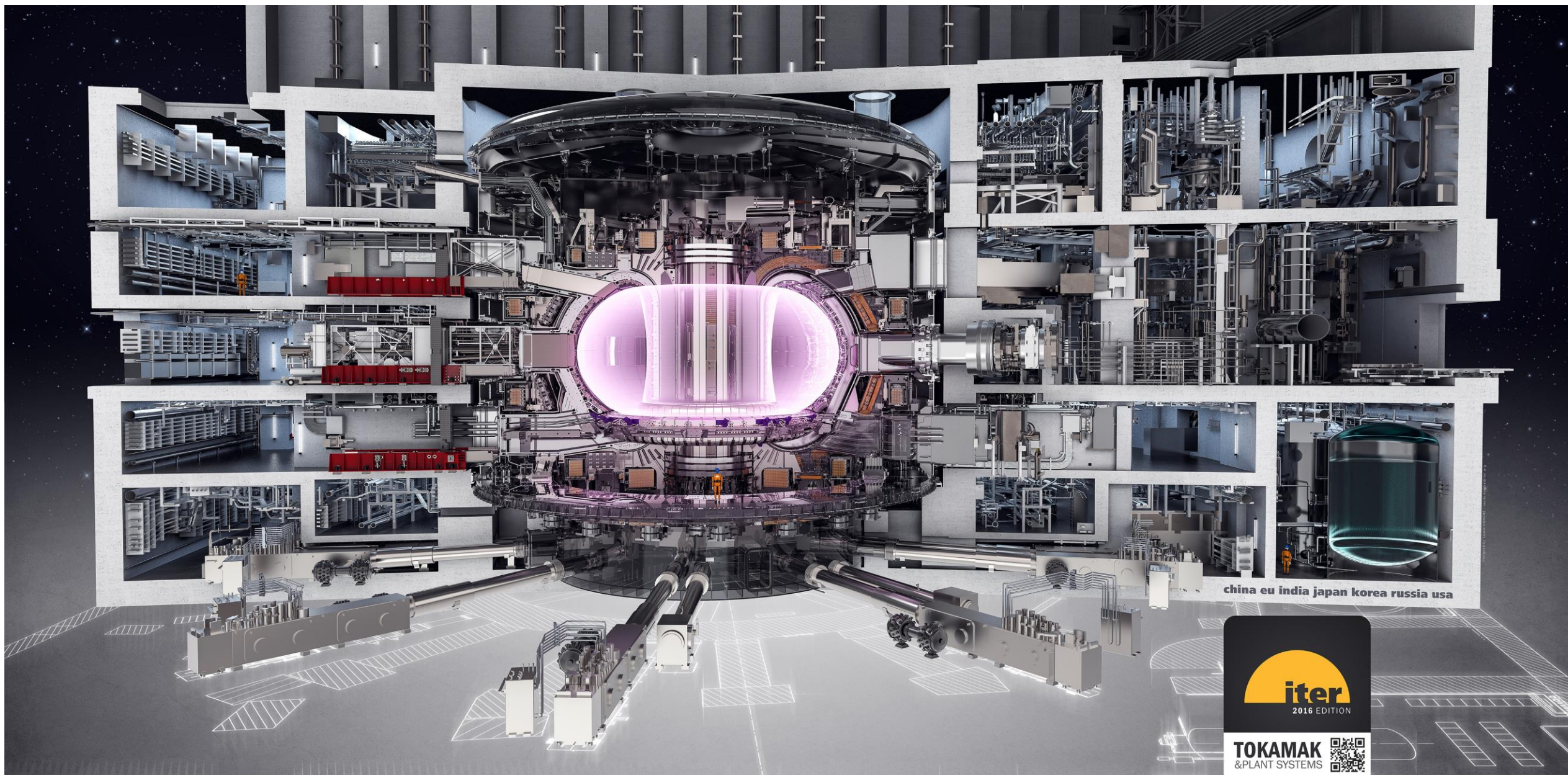As we are starting up there are little in terms of conclusions…

Summary

Fair for Fusion starts september 1st aiming to:

- Maximise the benefit of open [fusion] data.
- List the necessary policies and restrictions for opening fusion data to the research community and the public.
- Define an architectural blueprint for a platform delivering open fusion data, <u>in collaboration with external bodies</u>
- Develop a suite of tools to support open access of fusion data
- Develop an open data demonstrator for use within the fusion community to promote the opportunities and ease of use of such a platform

We will seek collaboration/exchanges with relevant partners and bodies

- Technical Advisory Committe (TAC) one venue: feel that you can contribute → make contact with me! Some prospective technologies/approaches already in proposal but very open for suggestions!
- there will be a web presence … soon: www.fair4fusion.eu

china eu india japan korea russia usa

iter 2016 EDITION

TOKAMAK &PLANT SYSTEMS