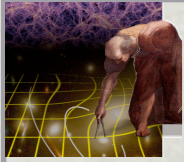# The Euclid Science Ground Segment Objectives
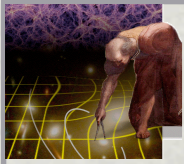
Euclid France - 2012

M. Sauvage, Ground Segment Scientist

Mission statement

- **Defined essentially in the Science Management Plan**
- **Detailed during interactions with the Euclid "stakeholders" (EC scientists, Instrument Dev. Teams, EST...)**
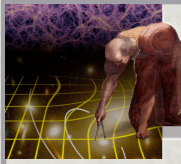
Mission statement

- **Defined essentially in the Science Management Plan**
- **Detailed during interactions with the Euclid "stakeholders" (EC scientists, Instrument Dev. Teams, EST...)**

  ✸ Provide science ready data products:

  - ✸ All calibrations performed, conversion to astronomical units (but no or little conversion to absolute units), "objects" (galaxies, stars) magnitudes and spectra extracted and catalogued, instrumental properties measured (e.g. PSF).

  - ✸ Science "interpretation" of the data is the task of the Science Working Groups. In some cases, it may be performed using the SGS systems.
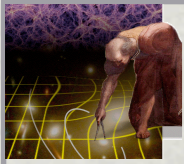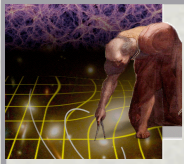
Mission statement

- **Defined essentially in the Science Management Plan**
- **Detailed during interactions with the Euclid "stakeholders" (EC scientists, Instrument Dev. Teams, EST...)**

  ✳ Provide science ready data products:

  - ✳ All calibrations performed, conversion to astronomical units (but no or little conversion to absolute units), "objects" (galaxies, stars) magnitudes and spectra extracted and catalogued, instrumental properties measured (e.g. PSF).

  - ✳ Science "interpretation" of the data is the task of the Science Working Groups. In some cases, it may be performed using the SGS systems.

  ✳ Assist SOC in the daily operations of the spacecraft:

  - ✳ Monitoring the health of the instrument on a daily basis.

  - ✳ Performing trend analysis and feeding back results to EST, Calibration, and Survey working groups for mission planning.

  - ✳ Overseeing the survey completion (e.g. assessing success/failure of tile observations) and feeding back results to EST, and Survey Working group.
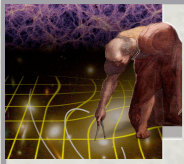
Mission statement

- **Defined essentially in the Science Management Plan**
- **Detailed during interactions with the Euclid "stakeholders" (EC scientists, Instrument Dev. Teams, EST...)**

  ✳ Provide science ready data products:

  - ✳ All calibrations performed, conversion to astronomical units (but no or little conversion to absolute units), "objects" (galaxies, stars) magnitudes and spectra extracted and catalogued, instrumental properties measured (e.g. PSF).

  - ✳ Science "interpretation" of the data is the task of the Science Working Groups. In some cases, it may be performed using the SGS systems.

  ✳ Assist SOC in the daily operations of the spacecraft:

  - ✳ Monitoring the health of the instrument on a daily basis.

  - ✳ Performing trend analysis and feeding back results to EST, Calibration, and Survey working groups for mission planning.

  - ✳ Overseeing the survey completion (e.g. assessing success/failure of tile observations) and feeding back results to EST, and Survey Working group.

  Classic role for a ground segment

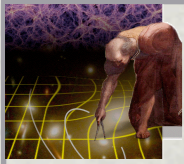# Mission statement (non classic)

The Euclid SGS will also process/ingest ground-based data, as well as generate (and process) simulated data.

Mission statement (non classic)

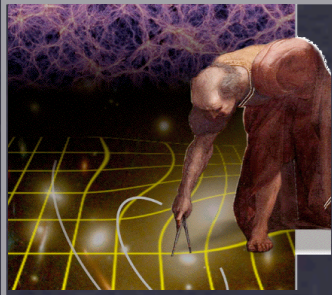**The Euclid SGS will also process/ingest ground-based data, as well as generate (and process) simulated data.**

* For the WL probe we need to study evolution of the signal with redshift.

  * On board we only have 4 bands for the $1.5 \, 10^9$ galaxies in the WL catalog.

* We rely on external surveys to provide the extra coverage for photometric redshift estimation.

* We cannot simply use their catalogs because we want to ensure the highest homogeneity of the Euclid dataset.

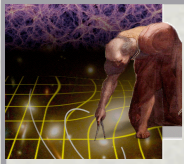 * Astrometric reference, photometric methods, PSF-dependency...

Mission statement (non classic)

The Euclid SGS will also process/ingest ground-based data, as well as generate (and process) simulated data.

* For the WL probe we need to study evolution of the signal with redshift.

    * On board we only have 4 bands for the 1.5×10^9 galaxies in the WL catalog.

* We rely on external surveys to provide the extra coverage for photometric redshift estimation.

* We cannot simply use their catalogs because we want to ensure the highest homogeneity of the Euclid dataset.

    * Astrometric reference, photometric methods, PSF-dependency...

* The main purpose of Euclid is to validate/falsify current cosmological models and fundamental physics theories.

* We cannot rely only on our data processing expertise to cancel every source of bias.

* Simulated surveys will be a key tool for the scientific analysis of the mission.

    * Requires instruments/mission simulators.

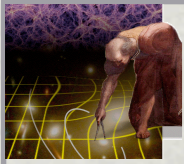   * Provides a capacity to make many realizations of the "Euclid experience" to track biases.

Structure and Position of the SGS in the Euclid

The SGS structure and teams

* The SGS is comprised of:

  * A Project Office:

    * F. Pasian, Manager - Ch Dabin, System Lead - M. Sauvage, SGS Scientist - O. Mansutti, Configuration Lead, C. Vuerli - PA/QA lead - A. Gregorio, IOT/PO coordination.

  * A System team:

    * Identifies and develops all services needed for the proper execution of the SGS tasks.

    * Proposes common tools for the SGS developments.

    * Studies the scenarios for the actual implementation of the SGS pipeline.

  * A series of Organization Units (10):

    * In charge of prototyping the different pipeline elements.

    * Structured along the different stages of the pipeline (but see later).

  * A series of National Science Data Centers

    * Providing developer expertise to implement the prototypes.

    * Providing computing and storage infrastructure to run the pipeline.

    * In close interaction with the System Team.

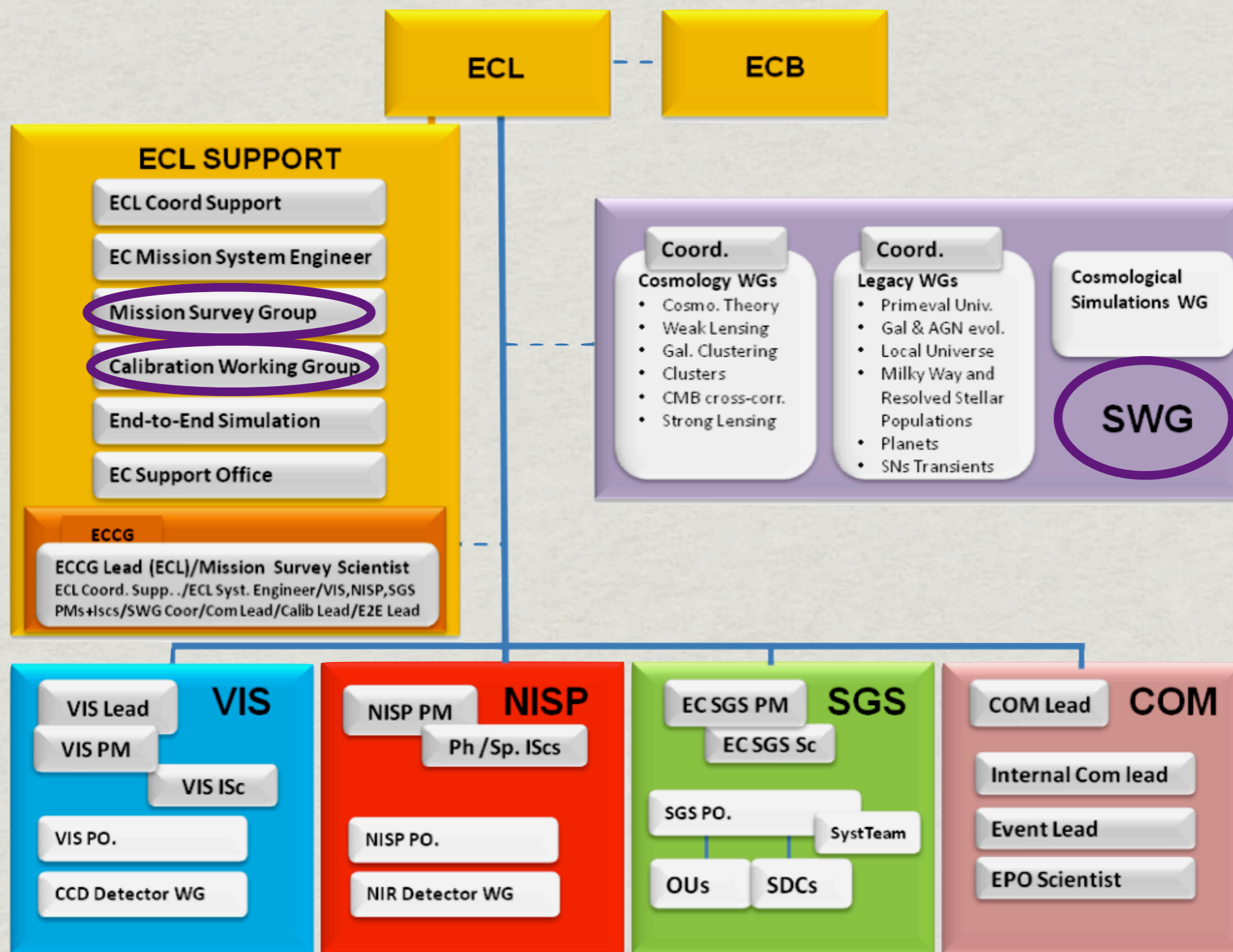* The SGS Organization Group gathers the PO, the OU and SDC leads, and deals with all development issues for the SGS.
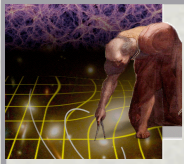
The SGS environment - EC
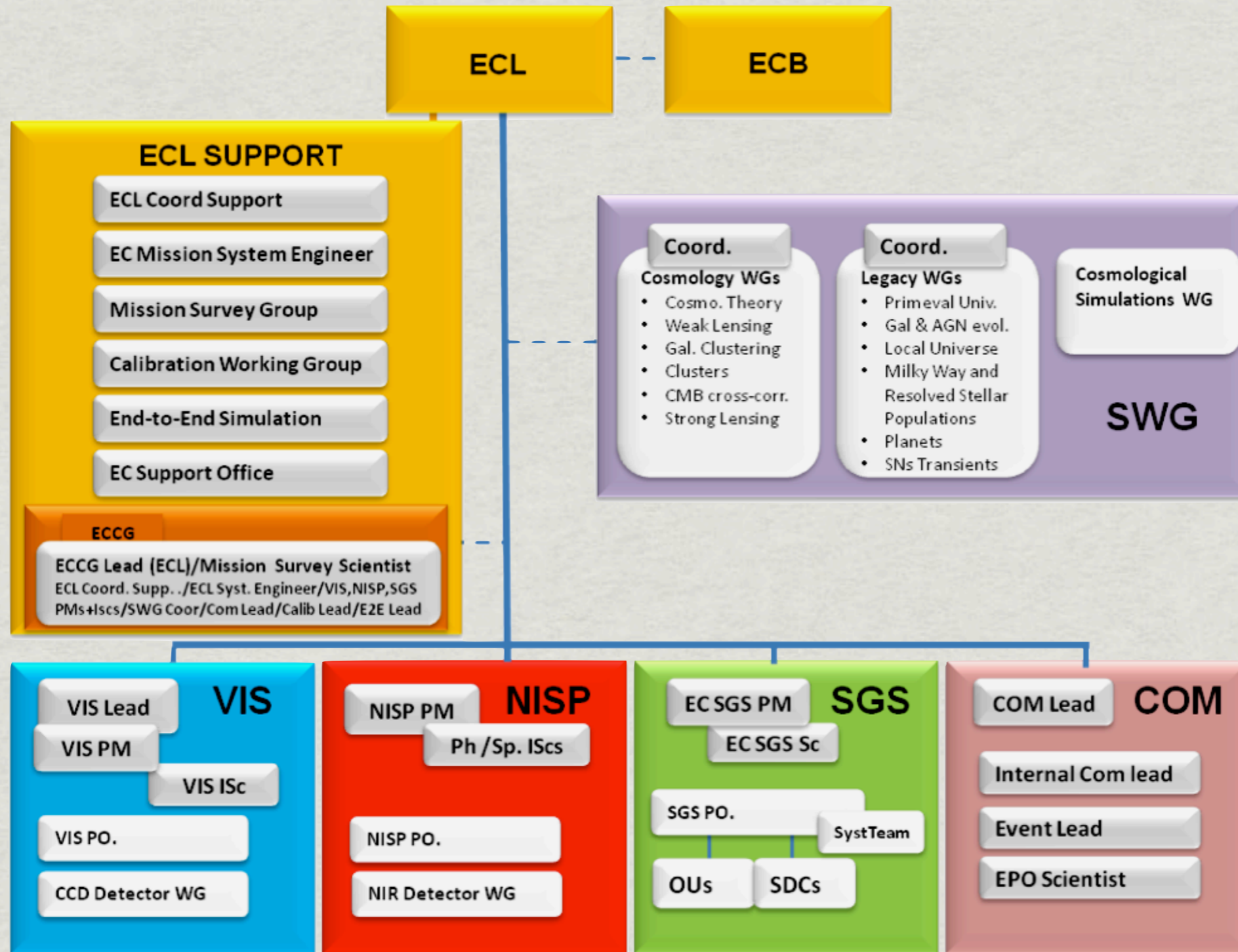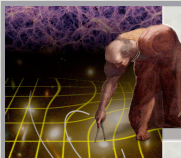
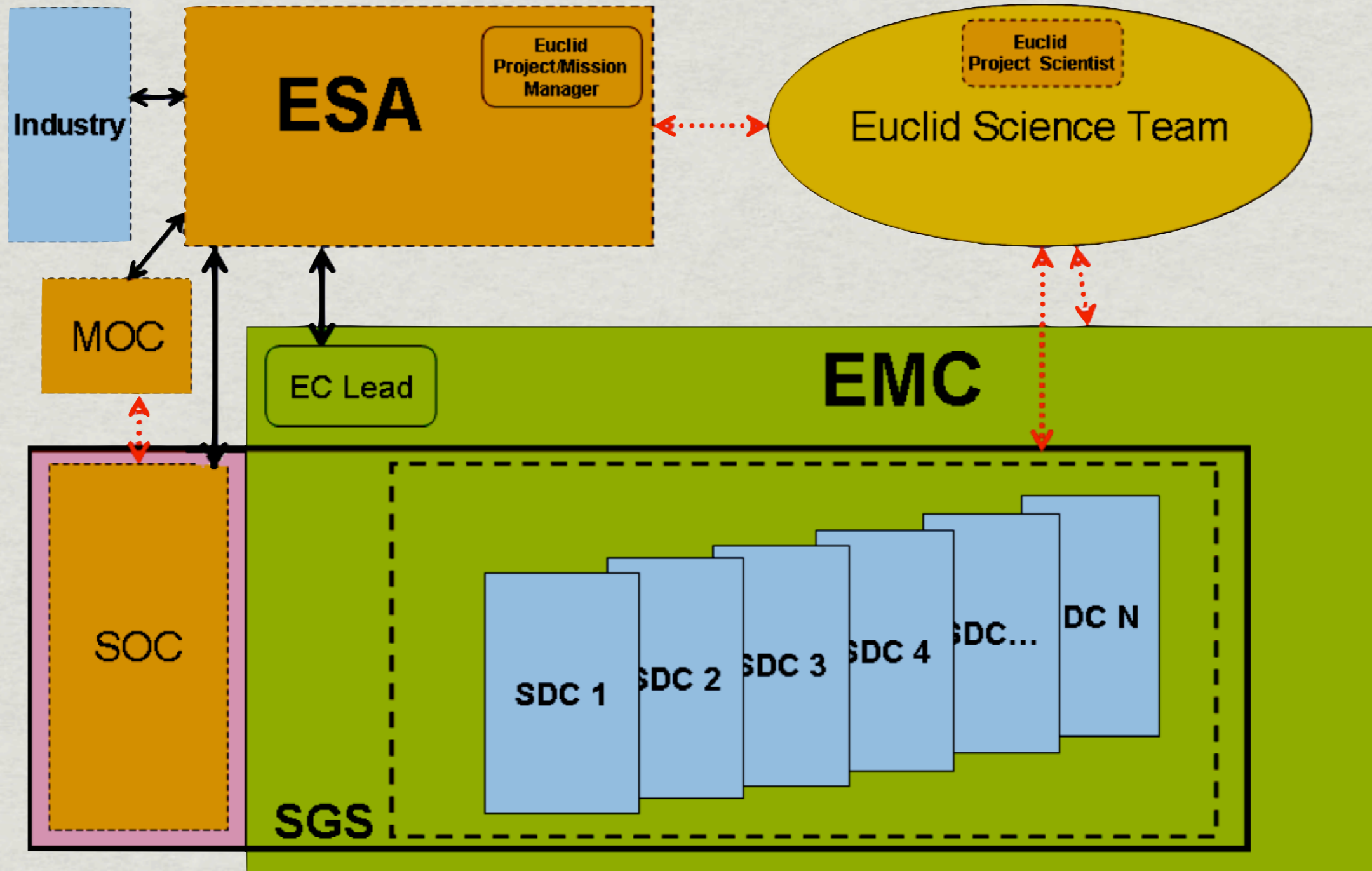**From the EC management plan**

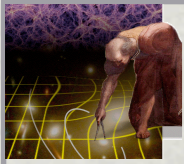# The SGS environment - EC

From the EC management plan

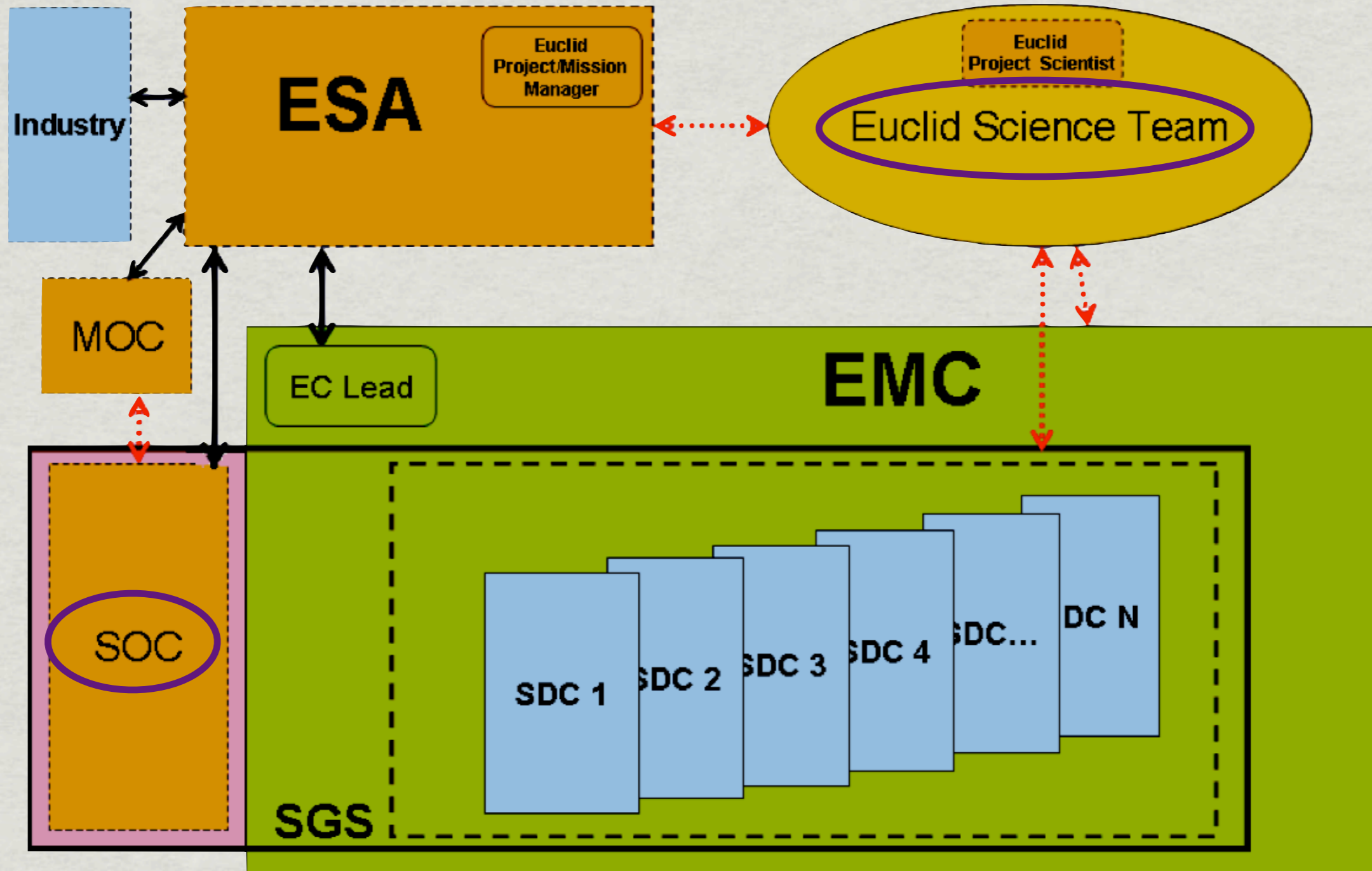# The SGS environment - ESA

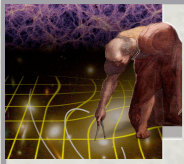## From the Science Management Plan

The SGS environment - ESA

From the Science Management Plan

Interactions - SOC

* For some aspects, the SOC is fully integrated in the SGS:

    * SOC executes the first stage of data processing (telemetry reformatting and association).

    * SOC participates in the design and realization of the SGS systems and architecture.

* W.r.t the SOC, the EC SGS is responsible for:

    * Delivering the "Level 1 pipeline" (i.e. from raw telemetry packets to raw observation "files").

    * Delivering elements to build the Quick-Release data products ("Level Q").

    * Defining the operations boundaries of the instruments.

    * Monitoring instrument health and reporting.

Interactions - PS+EST

* Essentially, the EST's role is to maximize the scientific output of the mission, w.r.t. the astronomical community.

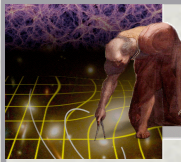* For this SGS this means setting a number of rules:

  * Defining which data products we are supposed to create and release.

  * Defining quality objectives for data products.
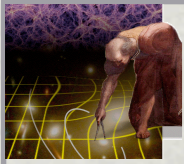
  * Defining data release schedule.

Next stop for the SGS: Preliminary Requirements Review

Reviews

Review	Subjects and Result	SGS model - Pipeline maturity status
July 2013: **PRR** Preliminary Requirements R.	Validate: schedule, management plans, organization, process and architecture principles, Pipeline scenario and full-fledged Data and Processing Flow document **Verification of the Reqs on the SGS**	Internal interfaces
Nov 2014: **SRR** System Requirements R.	External interfaces and pipeline interfaces **Approve the SGS requirements**	Prototypes
May 2016: **PDR** Preliminary Design R.	Verification of the preliminary design: understanding what is needed for end-to-end data processing, ability to cope with simplified (simulated) data. **Performance assessment**	Breadboard model
Nov 2017: **CDR** Critical Design R.	Compatibility tests status and reports and external interfaces **Final design and infrastructure**	Demonstration model
May 2019: **QR** Qualification R.	Integration of pipelines on SOC/SDCs infrastructures Readiness of operational interfaces **End-to-End test status and reports**	Operational model 1
Feb 2020: **ORR** Operational Readiness R.	Ensure full SGS readiness for in-orbit operations **Authorize utilization for space segment in-orbit operations**	Operational model 2
Jan 2021: **D**ata **P**rocessing **R**eadiness **R**.	**Ensure the SGS system readiness for public data release**	=

The objectives

* In short:

  * Validate schedule, management plans, organization, process and architecture principles.
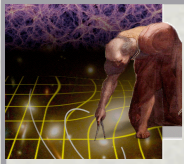
  * Verification of SGS requirements.

* In more explicit terms, we need to achieve the following milestones:

  * Release preliminary management, engineering and product assurance plans.

  * Release technical requirement specifications.

  * Identify the system and operations concepts and confirm their technical and programmatic feasibility.

Detailed OU perspective

Technical requirements specification

* GDPRD and SIRD flow down, associated status and traceability matrix:

    * Will be covered in the data processing flow document (see later).

* Preliminary SGS requirements definition:

    * All requirements placed *on* the SGS lead in turn to requirements placed by some units of the SGS onto other units of the SGS.

        * At the PRR we need to demonstrate compliance w.r.t. *external requirements*, and provide *preliminary* definitions for the *internal requirements*.

    * In an OU-specific perspective:

        * Data processing, performance and infrastructure requirements.

        * Archiving requirements.

        * Simulation data production requirements and sizing.

        * External data production requirements and sizing (other than those already contained in the GDPRD).

Data processing flow document

❊ For the Preliminary Requirements Review we need to demonstrate complete mapping between our external requirements and our organization.

❊ To come up with a development plan, we need to establish a global overview of the data processing activities of the SGS (the "big picture").

  ❊ The complete Euclid pipeline is *definitely not* a chain of OU-developed pipelines...

❊ This will be integrated in a Data Processing Flow document

  ❊ Part 1: for WL and GC linking requirements to data processing actions

  ❊ Part 2: Organization of the data processing actions into a global flow.

❊ Follow this on Redmine:

  ❊ http://euclid.roe.ac.uk/projects/data-quality-toolt?jump=welcome
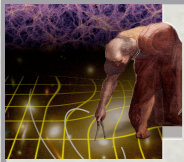
Overall open issues
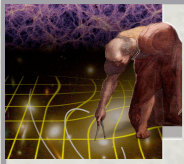
* Two models for the pipeline implementation:

  * Distribute by functions: allows optimal specialization hardware/software, generate heavy traffic.

  * Distribute by sky areas: reduced and possibly optimal traffic, idle time for some centers while survey progresses, not optimal with respect to cadencing of pipeline stages.

~2 Pb

Feedback loops
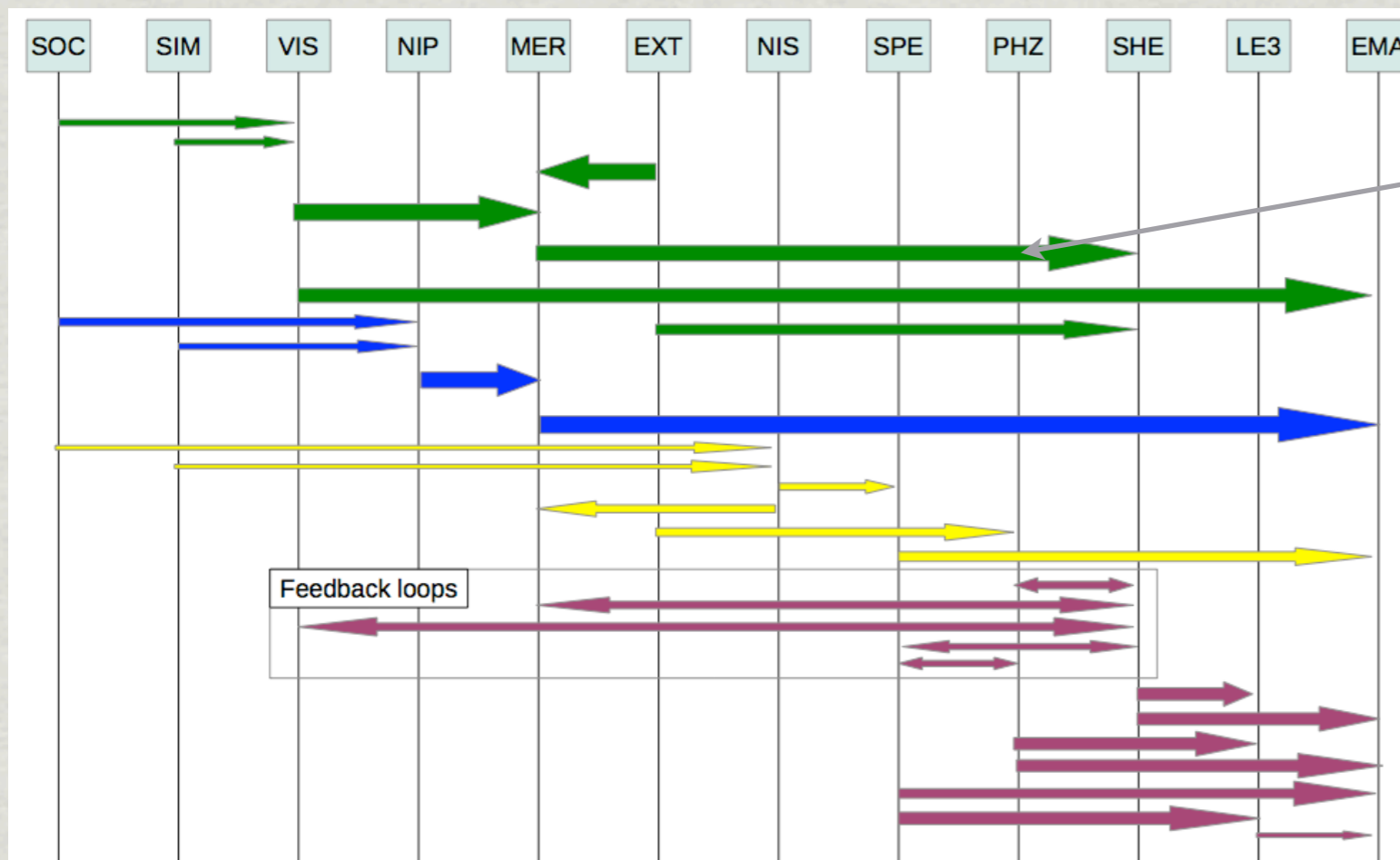
} PDFs can blow up catalog size "infinitely"

Study by K. Noddle from the System Team

Overall open issues

✳ Cadence is highly variable at different stages of the pipeline:

 ✳ At the first stages, high cadence is needed (e.g. transients detection, survey execution monitoring, health monitoring).

 ✳ At late stages, high cadence does not make sense (e.g. machine learning systems for photo-z prefer large samples, correlation functions do not need recomputing unless volume has increased substantially).

✳ Data product levels are broadly defined in the SMP (from 1-raw to 3-science ready).

 ✳ We need to map those categories onto the actual products of the pipeline.

✳ Data releases:

 ✳ What data products do we release and following which schedule?

 ✳ We need a working definition of level Q, the Quick-Release data product.

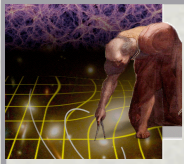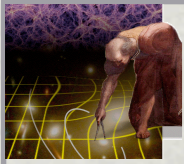 ✳ To be defined shortly by the EST (proposals are on the table).
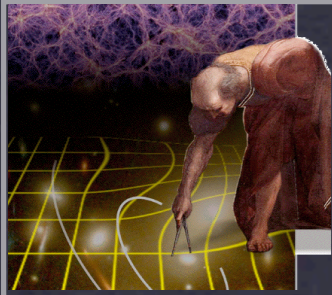
OU-EXT & OU-MER
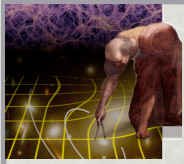
OU-EXT role

* To bring into the Euclid archive system, the data that are essential for the success of the primary cosmological probes.

    * g,r,i,z photometry for the whole survey area, at a depth that allows evaluation of a phot-z for all the objects of the survey.

        * More galaxies than strictly the WL sample.

        * Stars to control the SED effects in the transfer of the PSF model to galaxies.

        * In volume this is the largest input to the system (possibly larger than the Euclid data themselves...).

    * Astrometric catalog (most likely Gaia) to tie all the imaging data together (VIS, NISP, "EXT").

        * Possibly we shall use as well the spectro-photometric data generated by Gaia (for the PSF model).

    * Samples of galaxy spectra to calibrate the WL tomographic bins.

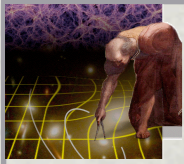   * Which sample is still a highly debated question.

OU-EXT activities

* They are driven by the photometric data.

* The baseline is that the EXT data is reduced by other consortia such that the EXT task is to make sure that the data conform to the requirements:

  * 1% relative photometric accuracy across the whole survey

  * 0.02% color "calibration" drift across smallest spatial scale of PSF model (1 fov).

  * At minimum, OU-EXT extracts the photometry from the coadded survey images.

* Open questions for EXT data and pipeline strategy:

  * Optimal detection/photometry on single epochs or coadded frames?

  * Representation of the EXT photometric data in the simulation branch?

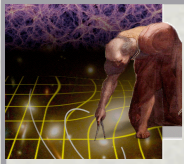 * Size of the spectroscopic sample for WL tomographic bin redshift calibration?

OU-EXT providers

- Only the Southern sky is covered:

  - DES (5000 deg^2) possible extension (e-DES) of 2500 deg^2.

  - KIDS/VIKINGS 1500 deg^2.

- For the northern sky, a number of options are being discussed:

  - Pan-STARRS (1,2) requires extra funding to be secured (possibly by Canada)

  - Subaru/HyperSuprimeCam requires agreement with Japan, 6000 deg^2 EC + 1500 deg^2 Subaru.

- Other options for northern sky include WHT and CFHT but are much less advanced.

OU-MER role

* Performs the difficult task of creating the object catalog from the different survey instrument (VIS, NISP-imaging, NISP-spectroscopy, EXT-imaging, *EXT-spectroscopy*).

* Can be seen as the main hub for the SGS pipeline:

  * photo-z can only be derived from merged ground and space based photometry.

  * PSF model requires a catalog of bona-fide stars with SED in the VIS band.

  * Galaxy shapes need to be "corrected" for the target SED.

  * Spectrum extraction will require object position (hence detection).

  * Spectrometric redshift measurements will often require photometric redshifts as a constraint.

OU-MER activities today

* Cataloguing working group initiated:

  * Studies the cataloguing options taking into account:

    * Requirements.

    * Different strategies coming from different groups (e.g. WL vs. Legacy).

    * Know issues and techniques (different PSFs, variable PSFs).

    * Photometric strategies depending on objectives.

    * ....

  * Follow project on Redmine

    * Led by A. Fontana and M. Sauvage