



ASTRON

Netherlands Institute for Radio Astronomy

LTA Development Meeting
Welcome



Logistics

- Wifi: eduroam, or wiguest.
- Teams: remote participants, and presenting. Details on Indico page (<https://indico.astron.nl/event/380/>).
- Lunches, coffee breaks, etc — right outside this room.
- WC, coffee machine, etc — round the corner
- Is everybody ok with photos?
- Who to contact if you need something: me, or is-secretariat@astron.nl.

Hotel

- Suggested hotel (Fletcher Landhotel De Borken) is ~20 minutes walk away.
- If you came direct to ASTRON by taxi and you need help with transport to the hotel, please let me know and we'll arrange a taxi and/or carpooling.

Dinner (Thursday evening)

- 18:00 at Onder de Eiken (<https://onderdeeiken.nl/>, Drift 20a, 7991AB Dwingeloo)
 - 30 minute walk from the hotel
- There will be a fixed price menu, and you can order your own drinks on our tab. Please be reasonable and don't bankrupt us! 🍺
- We can help arrange taxis / carpooling — please let me know if you need transport.
- *Show of hands so we can confirm numbers.*

Why are we here?

- *Social and technical* reasons.
- *Technical:*
 - The LOFAR LTA is already among the largest and most complex astronomical data archives. And it's getting larger and more complex, and user needs are growing. We need to start converging on how to manage those challenges.
- *Social:*
 - The best way to manage a challenge of this sort is by building a mutual understanding and rapport. We're here to get to know each other and understand who we are working with in the future.

Structure & Agenda

- <https://indico.astron.nl/event/380/>
- Present & future of LOFAR as an instrument.
- Current status of LTA sites.
- Ideas about what comes next.
- Learning from other facilities — Rubin, SKA.
- Informal; lots of time for discussion.
- *Note that this is an open (ish) meeting.*

Who are we?

- Table round of brief (~30 second) introductions.

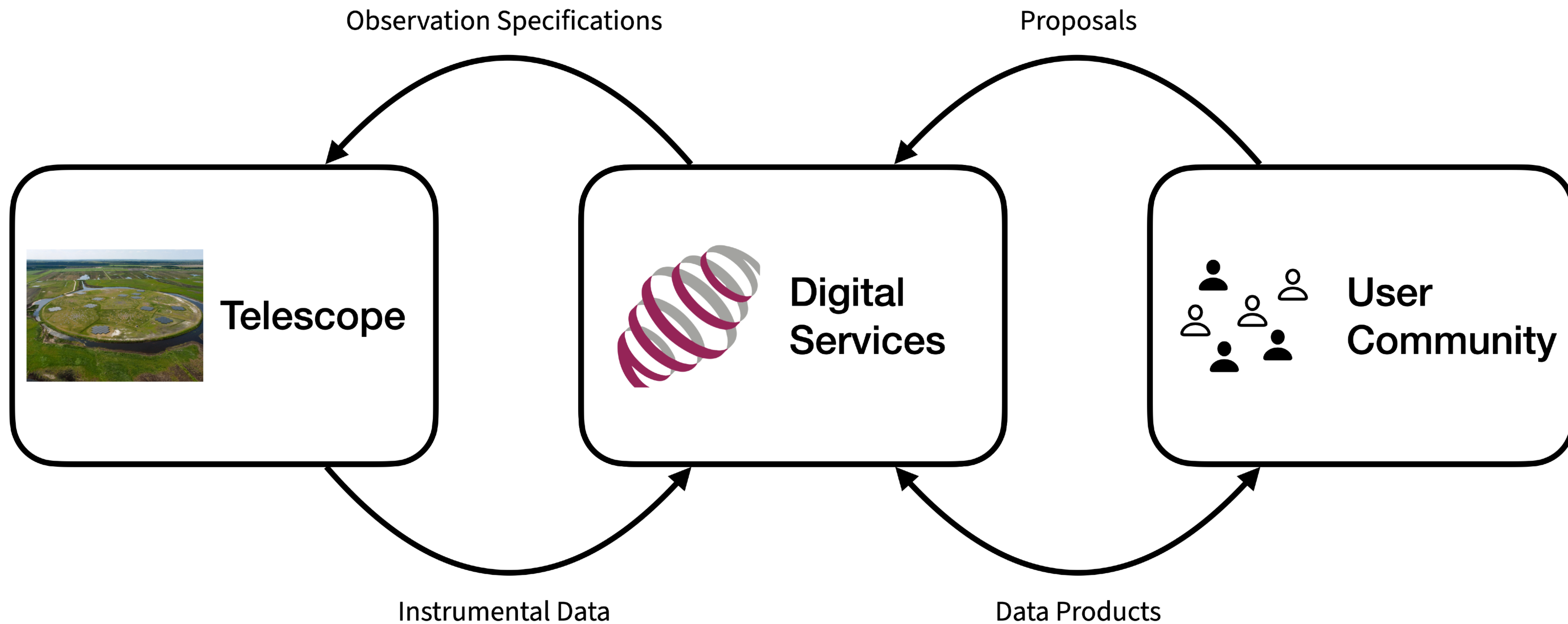
A Vision for LOFAR Digital Services

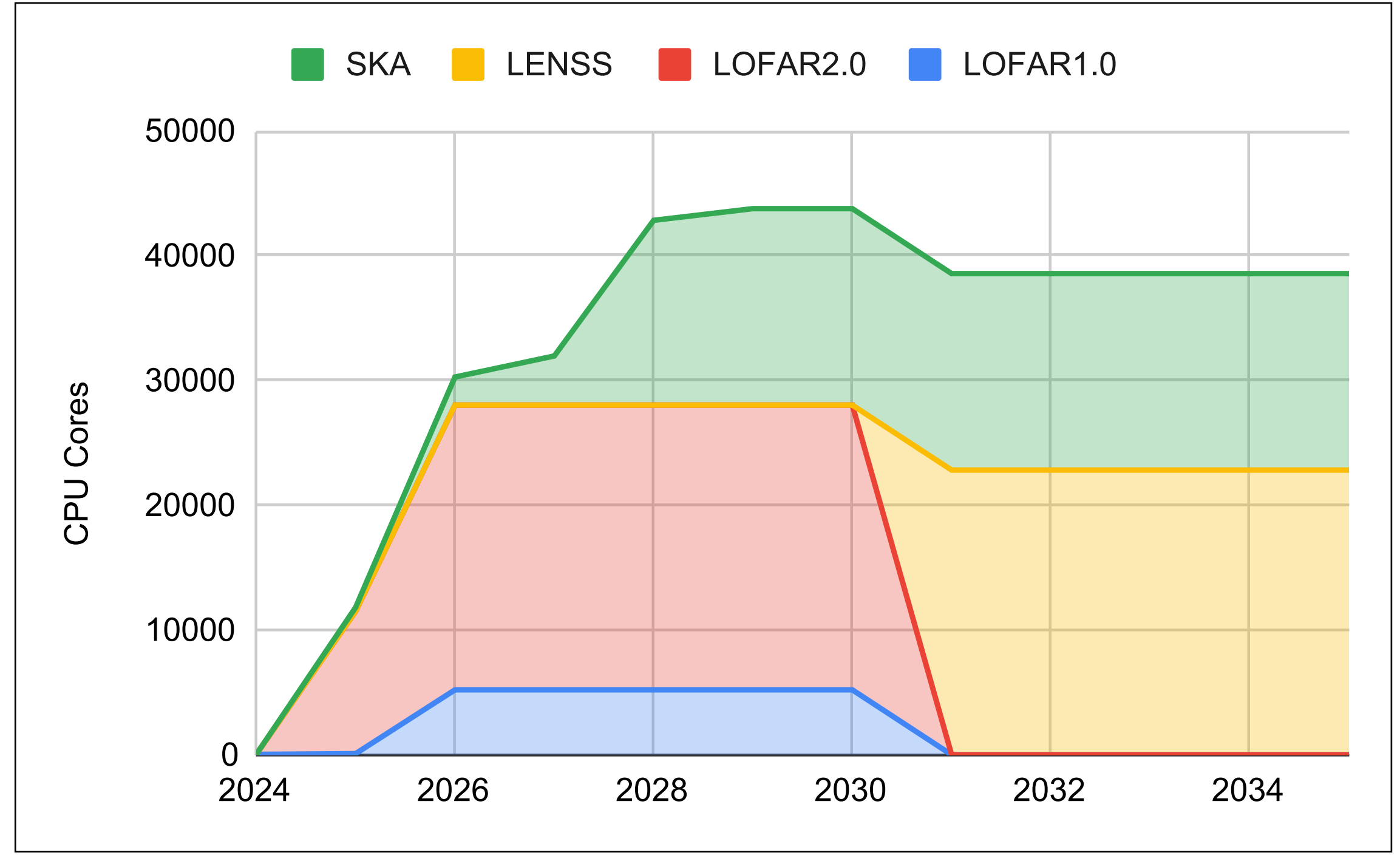
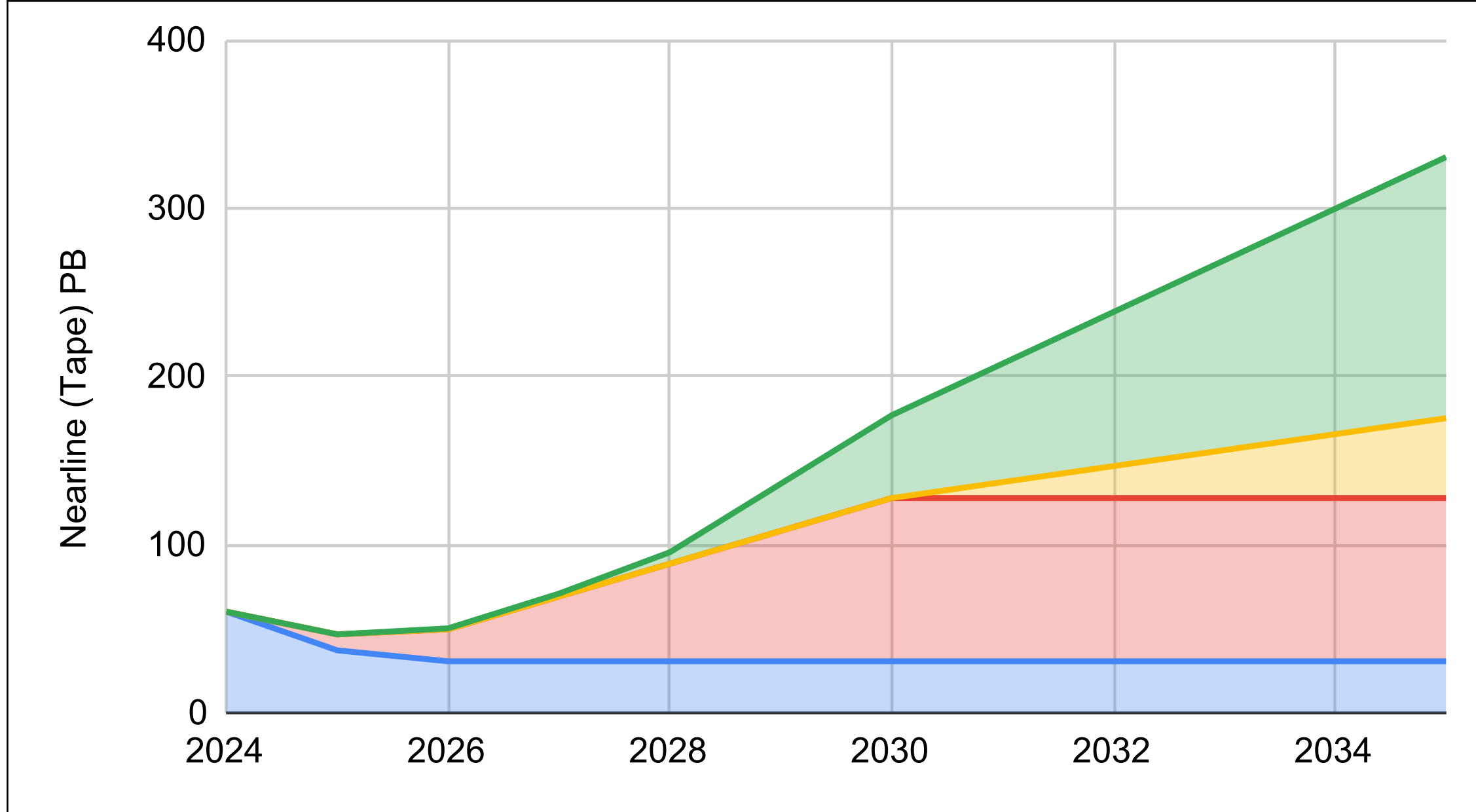
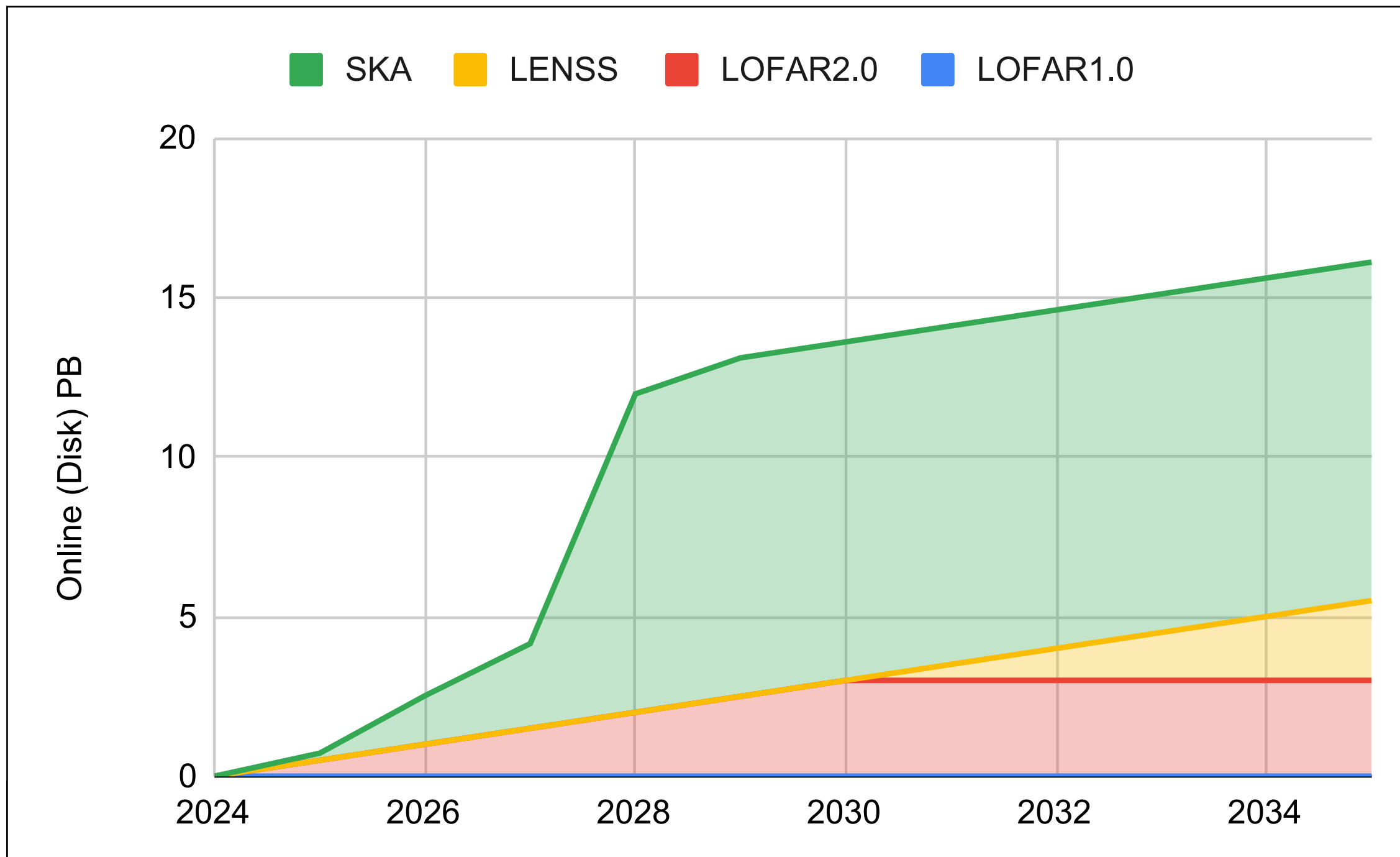
John D. Swinbank
swinbank@astron.nl

ASTRON

Netherlands Institute for Radio Astronomy









Proposal
Management



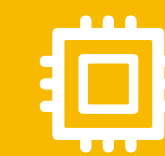
Archiving
& Curation



Scientific
Pipelines



Digital
Services



Managed
Processing



Discovery
& Access



Interactive
Data Analysis



User Pipeline
Execution



Proposal
Management



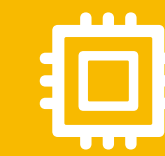
Archiving
& Curation



Scientific
Pipelines



Digital
Services



Managed
Processing



Discovery
& Access



Interactive
Data Analysis



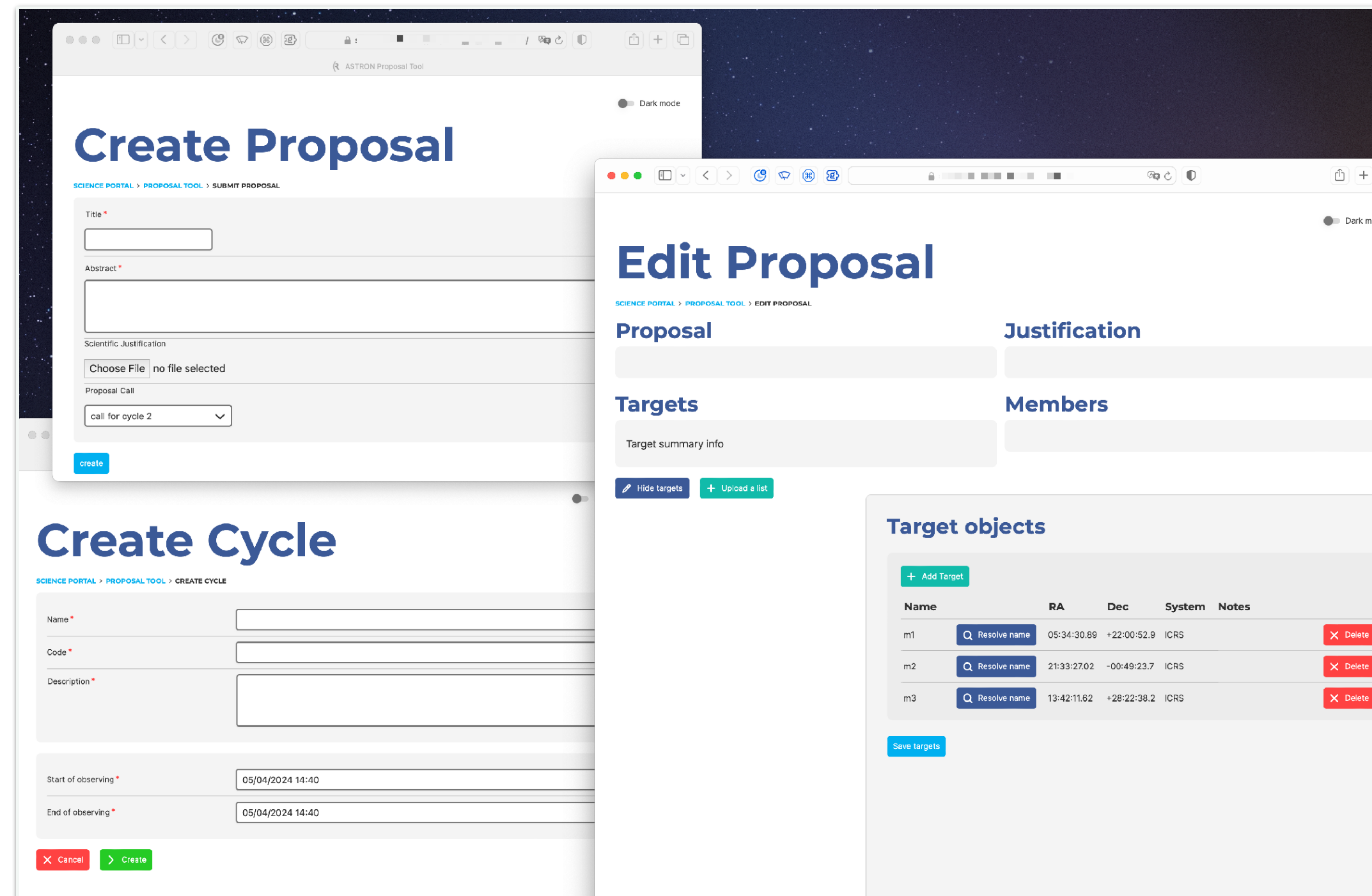
User Pipeline
Execution

COMING LATER

COMING LATER

Proposal Management

- Replacement for the Northstar tool used in LOFAR1, which is obsolete & unmaintainable.
- This also includes new user/group management systems: “federated authentication & authorization”.
- Usable / testable / commissionable version ~this summer.



Federated AAI: SURF SRAM

Logging in with your own institutional account

The image shows a 'Sign in to your account' form with fields for Username and Password, and a 'Sign In' button. Below it are options for 'SRAM acceptance environment' and 'SRAM prod'. To the right is a 'Proposal Overview' page showing a table of collaborative organizations with columns for Title, Collaborative Organization (with IDs), and Actions (Edit, View, Write). A 'SURF' logo is overlaid on the login form.

Collaboration (project team) management

This block contains three screenshots of the collaboration management interface. The top-left screenshot shows the 'Create new collaboration' form with fields for Name, Short name, Description, and Invitees (comma separated), with 'Submit' and 'Cancel' buttons. The top-right screenshot shows a 'Welcome to My gr8 collaboration' page with a notification 'You are invited to become member of this collaboration', a 'Purpose of the collaboration' section, and a 'Proceed to My gr8 collaboration' button. The bottom screenshot shows the 'My gr8 collaboration' overview page with tabs for About, Admins, Members, Groups, Services, Join requests, and Service tokens. The 'Members (2)' section lists members with their names, emails, institutions, roles, and membership expiration dates.

Consent for GDPR & Acceptable use

The image shows two pages related to consent. The left page is titled 'Review your information that will be shared' and lists personal information: Display Name (Hanno Holties), Full Name (holties), First name (Hanno), Surname (Holties), and Email address (holties@astron.nl). It also shows an identifier and logos for 'eduTEAMS Service' and 'SURFconect'. The right page is a 'Welcome to My gr8 collaboration' page with a notification, a 'Purpose of the collaboration' section, and a 'Proceed to My gr8 collaboration' button at the bottom.

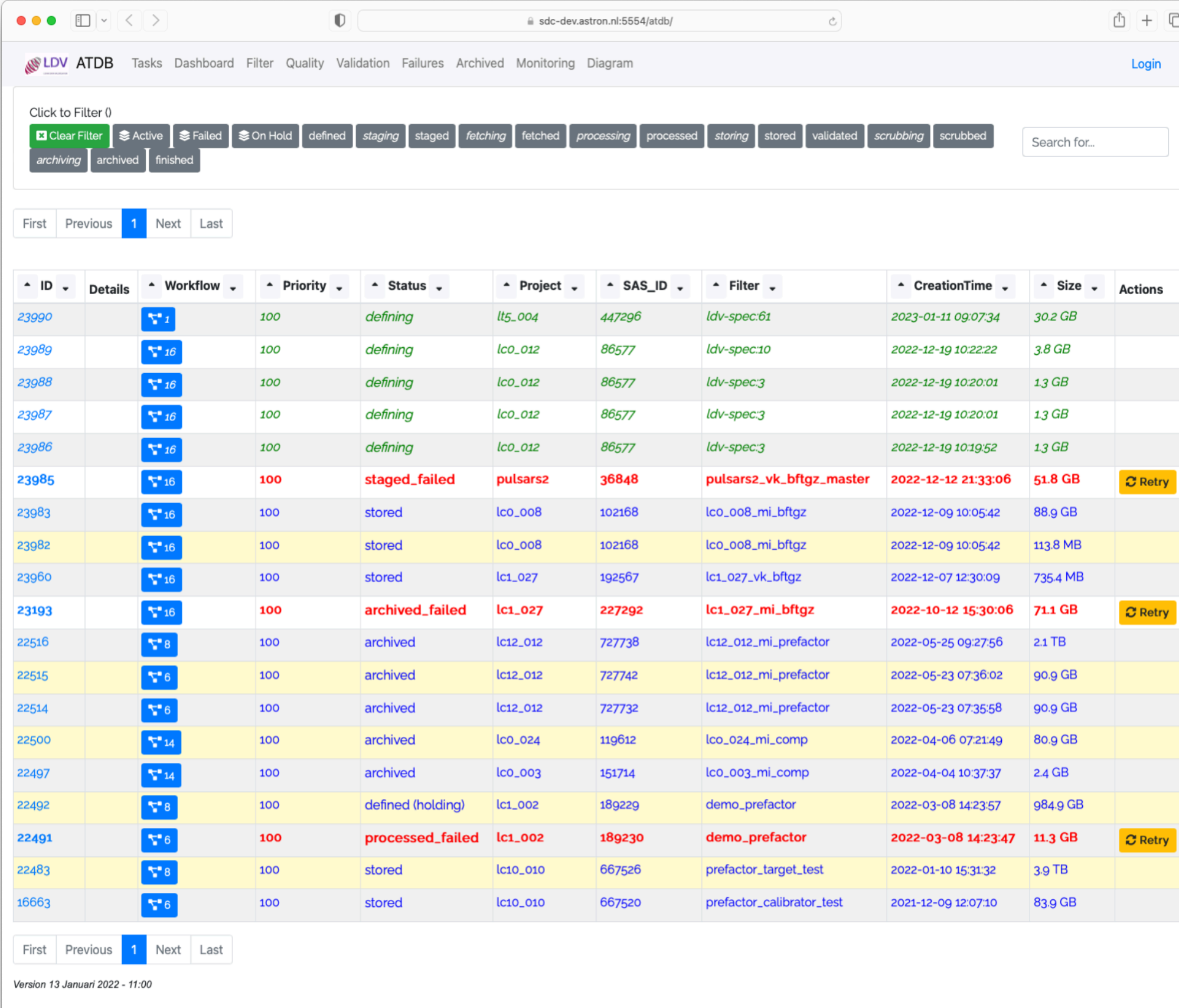
Archiving & Curation

Product Type	Example	Retention Period
Raw	Unprocessed visibilities	Not retained
Instrumental	Flagged & compressed visibilities	O(18 months)
Intermediate	Direction-independent calibrated visibilities	O(18 months)
Advanced	Image cubes	Indefinite
Special Cases	Unique observations that cannot be repeated	For discussion

- LTA Support for LOFAR ERIC agreed data policy.
- LTA support for “advanced” data products (e.g. images).
- Goal: the ability to ingest advanced / science-ready data products generated by the wider community.
 - Including management of data rights.
 - Become a “hub” for access to LOFAR data, wherever it is generated.

Managed Processing

- Execute predefined pipelines “at scale” against data in the LTA.
- Capability developed in the context of the LOFAR Data Valorization effort, currently running at SURF.
- Future work:
 - Scale to other LTA sites (Jülich, Poznań, maybe more).
 - Increased automation.
 - Incorporate more pipelines.
 - Polish & user enhancements.

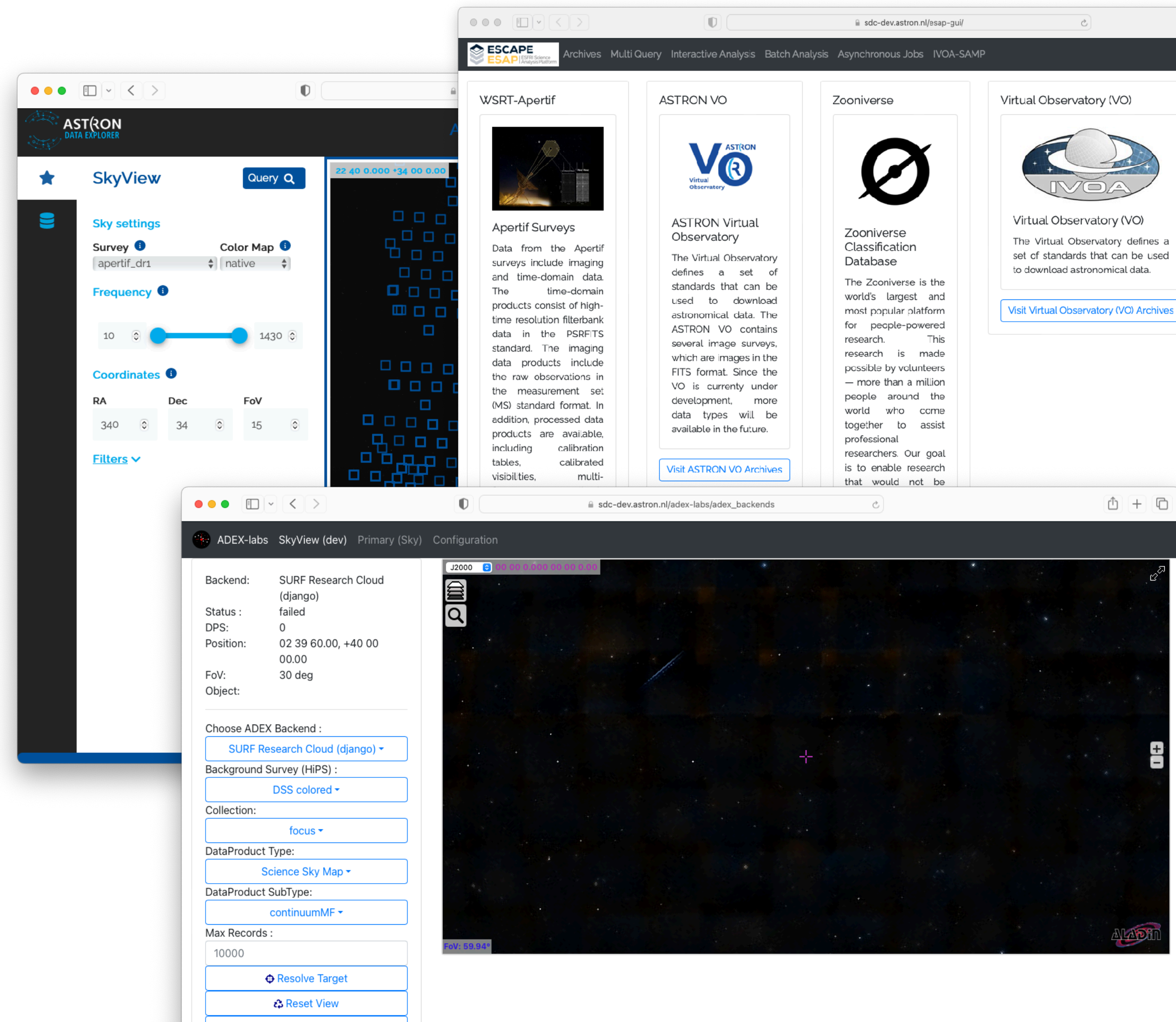


The screenshot shows the LDV ATDB web interface. The top navigation bar includes 'LDV ATDB', 'Tasks', 'Dashboard', 'Filter', 'Quality', 'Validation', 'Failures', 'Archived', 'Monitoring', and 'Diagram'. A search bar is on the right. Below the navigation, there are filter buttons for 'archiving', 'archived', and 'finished'. A table of tasks is displayed with columns for ID, Details, Workflow, Priority, Status, Project, SAS_ID, Filter, CreationTime, Size, and Actions. The table contains 20 rows of task data.

ID	Details	Workflow	Priority	Status	Project	SAS_ID	Filter	CreationTime	Size	Actions
23990			100	defining	lt5_004	447296	ldv-spec:61	2023-01-11 09:07:34	30.2 GB	
23989			100	defining	lc0_012	86577	ldv-spec:10	2022-12-19 10:22:22	3.8 GB	
23988			100	defining	lc0_012	86577	ldv-spec:3	2022-12-19 10:20:01	1.3 GB	
23987			100	defining	lc0_012	86577	ldv-spec:3	2022-12-19 10:20:01	1.3 GB	
23986			100	defining	lc0_012	86577	ldv-spec:3	2022-12-19 10:19:52	1.3 GB	
23985			100	staged_failed	pulsars2	36848	pulsars2_vk_bftgz_master	2022-12-12 21:33:06	51.8 GB	Retry
23983			100	stored	lc0_008	102168	lc0_008_mi_bftgz	2022-12-09 10:05:42	88.9 GB	
23982			100	stored	lc0_008	102168	lc0_008_mi_bftgz	2022-12-09 10:05:42	113.8 MB	
23960			100	stored	lc1_027	192567	lc1_027_vk_bftgz	2022-12-07 12:30:09	735.4 MB	
23193			100	archived_failed	lc1_027	227292	lc1_027_mi_bftgz	2022-10-12 15:30:06	71.1 GB	Retry
22516			100	archived	lc12_012	727738	lc12_012_mi_prefactor	2022-05-25 09:27:56	2.1 TB	
22515			100	archived	lc12_012	727742	lc12_012_mi_prefactor	2022-05-23 07:36:02	90.9 GB	
22514			100	archived	lc12_012	727732	lc12_012_mi_prefactor	2022-05-23 07:35:58	90.9 GB	
22500			100	archived	lc0_024	119612	lc0_024_mi_comp	2022-04-06 07:21:49	80.9 GB	
22497			100	archived	lc0_003	151714	lc0_003_mi_comp	2022-04-04 10:37:37	2.4 GB	
22492			100	defined (holding)	lc1_002	189229	demo_prefactor	2022-03-08 14:23:57	984.9 GB	
22491			100	processed_failed	lc1_002	189230	demo_prefactor	2022-03-08 14:23:47	11.3 GB	Retry
22483			100	stored	lc10_010	667526	prefactor_target_test	2022-01-10 15:31:32	3.9 TB	
16663			100	stored	lc10_010	667520	prefactor_calibrator_test	2021-12-09 12:07:10	83.9 GB	

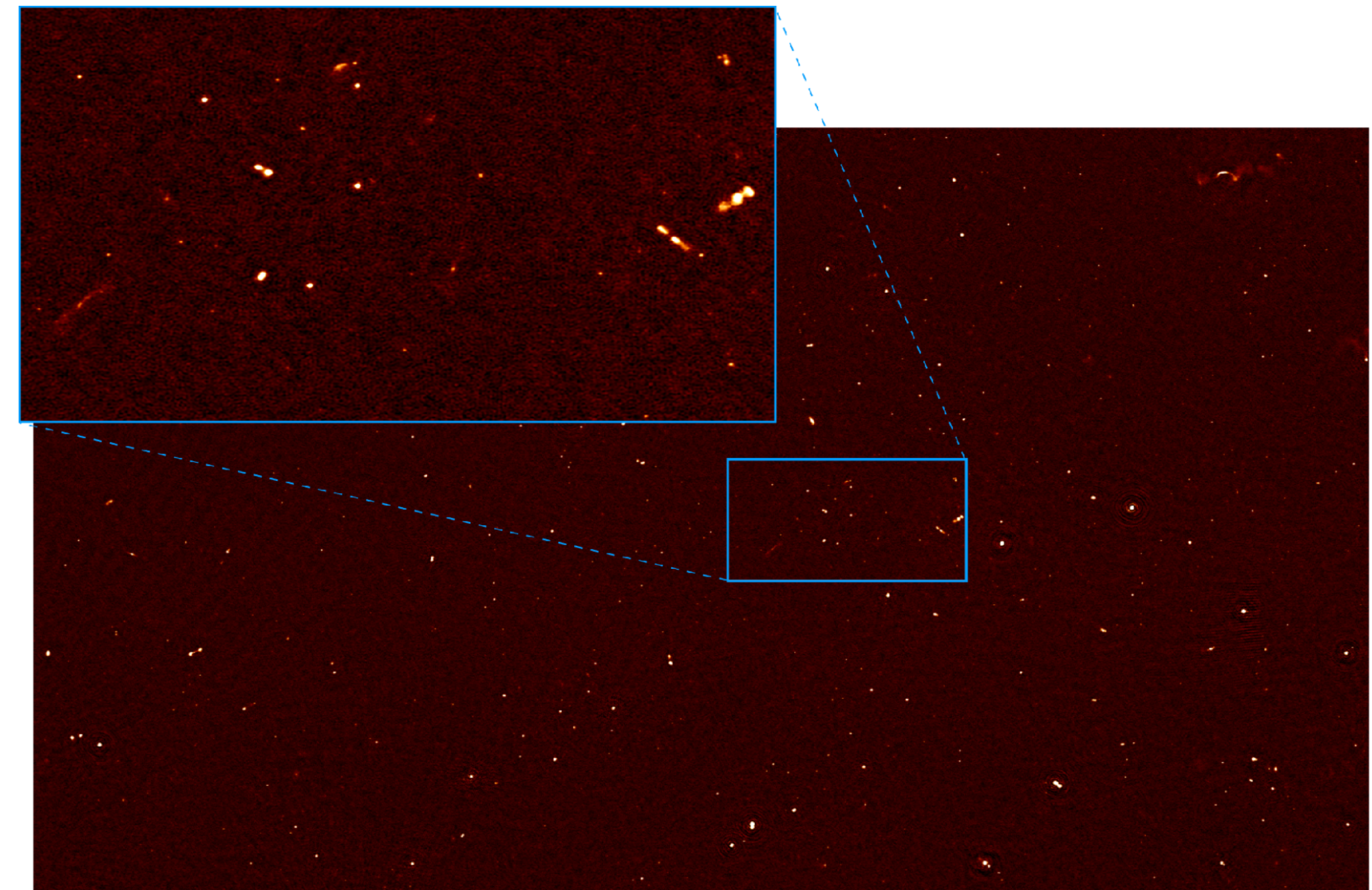
Discovery & Access

- Software solutions to make archive access as robust & reliable as possible...
- ...in tandem with LOFAR ERIC service level agreements with data centres.
- Upgraded archive interface: “ADEX”.
- Pervasive use of Virtual Observatory interfaces for publishing data.
- Aiming for a fully “FAIR” compliant archive:
 - Findable, Accessible, Interoperable, Reusable
 - <https://force11.org/info/the-fair-data-principles/>



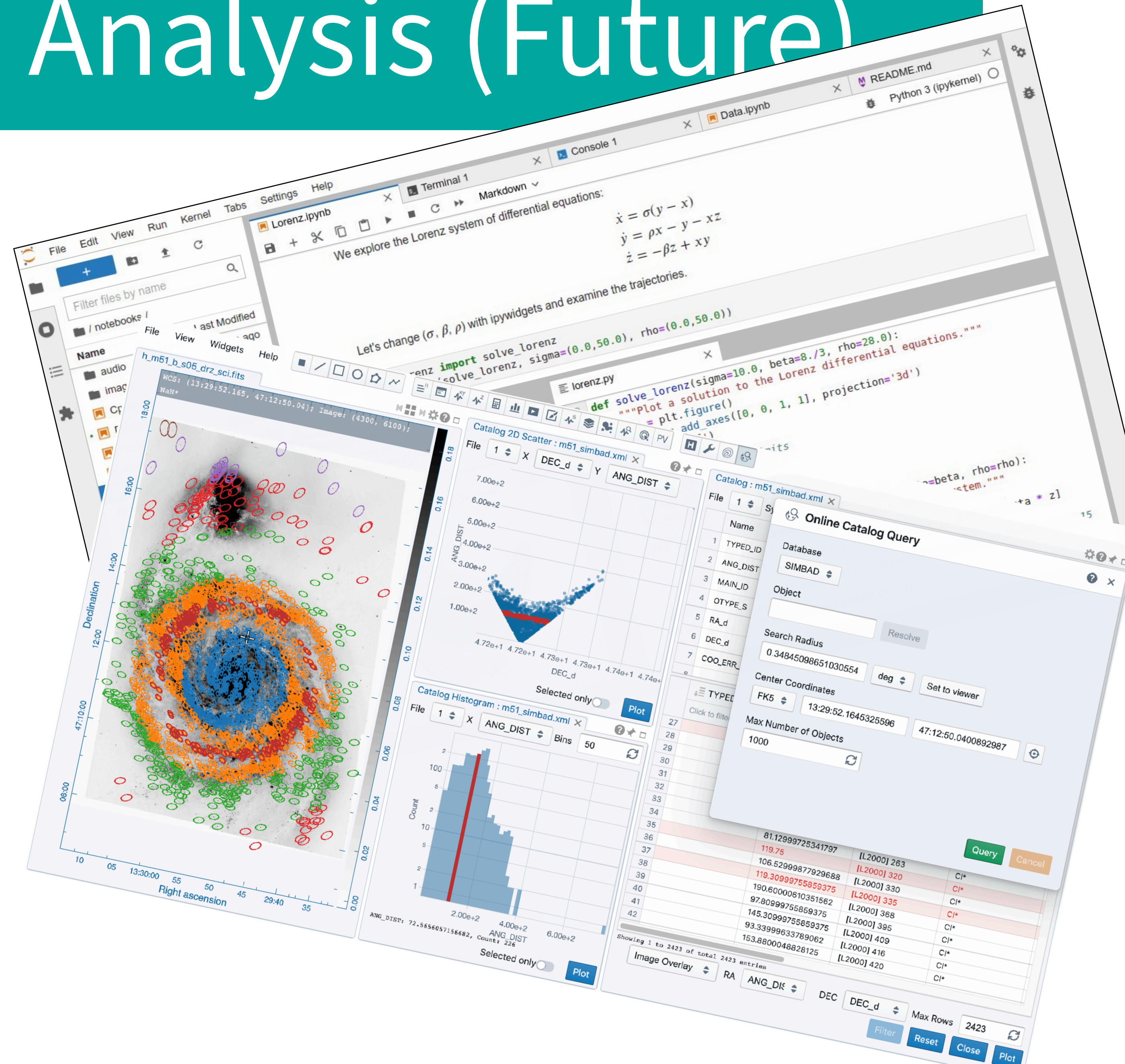
Scientific Pipelines

- The Observatory-supported portfolio for LOFAR2.0:
 - Pre-Processing
 - LINC (direction-independent calibration)
 - Rapthor (direction-dependent calibration)
 - VLBI (postage stamps & wide field)
 - PULP (known pulsars)
 - TraP (image plane transients; *stretch goal*)
- For cost & science productivity reasons, processing plans/pipelines must be in place before observations start.
- Not “ASTRON's pipelines”, but *community* pipelines; we work together to make them effective.



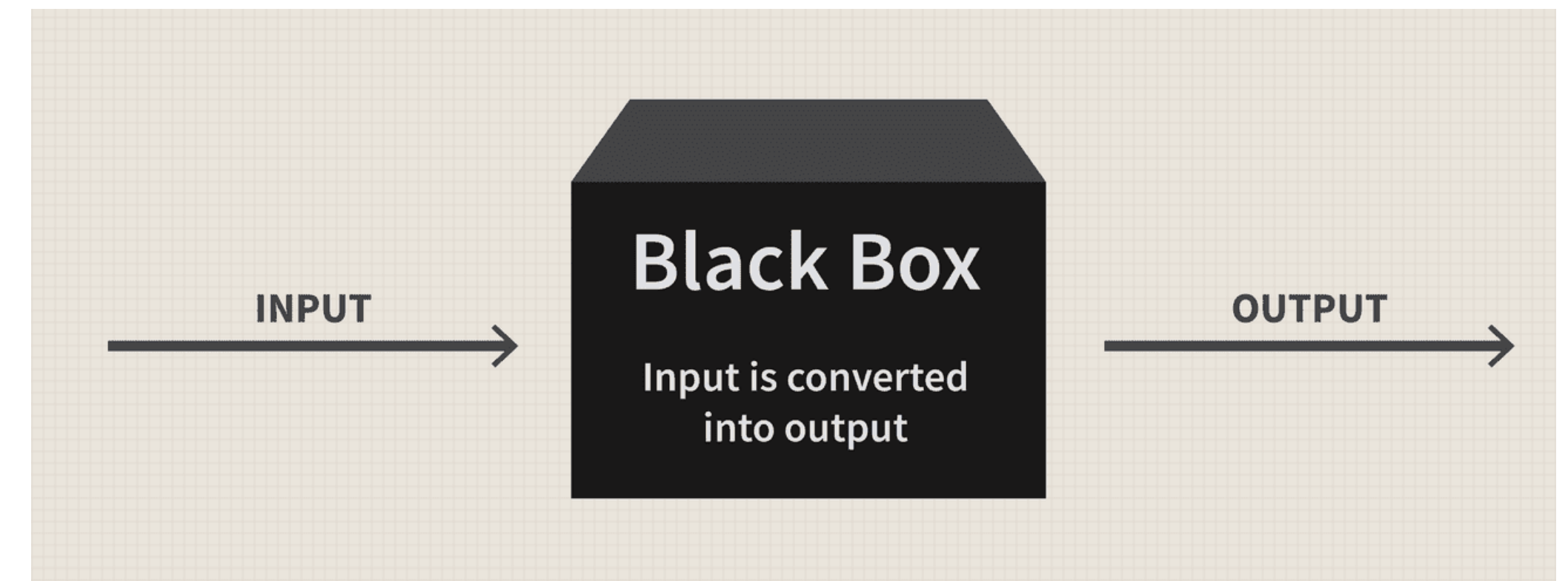
Interactive Data Analysis (Future)

- Common expectation: Jupyter notebooks running next to the data.
 - Jupyter notebook: interactive browser-based environment including live code, text, figures, etc.
- Also: “legacy” graphical applications (CASA, TOPCAT, ...), though e.g. VNC/remote desktop connections.
- Also also: command line applications, through e.g. SSH.
- Implies the existence of shared, persistent storage for work in progress, output products, etc.

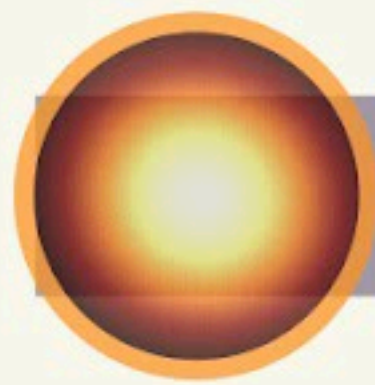


User Pipeline Execution (Future)

- Whatever observatory-supported pipelines are available, there will always be new science cases that aren't supported.
- Enable them, while minimizing risk to our operational system.
- Provide an API against which pipelines can be developed.
- Provide a “black box” system with appropriate quotas etc for executing untrusted payloads.



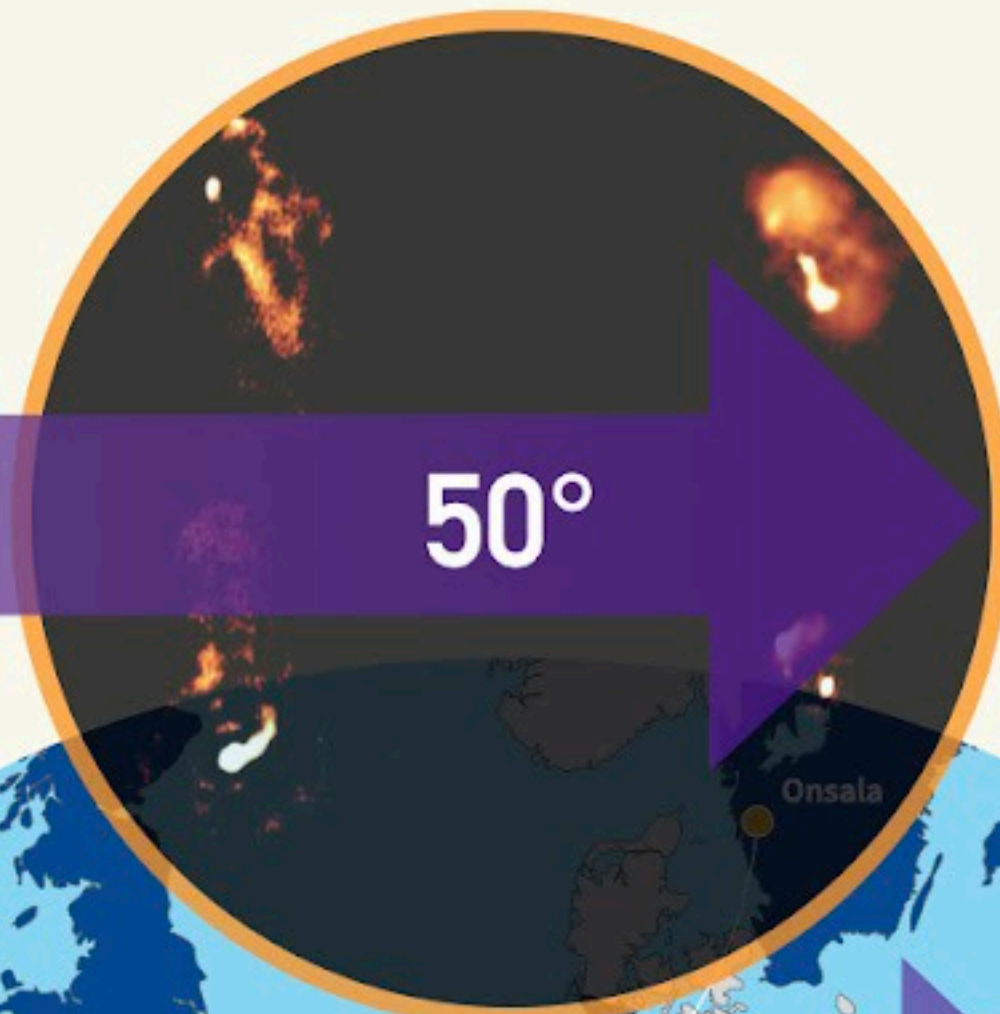
Carbon neutral Science



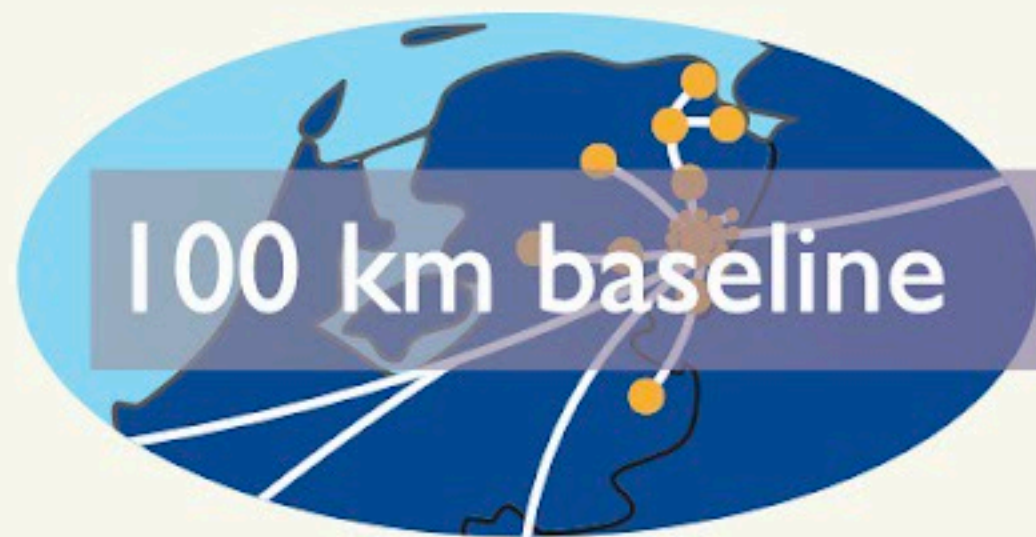
12°

field-of-view

x4



50°



100 km baseline

Resolution

x20



2000 km baseline



20 days

Speed to science

x40

12 hours

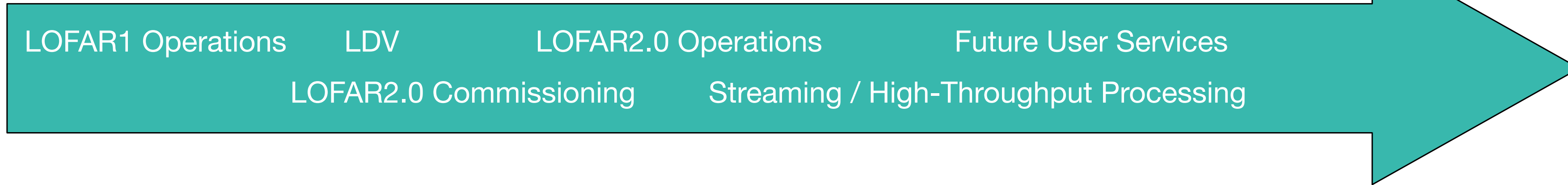


LOFAR2.0

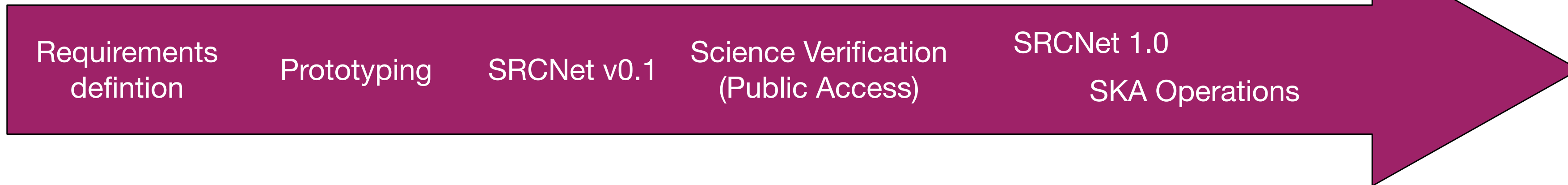
LOFAR2.0+LENS

2022 2023 2024 2025 2026 2027 2028 2029 2030

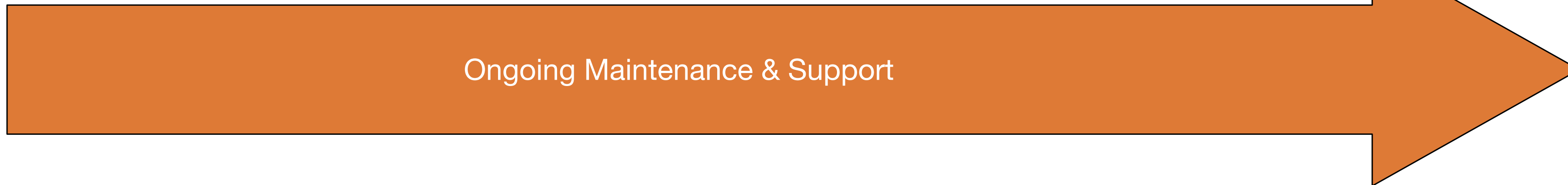
LOFAR Digital Services



SKA Regional Centre Network

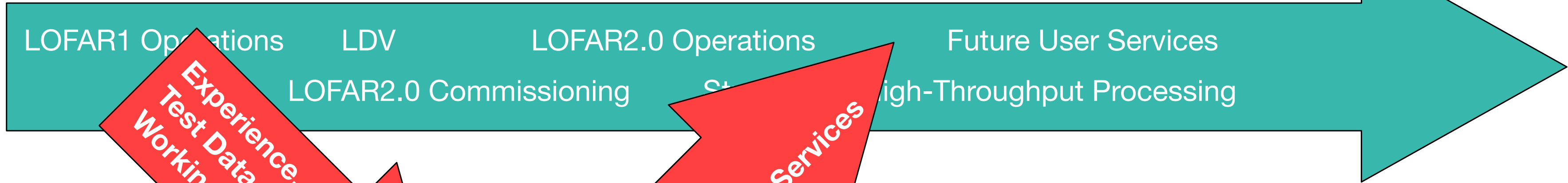


Apertif Long Term Archive



2022 2023 2024 2025 2026 2027 2028 2029 2030

LOFAR Digital Services



SKA Regional Cent



Apertif Long Term Archive



Experience, Test Data, Working Services

Technology, Next Generation Services

Conclusions

- The volume of data already produced by LOFAR is enormous...
- ...and it'll be even bigger in the LOFAR2.0 future.
- That presents us with challenges, both in terms of how we manage the data volume, but also in terms of the range of services we need to provide to end users.
- Addressing those challenges is something we can best address together, and we can also look to learn from parallel efforts in SKA.