Multi-dimensional Data Priority Area

Mark Allen
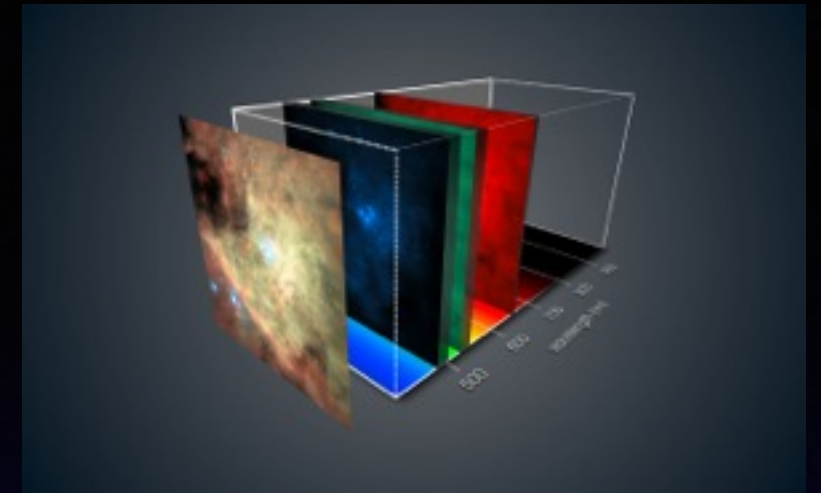
EURO VO

CoSADIE Project

Science Priority Areas

Multi-dimensional Data

Radio astronomy, Integral Field Spectroscopy, high energy, polarization, simulation, data mining datasets + ...

- Need to ensure that these are accessible and useable within the VO

Status

- Focus Sessions (Heidelberg May 2013)

  ‣ engagement with projects, minimal requirements

- Prototype demos (Hawaii Sept 2013)

  ‣ Multiple approaches, ~agreement on stds. needed

- 1st stds. to satisfy minimal requirements

- aiming for RFCs < May Interop 2014

- Follow-up Focus Session (Madrid May 2014)

Projects engaged

- ALMA
- LOFAR
- SKA
- ASKAP
- JVLA / NRAO
- MUSE

- CALIFA
- LSST
- CTA

CALIFA 1st DATA RELEASE

VO already in use/plans

- ALMA - using OpenCADC TAP, voview, will use SAMP, ObsCore, SIAPv2

- CyberSKA - VO access option via CADC

- CALIFA - data access via TAP and SSA

- MUSE - VO publishing via AstroWise

- ASKAP - all data through VO protocols

- CoRoT - avail from SVO, Kepler - avail from MAST
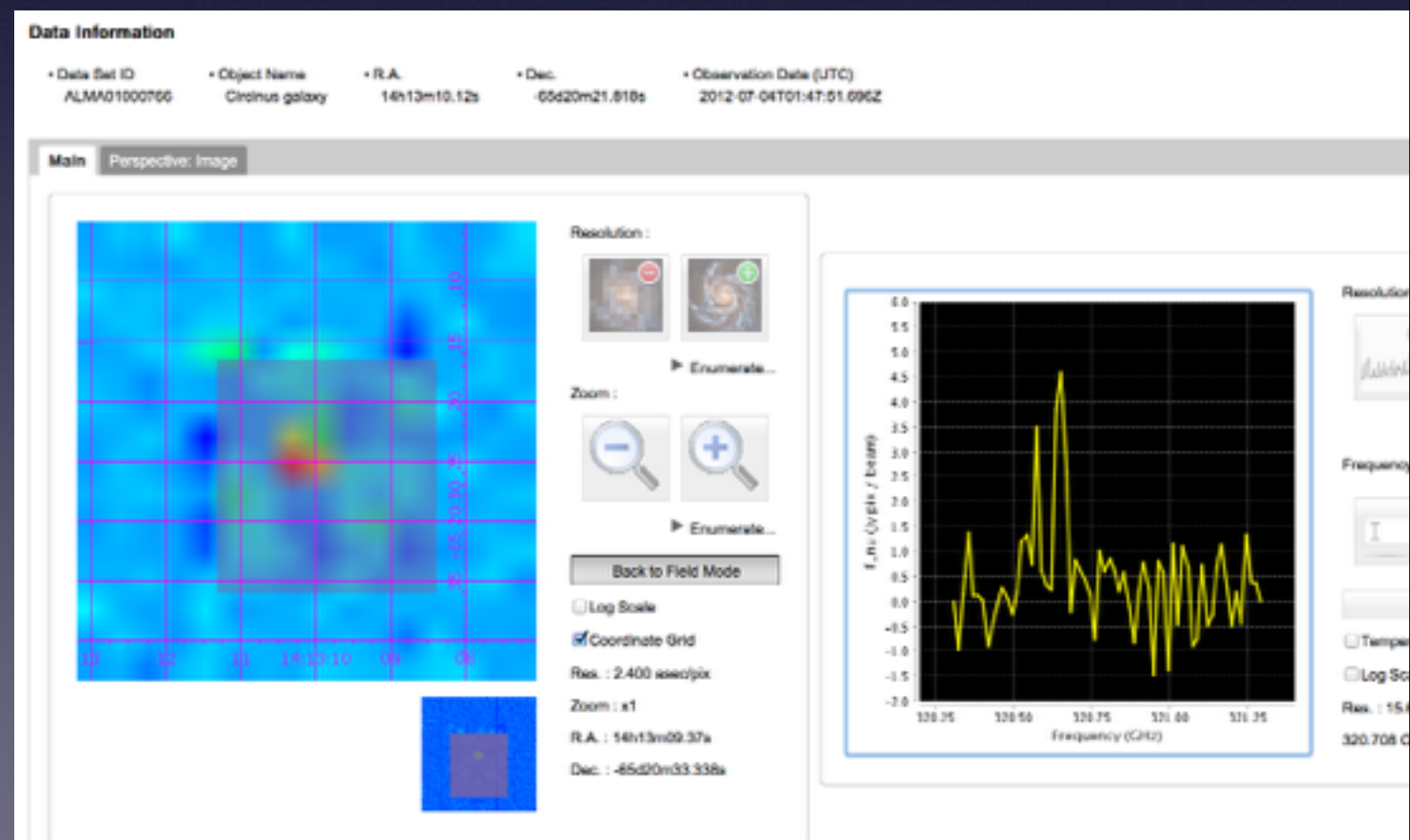
- LOFAR - VOEvent broker

- Identify implementation hurdles
  - description/access to cubes } in progress
  - query by time parameters
  - expense
- mismatch of expectations/approach
  - implementation of 'standards' cf. 'libraries'
  - *"Rough consensus and running code"*
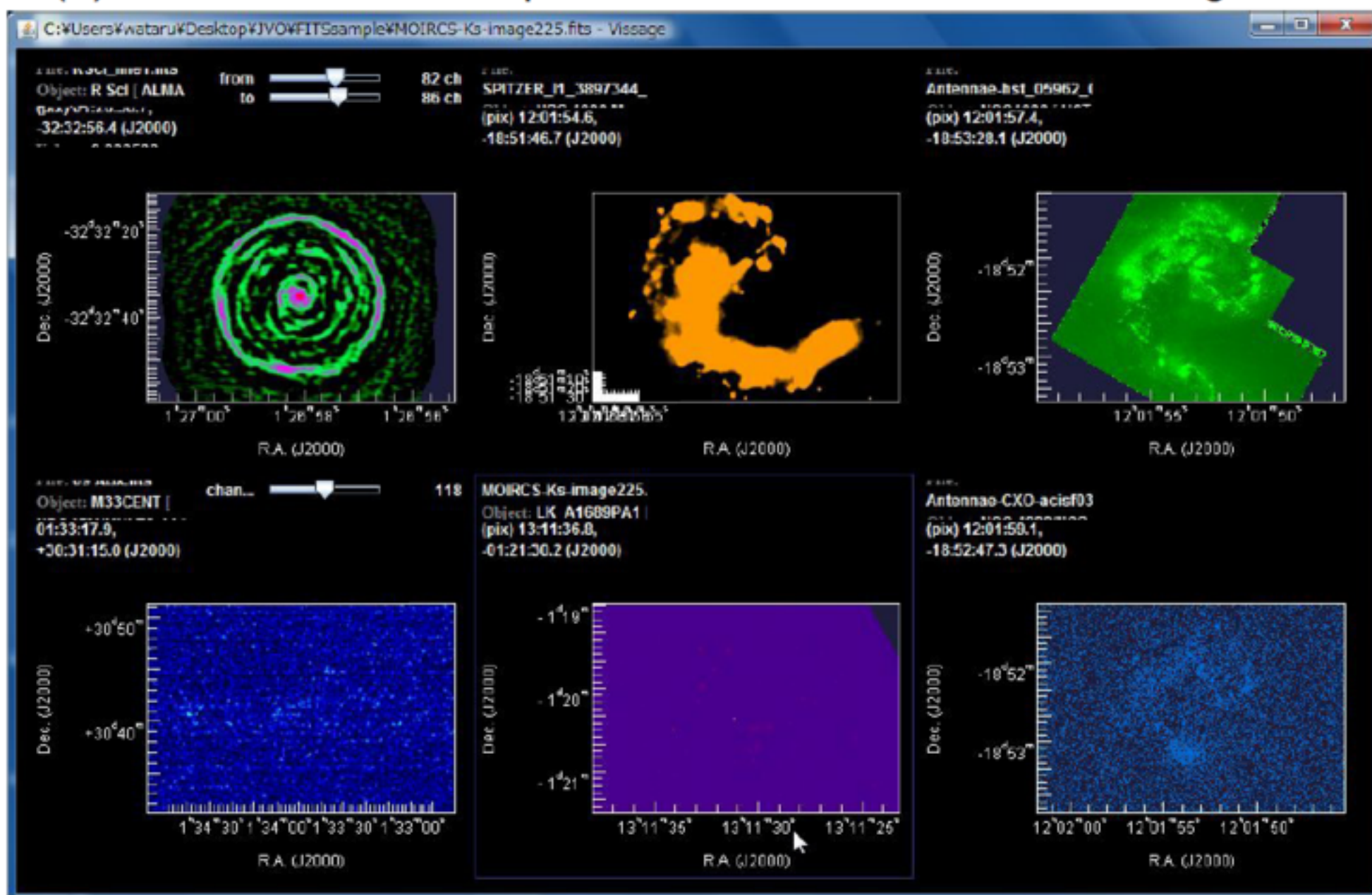 - Java and Python...
  - Importance of ...

Prototype demos
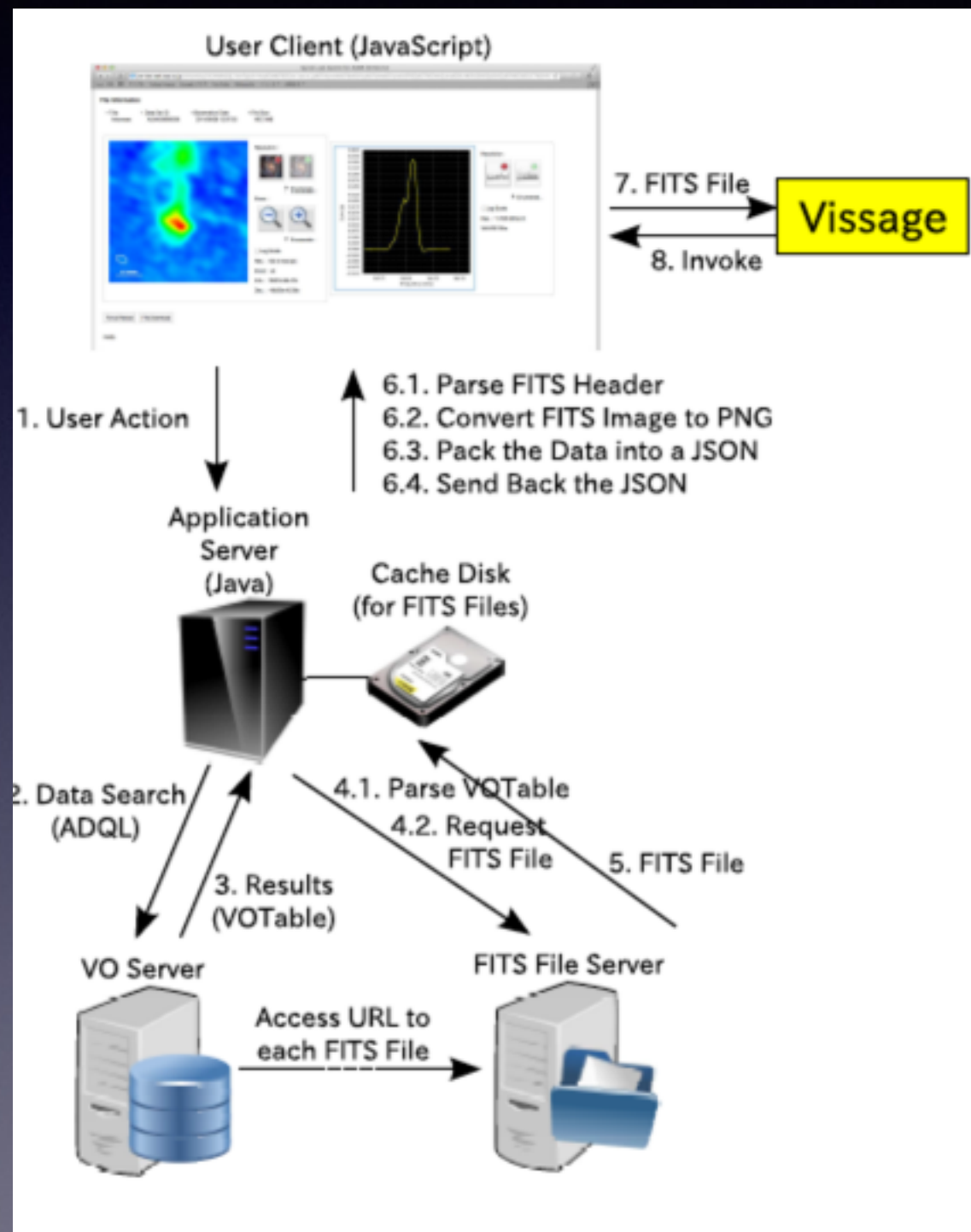
- JVO ALMA VO Portal and Vissage

Demonstration

(9) ALMA / NRO45m / Spitzer / Subaru / HST / Chadra images

ImageDM metadata for the cube: SIA2Query response

- VAO development

- http://vaosa-vm1.aoc.nrao.edu/ivoa-dal/siav2query.html

Minimal requirements

- **Data Discovery (Query)**
  - A service shall be able to receive queries regarding its data collection(s) from a client, with the client placing one or more of the following constraints:
    - RA,Dec
    - Frequency/wavelength
    - Polarization states
    - Spatial size
    - Angular resolution
    - Integration time
    - Time of observation
  - A service shall return to the client a list of observations, and the corresponding metadata for each observation, meeting the user-imposed constraints. In the event that the user places no constraints, the entire list of observations, and the corresponding metadata for each data set, shall be returned. In the event that no data meet the user's constraints, the service shall indicate the absence of any matches.
- **Data Access**
  - Once a user has the list of observations that satisfy the constraints, they select all or a subset of the observations and:
    - Download the complete science data for each of the selected observations (the service shall return the complete multi-dimensional science data and metadata for each selected observation) or;
    - Download simple cutouts of the science data for each of the selected observations (the service shall be able to extract and return a user-specified subset of the complete multi-dimensional science data and metadata for each selected observation).
- **Simple Cutout**
  - For a simple cutout, the user-specified subset is restricted to be a contiguous interval within each dimension of the multi-dimensional science data. The user should *not* be allowed to specify subsets with "gaps" or resampling or anything like that.
    - Spatial: a circle (a coordinate and a radius)
    - Energy: one interval (from energy1 to energy2)
    - Time: one interval (from time1 to time2)
    - Polarization: a list

Mostly accepted as reasonable requirements

- Data Discovery, Data Access

Some discussion about spatial cutout

- Cutout: For a simple cutout, the user-specified subset is restricted to be a contiguous interval within each dimension of the multi-dimensional science data. The user should *not* be allowed to specify subsets with "gaps" or resampling or anything like that.

  - Spatial: a circle (a coordinate and a radius)
  - Energy: one interval (from energy1 to energy2)
  - Time: one interval (from time1 to time2)
 - Polarization: a list

TCG to WGs :

Practically speaking, the endorsement of the minimum requirements for the first version of the multi-dimensional data access standards means:
- The preparation of the standards cannot be held up by discussion of "features" that are not necessary to meet the minimum standards
- The WGs should be thinking in an agile sense where subsequent versions of a given standard with more "features" come rapidly after the first version.

Standards that need to be RFC-ready by the May InterOp:
1. SIAv2 (query capability only)
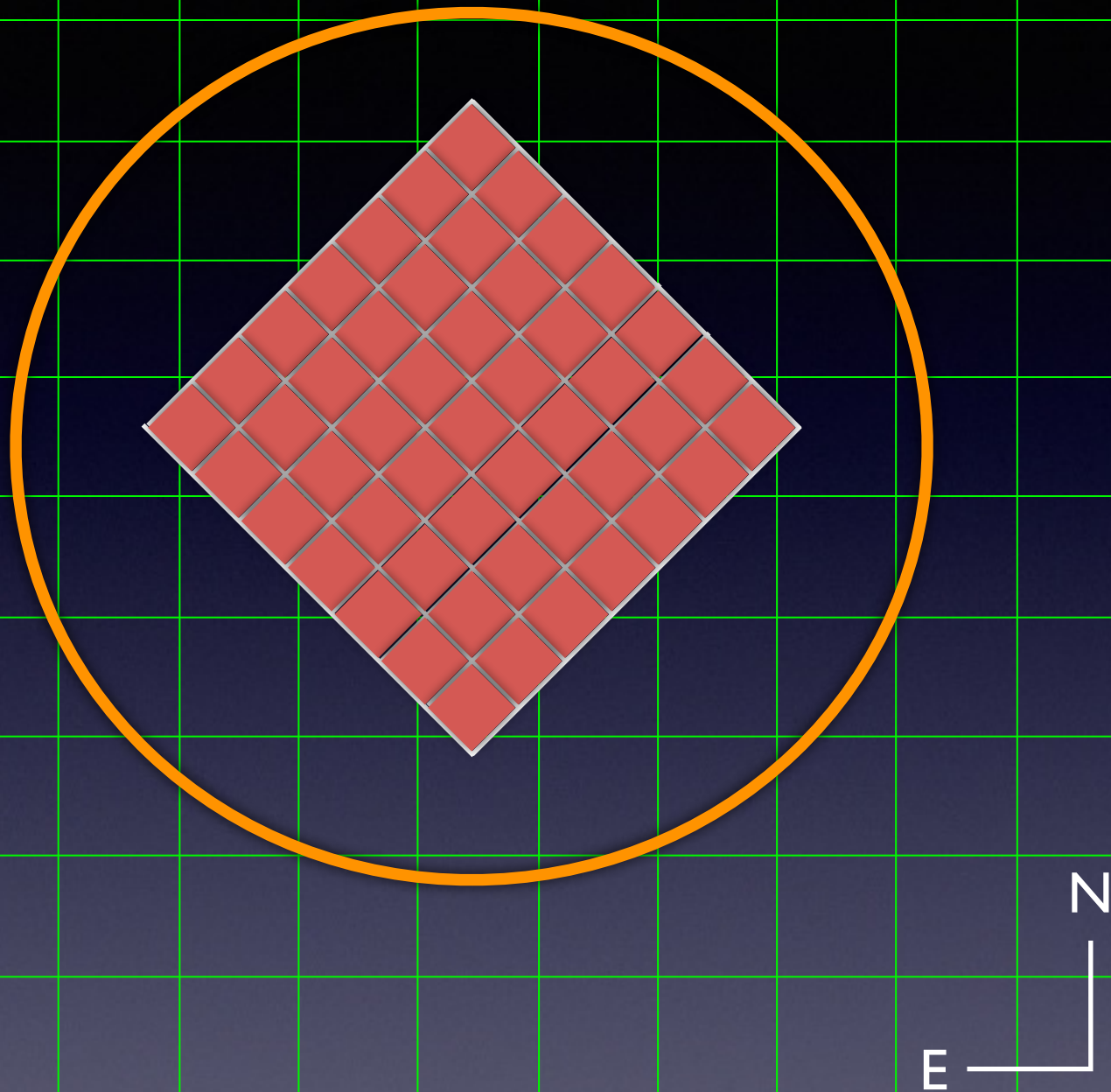2. DataLink
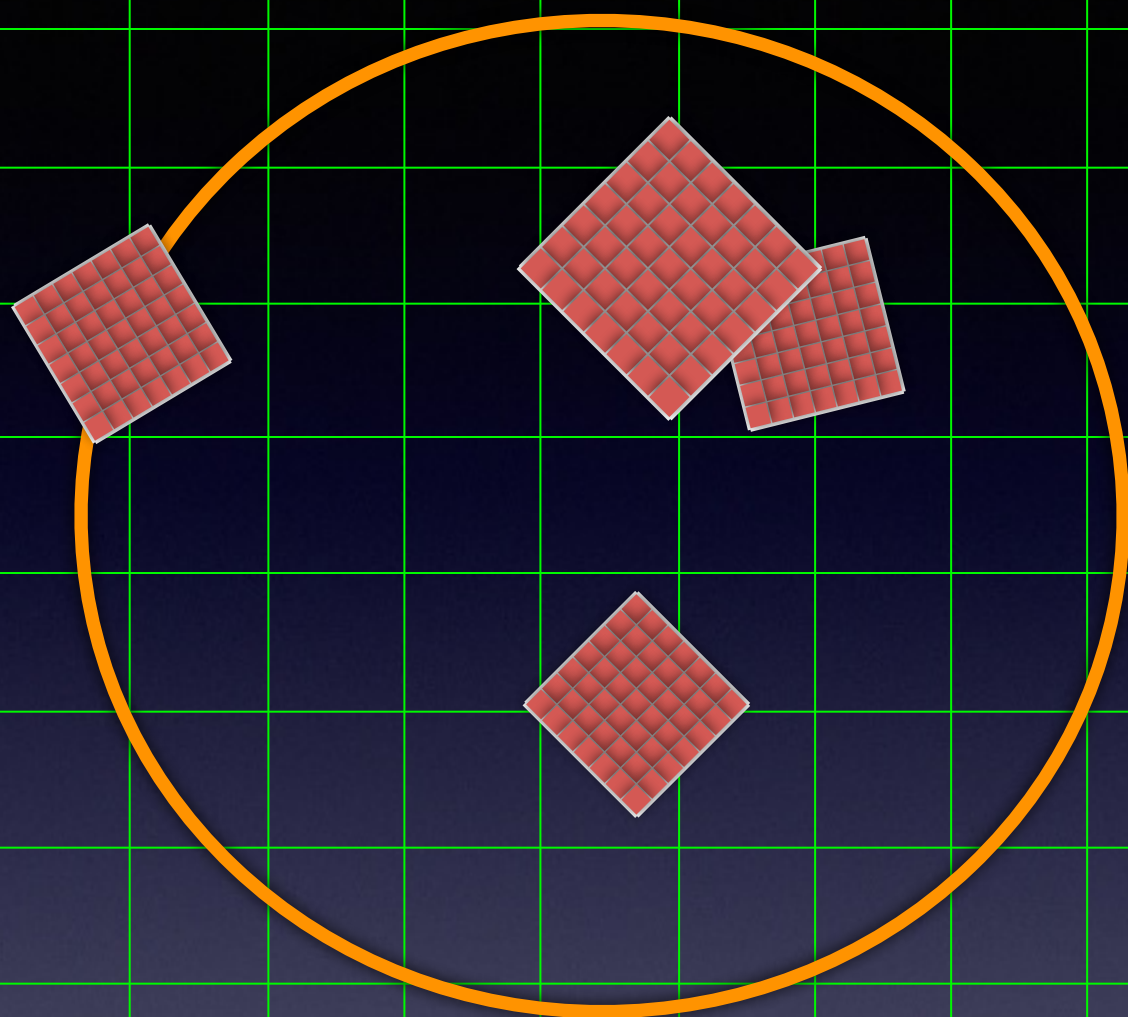3. AccessData (for simple cutouts only)

Cutouts
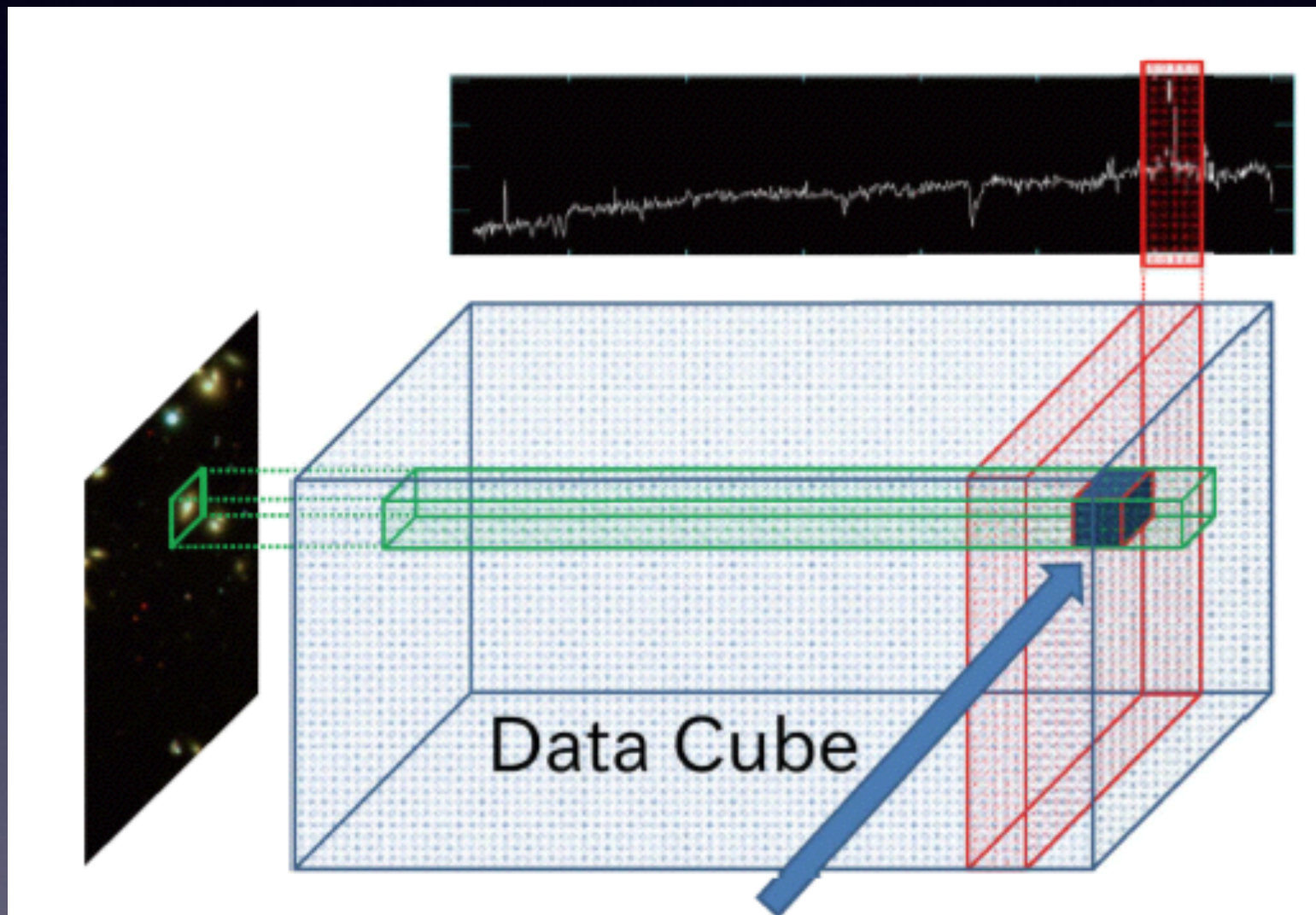
- simple cutout (without re-computing), along the cube axes, no gaps, transforms
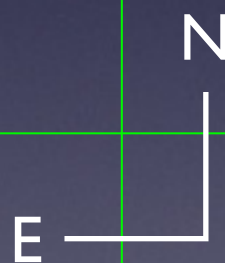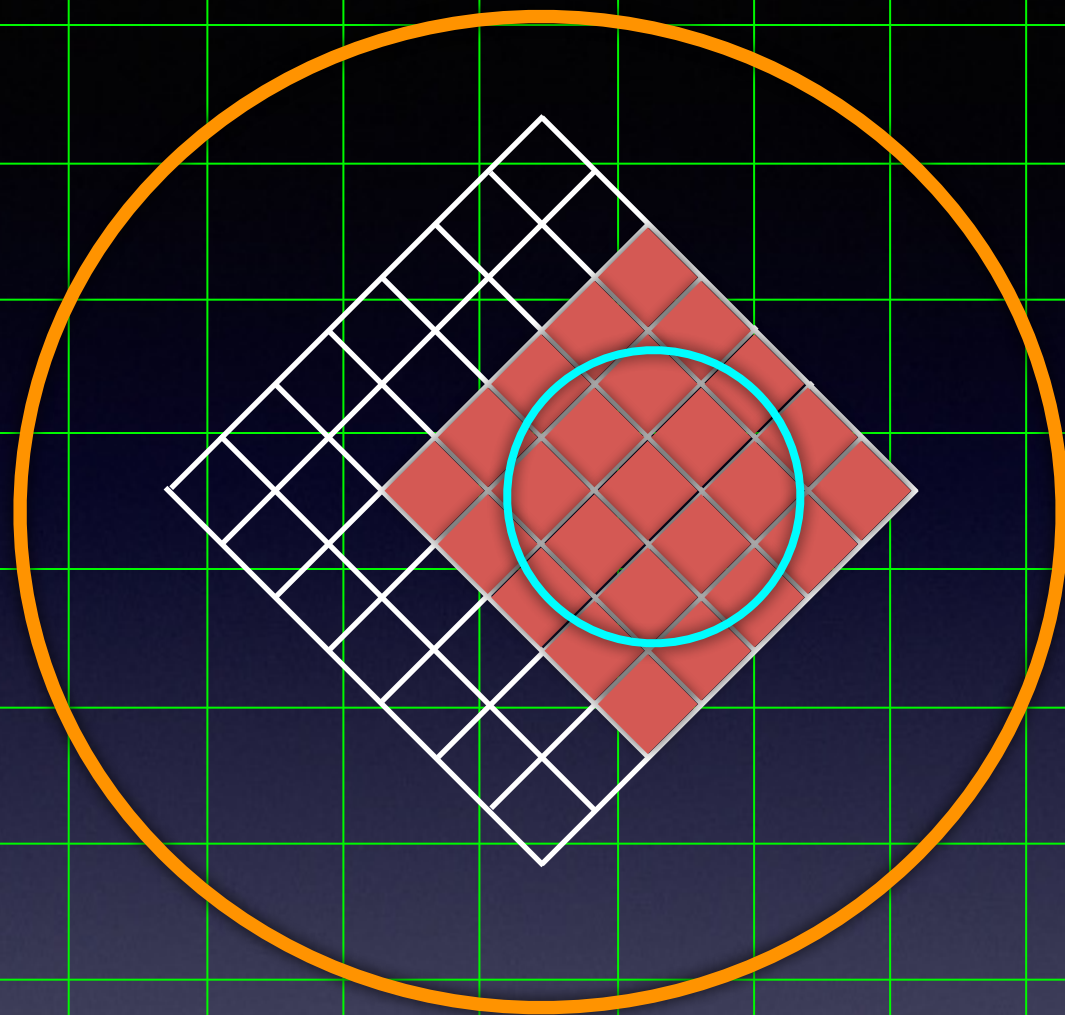
Considering cutouts simple to complex

- Data Discovery
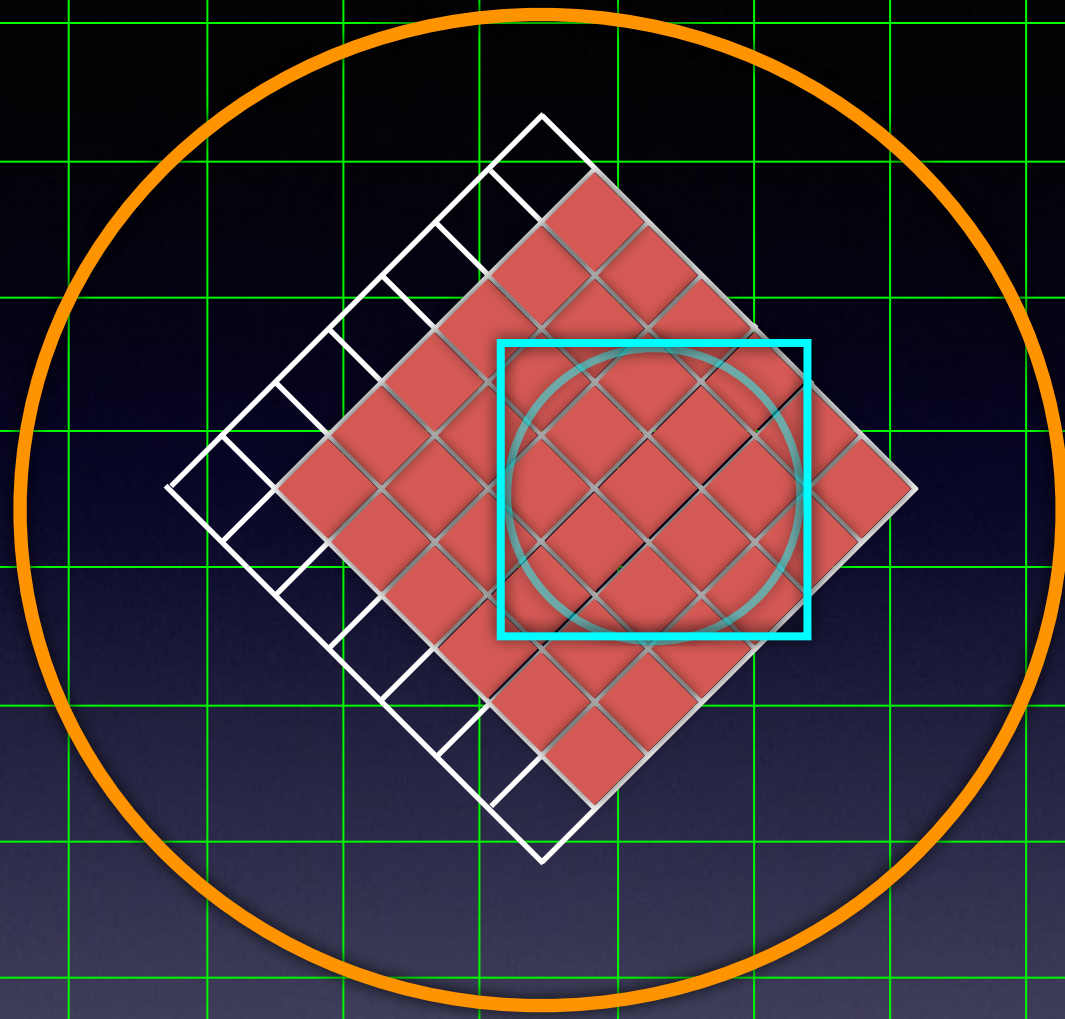
- Available Data
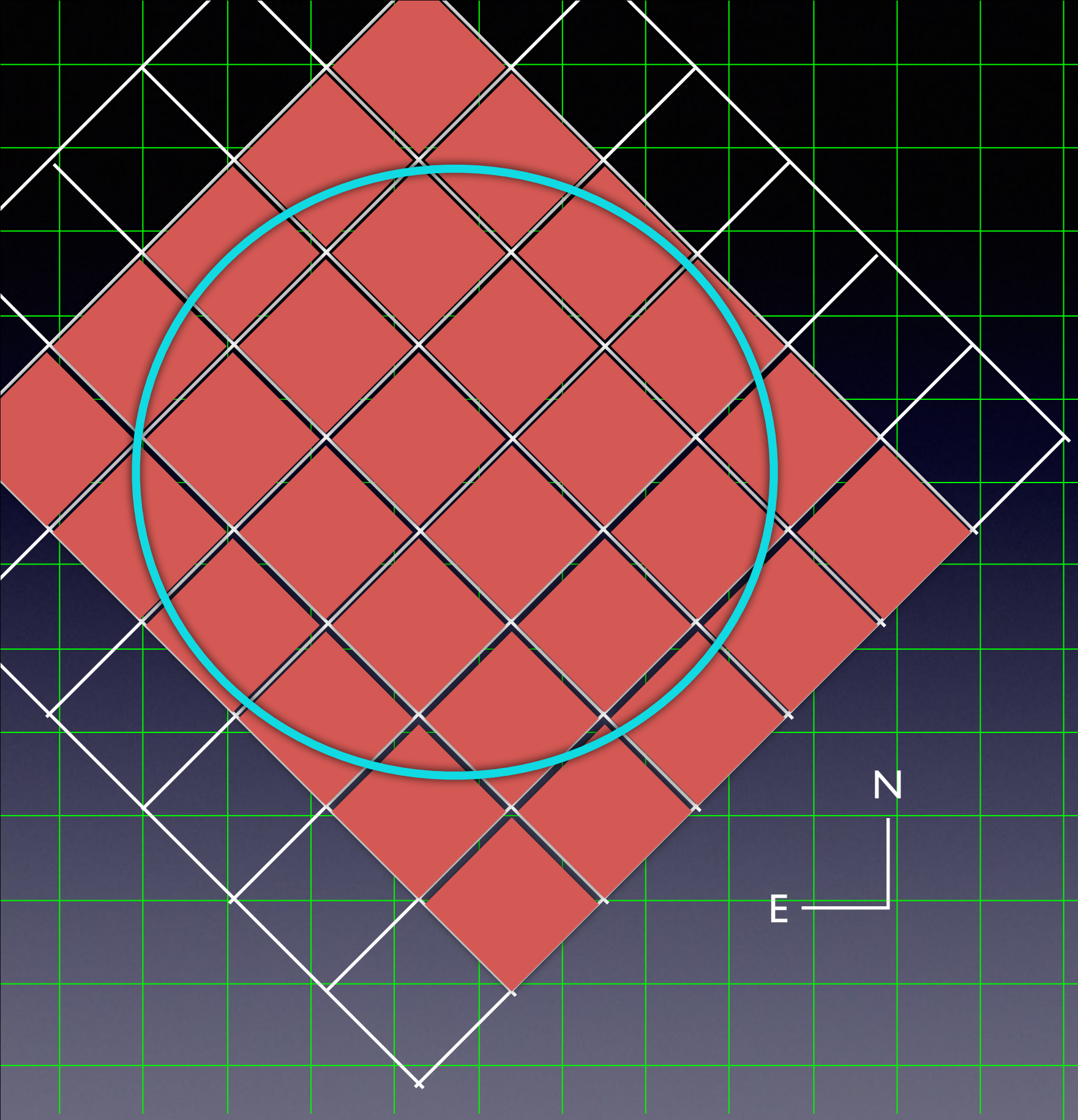
- Download

- Data Discovery

- Available Data

- Download

N

E

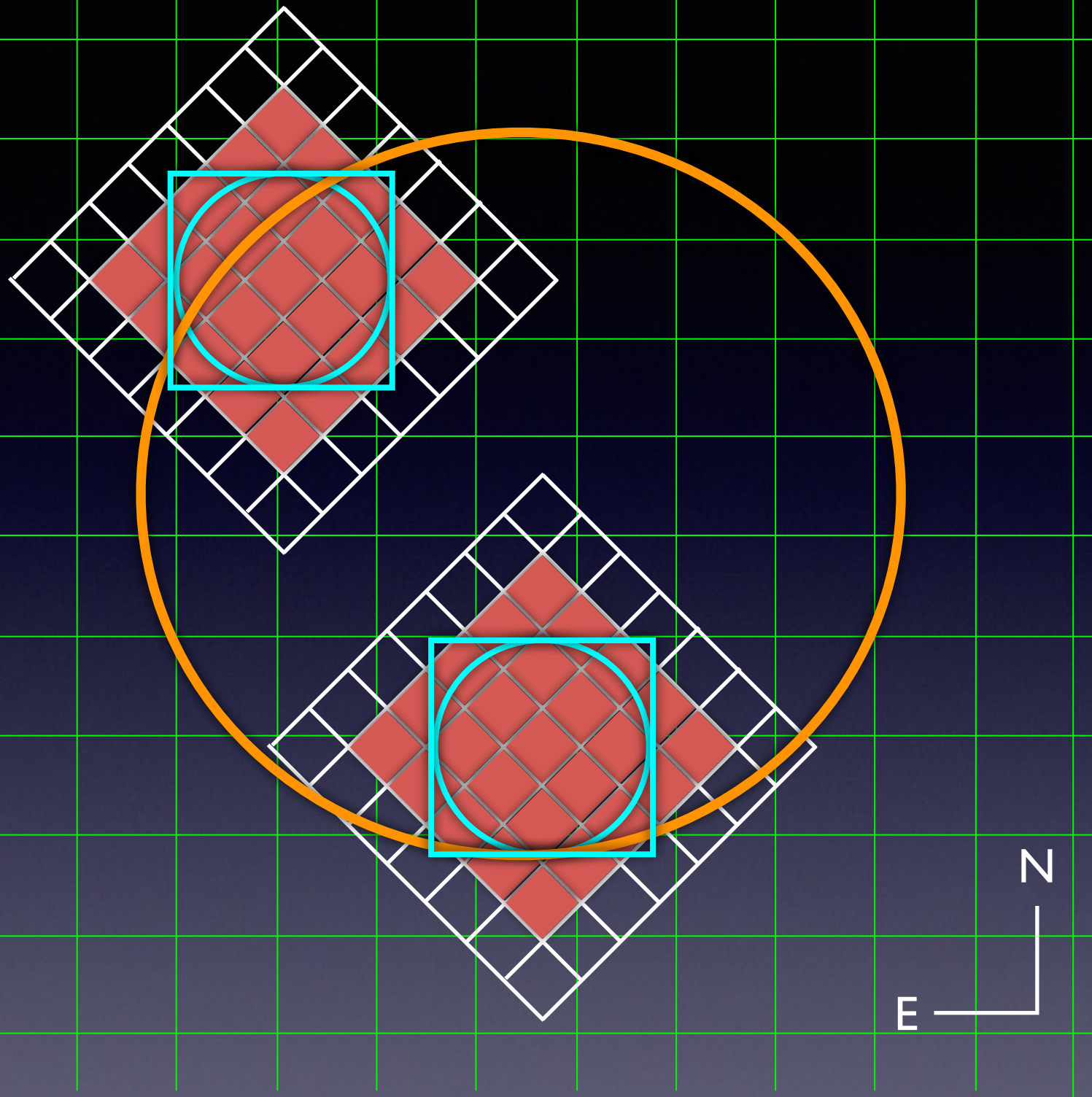- Data Discovery
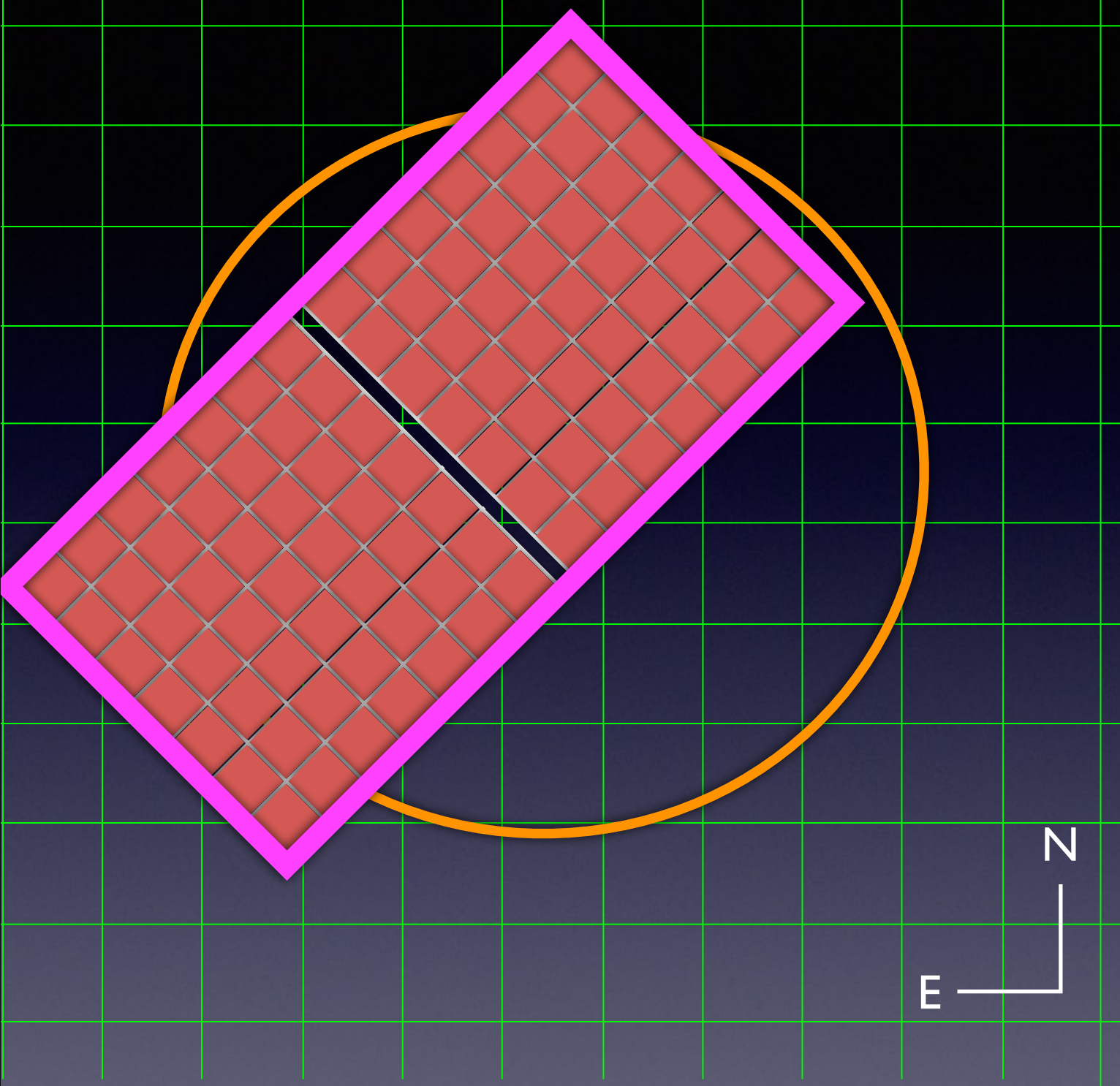- Available Data
- Cutout (circle)
- Download

- Data Discovery
- Available Data
- Cutout (rectangle)
- Download

- **Data Discovery**
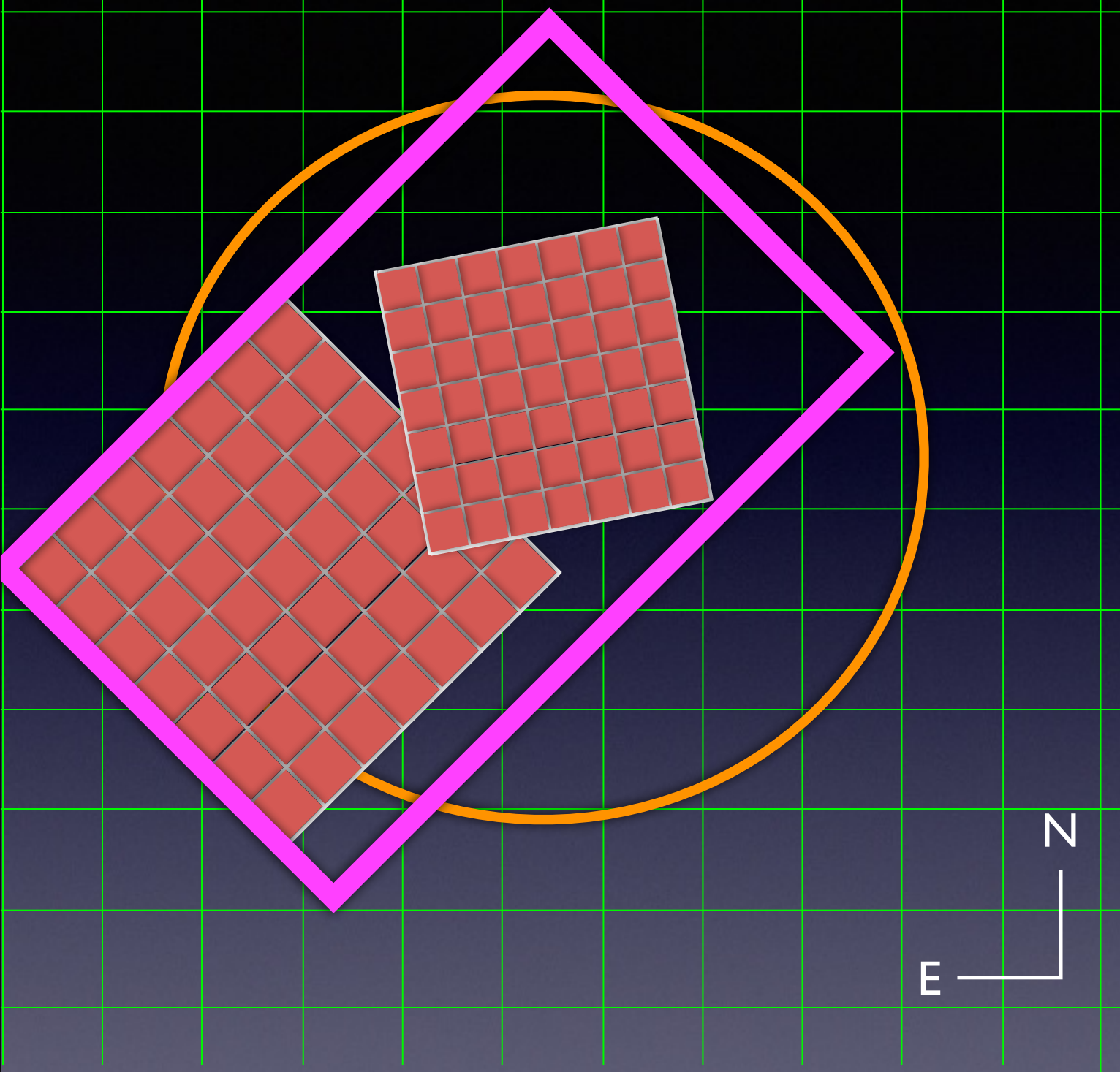
- **Available Data**

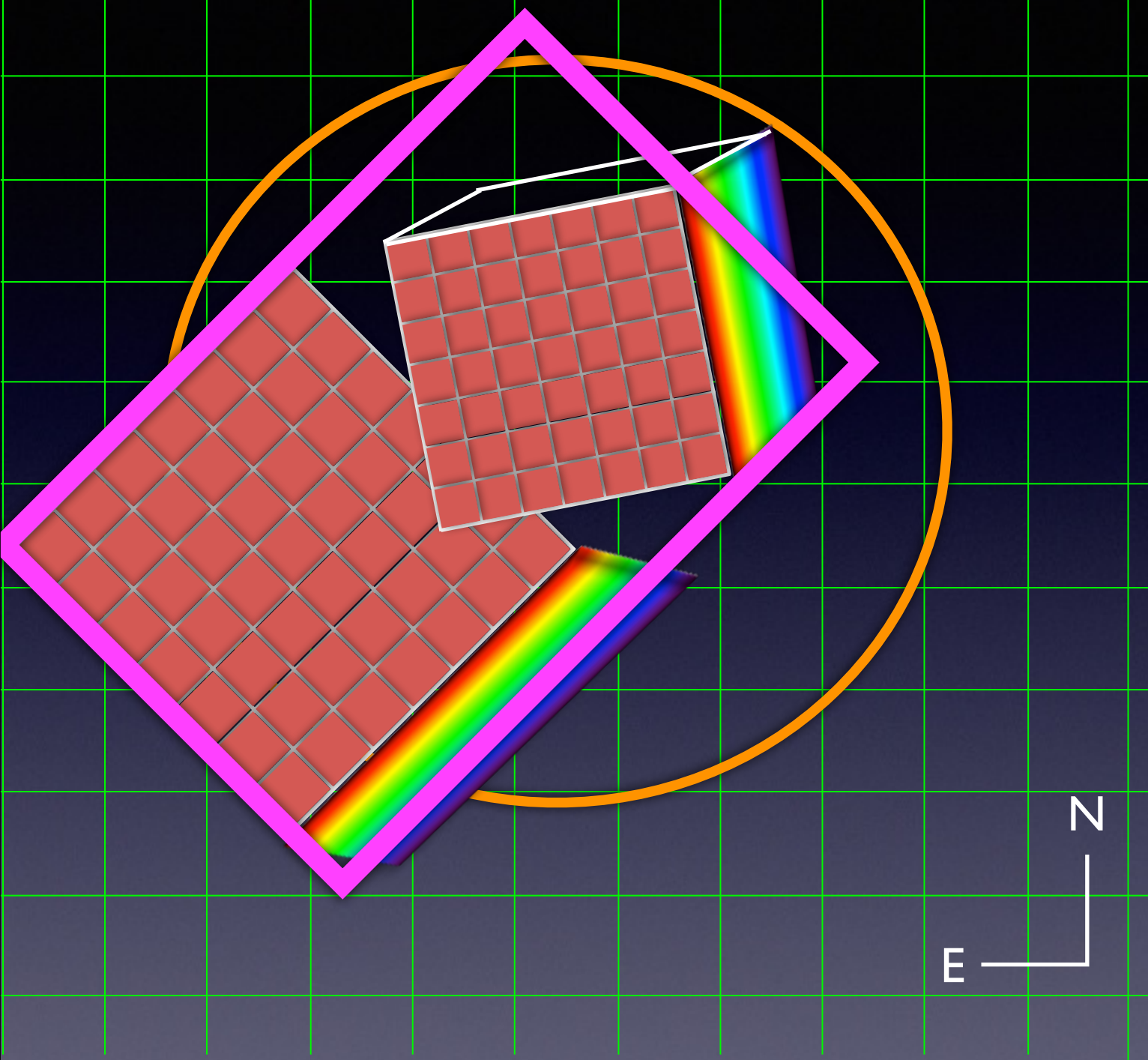- **Cutout = Data discovery cone**

- **Download**

- Data Discovery

- Available Data

- Cutout (circle) - what's valid?

- Download

- Data Discovery

- Available Data

- Simple Mosaiced product

- Download

N

E

- Data Discovery

- Available Data

- More complex mosaiced product

- Download

- Data Discovery
- Available Data
- More complex mosaiced product
- Download

Focus Session Madrid

The topics of this focus session at the 2014 May IVOA Interoperability Meeting include

- Summary of the progress towards the protocols and prototype implementations.

- Identification of the next steps in the widespread adoption of the IVOA standards and services.

- Description of the additional functionality or enhancements likely to be needed. What are the highest priorities?

- Invited presentations / Panel Discussion

- Follow-on WG sessions and break-outs

expected topics

- Address the 'rough consensus and running code' needs of projects

- Implementation schedule

- Implementation assistance - what form?

- Tools - SAMP enable cube viz. tools

- Remote Visualisation goals