BROCKMANN CONSULT

SNAP goes DataCubes

Marco Peters, M. Böttcher, C. Brockmann (Brockmann Consult)

C. Cara (CS Romania), L. Veci (SkyWatch) , F. Douziech (CS France)

M. Engdahl, S. Plummer, E. Volden (ESA)

ESA Living Planet Symposium 2022, Bonn, Germany

Overview

BROCKMANN CONSULT

SNAP

- **BEAM** (est. 2001) - ESA toolbox for the optical sensors on Envisat

- **NEST** (est. 2008) - ESA SAR toolbox and build on top of BEAM

- **SNAP** (est. 2014) - ESA started the new toolbox development for the upcoming Sentinel platforms

  - Development on a **common base**
  - SNAP leverages on the **heritage of BEAM and NEST**

- SNAP is build on **20 years of experience** in EO software development and EO data processing & analysis

BROCKMANN CONSULT

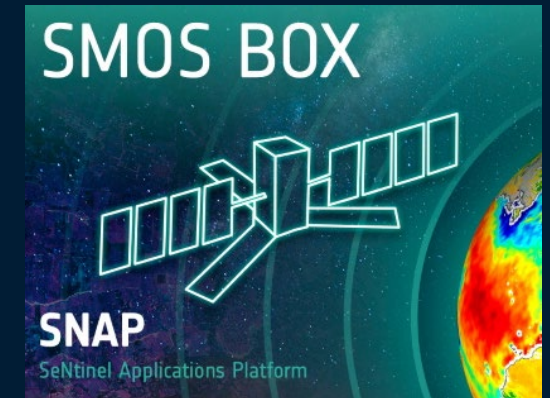SNAP   ARD & Data Cubes   BalticAims
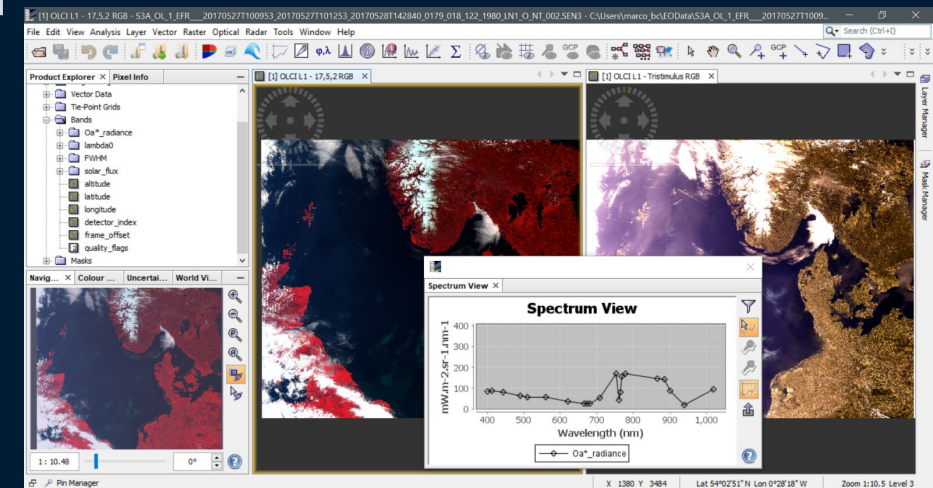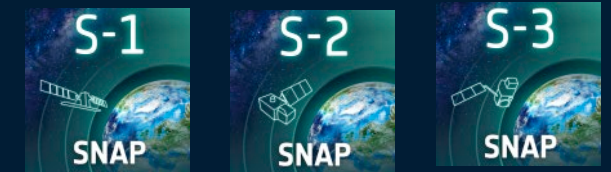
SNAP

- o More Toolboxes
  - Proba-V
  - SMOS
  - Radarsat
  - Chris-Proba
- o SeaDAS (NASA)
  - Built on to of SNAP
  - With own adaptations & extensions

BROCKMANN CONSULT

SNAP   ARD & Data Cubes   BalticAims

SNAP + zarr = cloud & data science

o Traditional Data Formats

- Made for local disk storage (dimap, netCDF, geoTIFF, …)
- Not effective on modern cloud-focused infrastructures (high level of parallelisation, distributed data access)

BROCKMANN CONSULT

SNAP

ARD & Data Cubes

BalticAims

SNAP + zarr = cloud & data science

o Rapidly growing user data science format

- Zarr fits perfectly to Python science data stack – numpy, xarry, dask

- jzarr is the Java implementation of the python zarr API (100% pure java)

- Bridges to openEO - an open API to connect R, Python, JavaScript and other clients to big Earth observation cloud back-ends in a simple and unified way

BROCKMANN CONSULT

SNAP   ARD & Data Cubes   BalticAims

SNAP + zarr = cloud & data science

○ SNAP 9 introduces a new data format which leverages the zarr data format as basis

- storage of chunked, compressed, N-dimensional arrays
- chunk arrays along any dimension
  - important for cloud optimised storage and processing
- compress and/or filter chunks
  - important for cloud optimised storage
- optimised I/O
  - Store arrays in memory, on disk, inside a Zip file, on S3, …
  - Read from / write to an array concurrently from multiple threads or processes
  - Organize arrays into hierarchies via groups

BROCKMANN CONSULT

SNAP   ARD & Data Cubes   BalticAims
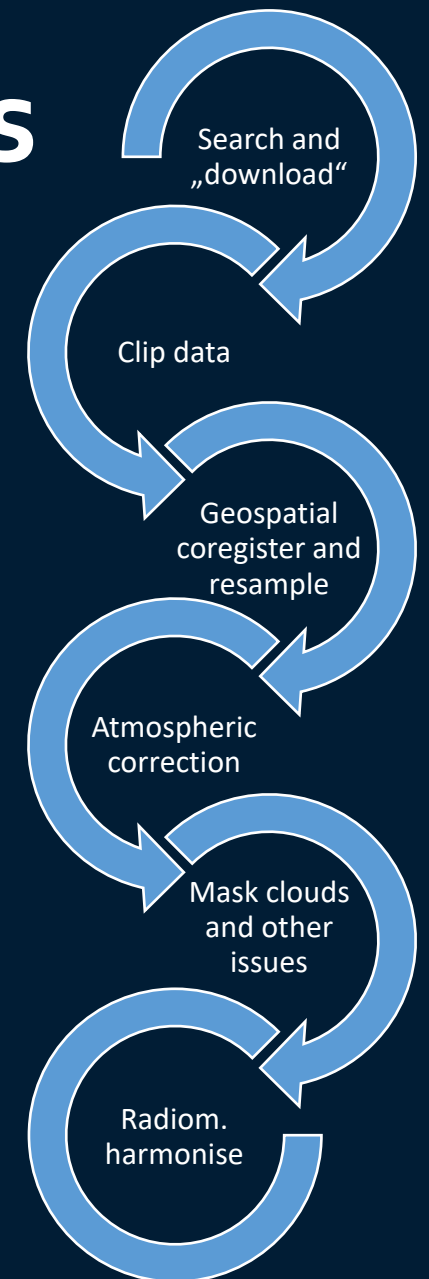
Analysis Ready Data & Data Cubes

Analysis Ready Data
Holmes, 2018

- Definition
  - time-series stacks of overhead imagery that are prepared for a user to analyse without having to pre-process the imagery themselves
- Wider context: cloud native geoprocessing
  - handle more of the data preparation work traditionally done by end users in their desktop environment
  - Benefit from data local processing in a cloud environment

PrePocessing Steps
Holmes 2018

Search and „download"

Clip data

Geospatial coregister and resample

Atmospheric correction

Mask clouds and other issues

Radiom. harmonise

SNAP

ARD & Data Cubes

BalticAims

BROCKMANN CONSULT

Analysis Ready Data & Data Cubes

CEOS Analysis Ready Data (CEOS-ARD)

o … are satellite data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets

   o CEOS Product Family Specification includes requirements on
   - General Metadata
   - Per-pixel Metadata
   - Radiometric and Atmospheric Corrections
   - Geometric Corrections

Analysis Ready Data & Data Cubes

- **Ready to use** data in space and time
- **Easy access** to multidimensional data sets
- **Easy analysis** of parameters in space and time
- **Easy visualisation** of time slices, time series,
- temporal statistics, …
- Data cubes support CEOS-ARD as well as **thematic space-time gridded data** (**"IRD – Interpretation Ready Data"**)
- **zarr format** optimally supports data cubes
- **xcube** open source EO datacubes based on zarr

BROCKMANN CONSULT

xcube

SNAP   ARD & Data Cubes   BalticAims
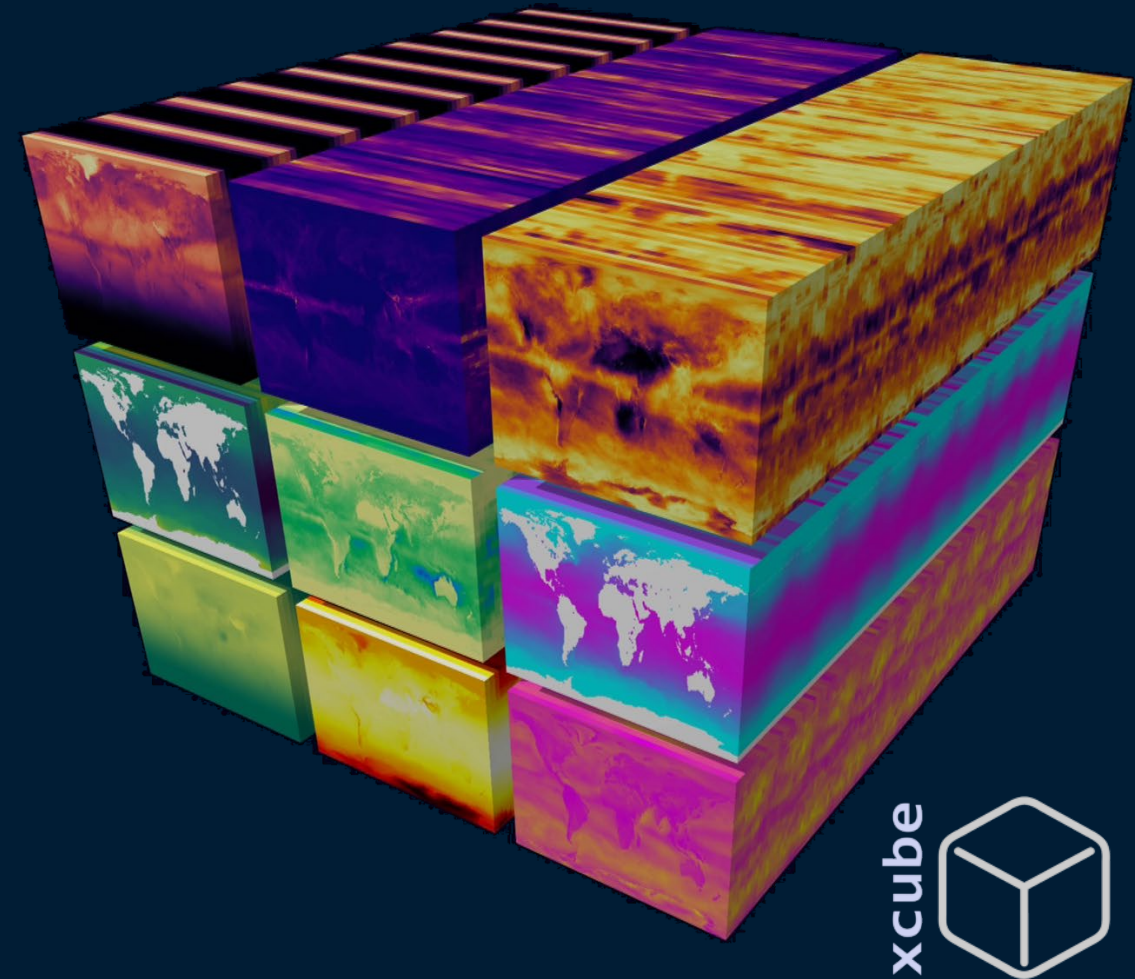
BalticAIMS – Example Application

o BalticAIMS = a data cubes based service supporting marine, coastal and land application in the Baltic Region (ESA Regional Initiative)

o Distributed processing in different processing centres

o Cubes are generated and merged in the cloud
- SNAP operators integrated into operational production systems
- Merging of EO information, Copernicus data, public data sources and user supplied data in regional data cubes
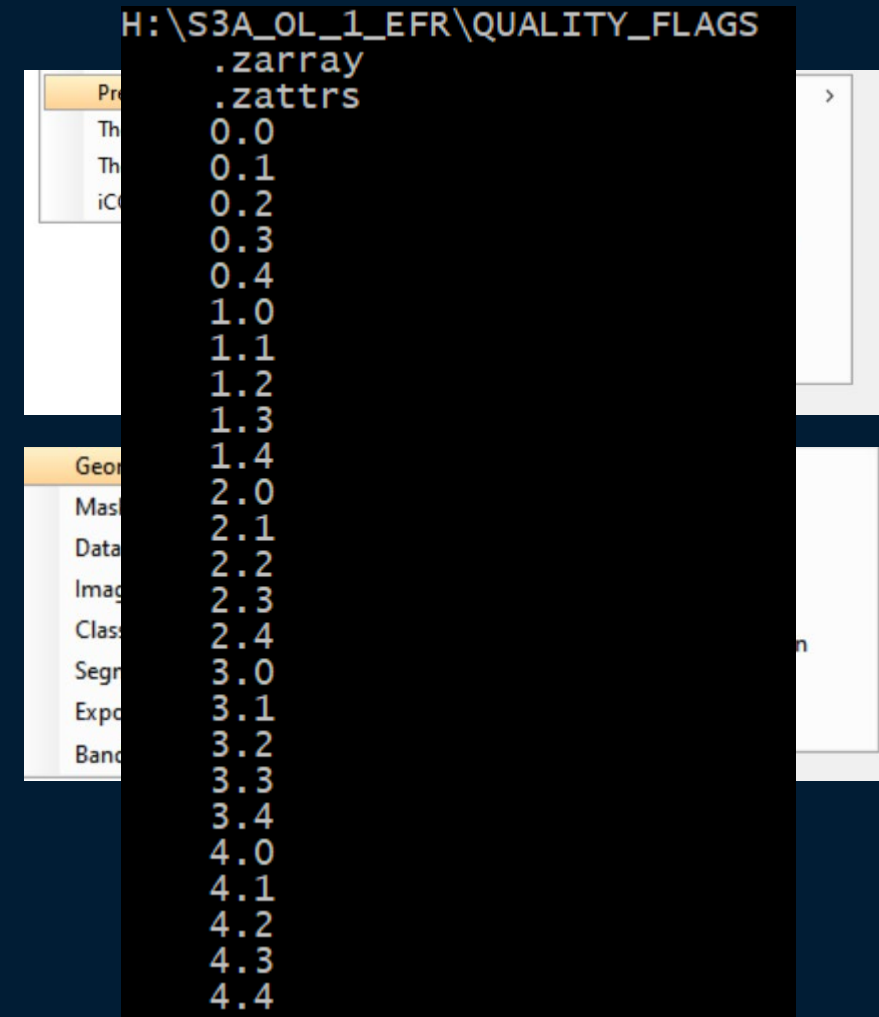- Integration of data cubes into user's tools (viewer, GIS)

SNAP   ARD & Data Cubes   BalticAims

BalticAIMS – Cube creation with SNAP

o Data processing with SNAP OPTICAL, RADAR and MASKS operators
  - Pre-Processing
    – Instrument specific corrections
    – Atmospheric correction (optical), denoising (SAR) , …
    – Pixel quality masks (cloud masking, …)
    – …
  - Thematic processing
    – Land vegetation parameters, chlorophyll concentration, wind, waves, …

o Spatio-temporal processing with SNAP rasters operators
  - Resampling, collocation, Level 3 binning …

o Writing zarr files as data cubes
  - Zarr writer with xcube conventions
  - Repeatedly adding time slices to a data cube
  - Bulk processing and concurrent composition of a data cube

H:\S3A_OL_1_EFR\QUALITY_FLAGS
.zarray
.zattrs
0.0
0.1
0.2
0.3
0.4
1.0
1.1
1.2
1.3
1.4
2.0
2.1
2.2
2.3
2.4
3.0
3.1
3.2
3.3
3.4
4.0
4.1
4.2
4.3
4.4

BROCKMANN CONSULT

SNAP   ARD & Data Cubes   BalticAims
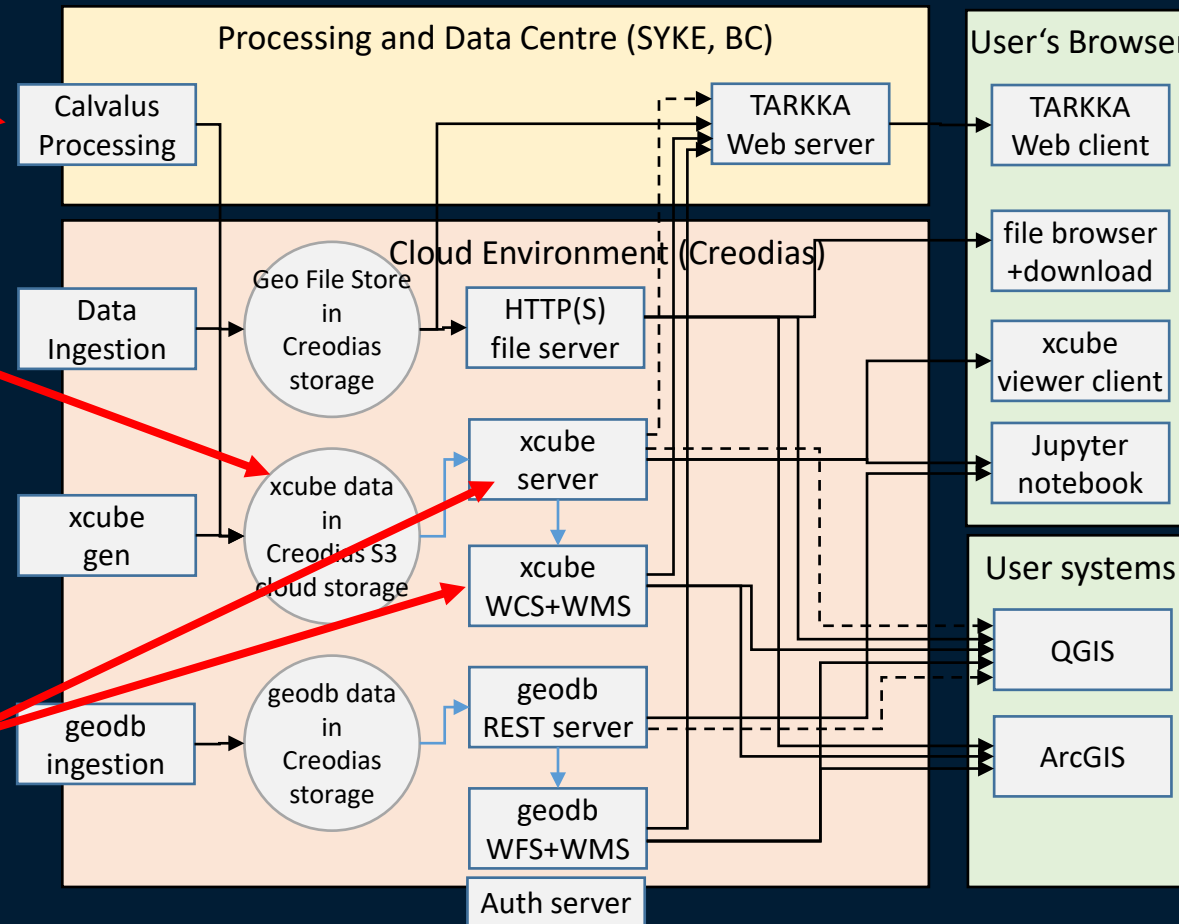○○○○○   ○○○   ○◐○○

BalticAIMS - System Architecture

SNAP operators to process EO data, zarr output

SNAP zarrs integrated with other data, which are generated with xcube-gen from other sources, such as Copernicus Services. Operated on cloud system.

All zarr cubes are published via xcube services

Access to information in zarr cubes in various clients, provdided by BalticAIMS service as well as integrated into user's systems

Processing and Data Centre (SYKE, BC)

Calvalus Processing

TARKKA Web server

Cloud Environment (Creodias)

Data Ingestion

Geo File Store in Creodias storage

HTTP(S) file server

xcube gen

xcube data in Creodias S3 cloud storage

xcube server

xcube WCS+WMS

geodb ingestion

geodb data in Creodias storage

geodb REST server

geodb WFS+WMS

Auth server

User's Browser

TARKKA Web client

file browser +download

xcube viewer client

Jupyter notebook

User systems

QGIS

ArcGIS

BROCKMANN CONSULT

SNAP   ARD & Data Cubes   BalticAims

BalticAIMS - Use Case Example

o HELCOM PLC subgroup (Pollution Load Compilations)
- investigates influence of nutrients on the Baltic Coastal Waters

o Required information:
- land use (EO via Copernicus LMS), river nutrient loads, river discharge, water quality (EO via SNAP processing)

o Operational processing using SNAP graphs in Calvalus
- writing zarr cubes

o xcube processing
- All source data spatial processing into cubes
- Serving zarr cubes via API and WMTS

o Public interfaces TARRKA and xcube Viewer

o Integration into QGIS demonstrating end-user tool integration
- German UBA as HELCOM PLC member
- Jupyter NB demonstrated, raised interest

BROCKMANN CONSULT

SNAP   ARD & Data Cubes   BalticAims
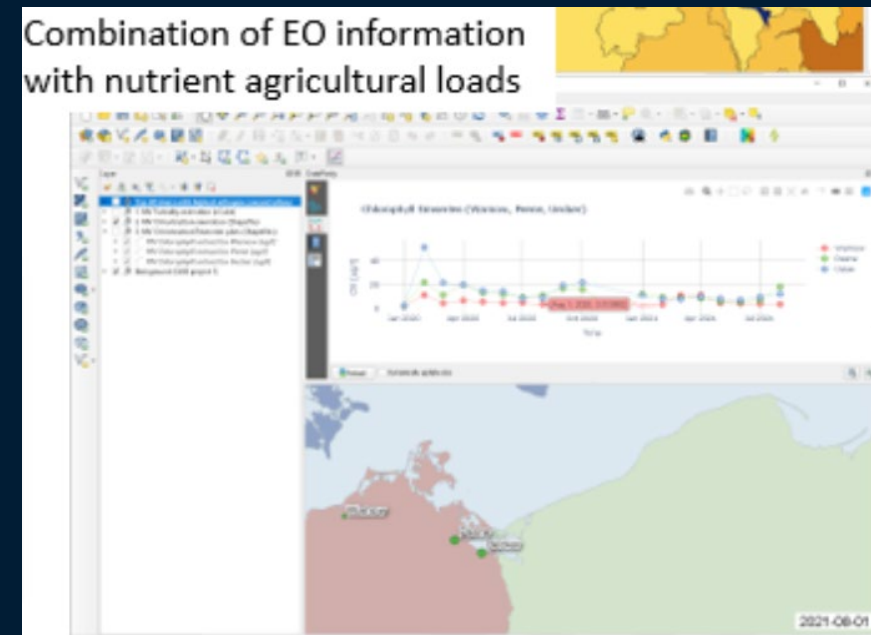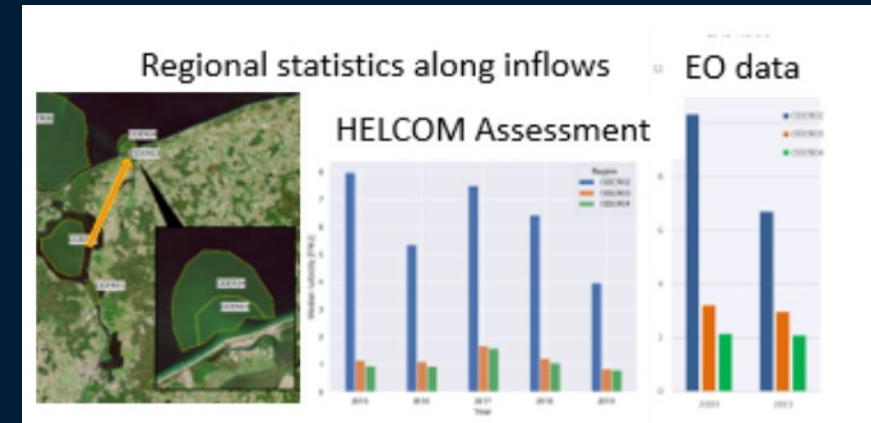
Summary

o   zarr is a format for storage of chunked, compressed, N-dimensional arrays, optimised for cloud applications

o   SNAP is providing zarr as output format already (trial) and will use it as standard output format in the future

o   zarr is very well suited to support ARD and data cubes in general

o   Prototype applications are successful

o   Two development lines for SNAP require future work
    1. SNAP is generating cubes
    2. SNAP uses cubes as input

Hvala   Gracias   Aitah   Obrigado

Takk   Koszonom   Thank you

Merci   SNAP
SeNtinel   Tak

Applications   Dziękuję   Blagodaram
Platform

Kiitos   Grazie   Danke

ευχαριστώ

BROCKMANN CONSULT