Project 56: Data Management

The realities of managing the data from 55 different projects in one national oceanographic program

Taco de Bruin, Leandro Ponsoni, Jennifer Welsh, Janine Nauw, Wil Sistermans,
Ronald de Koster, Jan Nieuwenhuis, Thomas Richter, Marten Tacoma, Richard Brink

NIOZ Royal Netherlands Institute for Sea Research (The Netherlands)

One National Program

During 2008 – 2015, the Dutch National Ocean and Coastal Research Program (Dutch acronym: ZKO) was carried out.

The scientific goals of the ZKO program focused on
- strengthening our understanding and knowledge of coastal development
- the role of biogeochemical cycles and material fluxes in relationship to water quality
- carrying capacity for a sustainable yield of the ecosystem
- changes in biodiversity
- the role of the sea and oceans in climate change
- the effects of climate change on the marine system

(from Final evaluation report of the ZKO program (Aug 2016))

Four Themes

The ZKO programme was divided into 4 different themes or regions:
- Changing Carrying Capacity (Wadden Sea area)
  - Line 1 - Policy-relevant Research
  - Line 2 - Monitoring
  - Line 3 - Hypothesis-driven Research
- North Sea
- Oceans
- Transnational Wadden Sea research (in cooperation with Germany)

Fifty-five Research Projects

Very diverse topics, with project ranging from *Science-policy interactions concerning the designation, management and use of Natura 2000 sites* and *Subtidal survey of selected bivalve species in the Western Wadden Sea* to *Benthic-Pelagic Coupling as Driver of Anomalies of Alkalinity, CO2 and pH in coastal seas* and *INATEX: Observations in the southwest Indian Ocean*

Countless Data-types

Data as varying as videos from foraging fish in the ocean to foraging birds on tidal flats. CTD and radar data. Primary production data and turnover rates of prokaryotes. Model data, observational data and reports from literature studies. Countless parameters such as alkalinity, CO2, pH, temperature, salinity, current velocity, biomass, sediment composition, etc., etc.
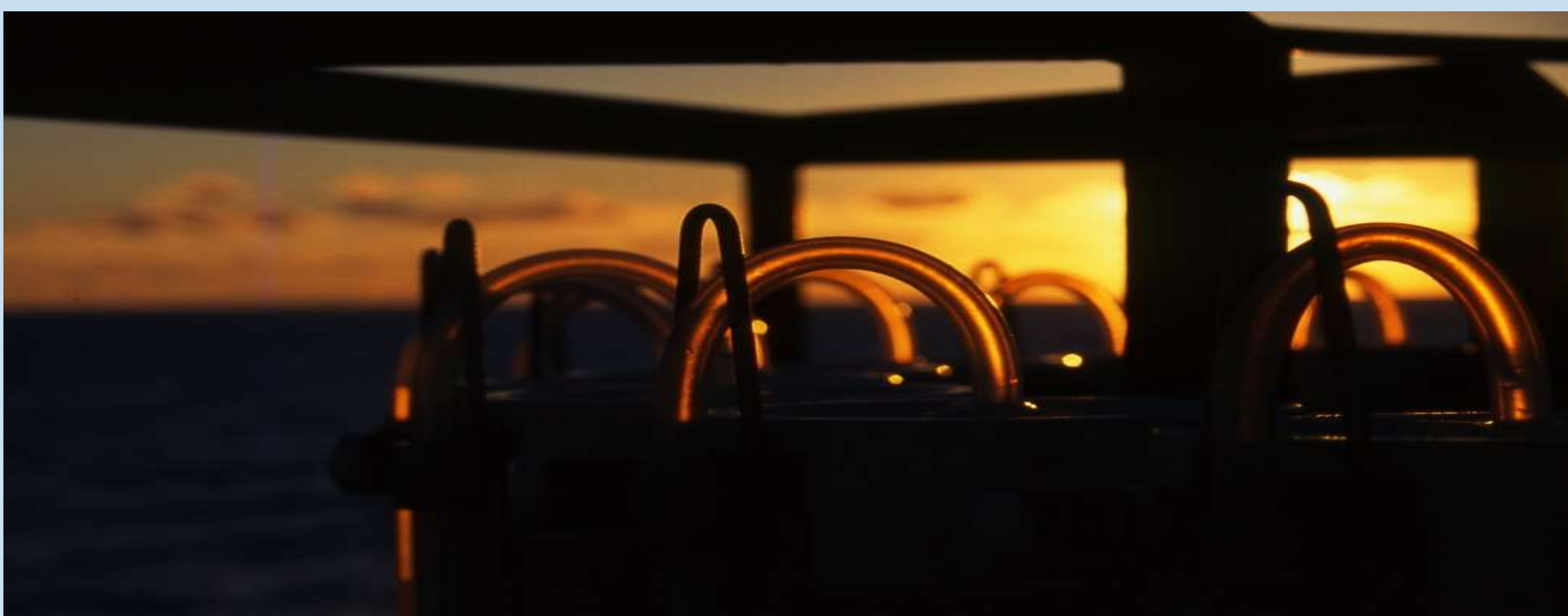
One ZKO Data Policy

Based on the IOC data Exchange Policy and WMO Resolution 40, the ZKO Data Policy in principle requires free and open access to all ZKO data.
During the life time of the ZKO program, the data should be available for all other ZKO projects. After the end of the ZKO program, all data should be freely and openly available.

Three ZKO Data Management Goals

The ZKO program had a Grand Vision of interacting and collaborating ZKO-projects
Three goals were set for ZKO Data Management:

- Archiving of the raw and final data , resulting from ZKO projects
- Exchange of data between projects during the program
- Free and open accessibility of the data after the end of the program

Project 56: Data Management

The ZKO Data Management goals were implemented, through 'project 56' ZKO Data Management
- A central archive to safeguard raw and final data was created
- A data access portal was built
  - The portal contains information on the program, the individual projects and the resulting datasets and publications.
  - All information is linked and can be browsed in various ways
  - Throughout the portal, SeaDataNet standards and vocabs have been used
  - Due to the diversity of the datatypes, the FTP protocol is used for data transfer
  - Publications and computer software can also be downloaded from the portal

URL: http://data.zkonet.nl

Results (and some issues)

- All the data requiring archiving were archived
- A dedicated ZKO data access portal was built and filled with ZKO data
- Almost no projects were interested in exchanging data with other projects
- As a result, data were only submitted after a project had ended
- 47% of the projects provided complete datasets, while another 12% is near completion. 16% of the projects did not collect any observational data (but did produce publications based on literature studies).
  25% of the projects did not deliver any data.
- Reasons for not delivering data are:
  - PI or PhD-student left with data, current data whereabouts unknown
  - PI or PhD-student unresponsive and communication halted
  - Flat refusal to submit data
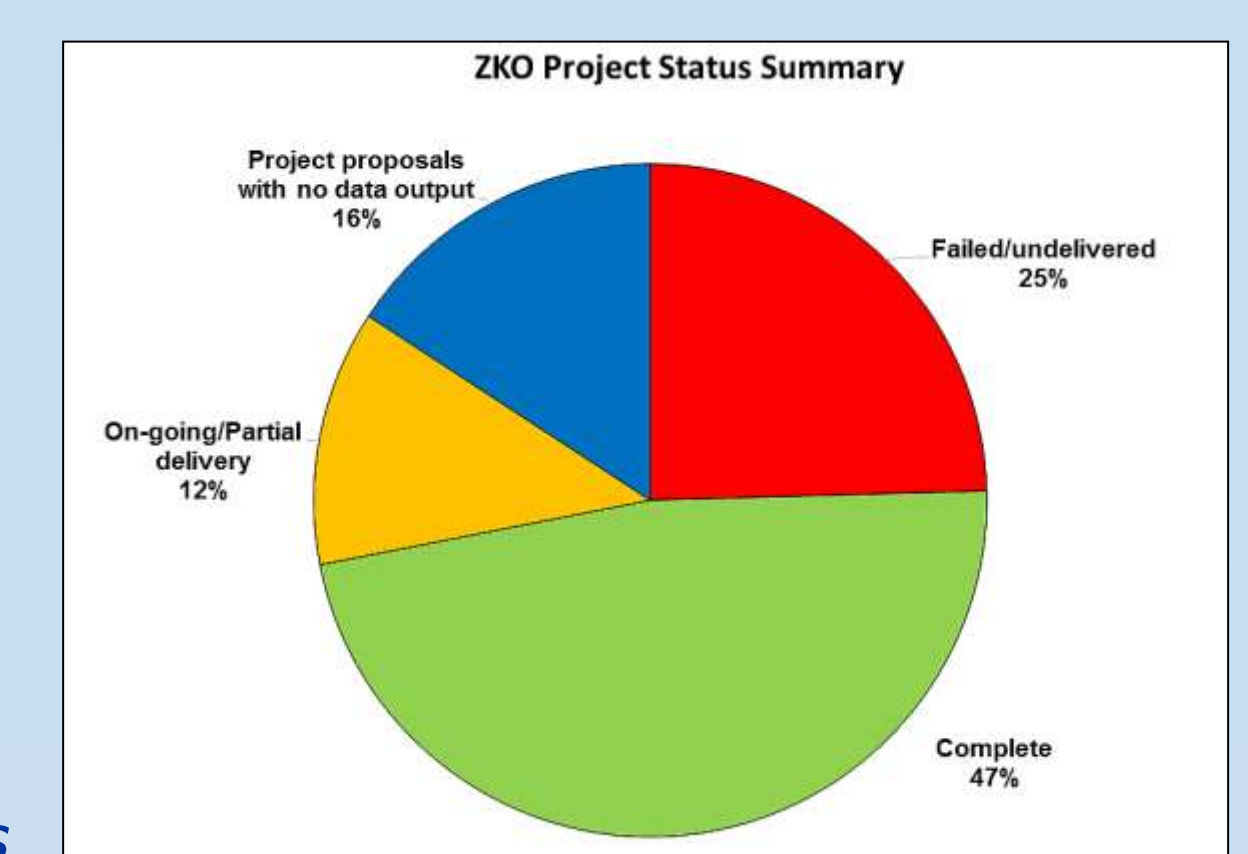
Overview of data submission status of all projects

Conclusions and lessons learned

- The large majority of ZKO data is safeguarded in the ZKO archive and available for (re)use at http://data.zkonet.nl

- Submission of data is not be expected until the very end of a project or even after the project has ended.

- A dedicated data manager is key to getting the data

- Precautions (in the form of a data submission protocol) should be taken to avoid loss of contact with the PI and/or the PhD student, who may have moved on after the project.

- There is a distinct role and responsibility for the funding agency in enforcing its Data Policy

- There is a mismatch in time between data management projects and research projects. While the latter are usually for 4 years, a data management project should start at the planning phase of a research project and continue long after the end of a project