



Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

How we support scientific research at INBO

FLAMES Annual Meeting 2017, KU Leuven

D. Brosens, P. Desmet, D. Du Seuil, I. Jansen, E. Lommelen, J. Loos, T. Milotic, T. Onkelinx, P. Quataert, F. Piesschaert, H. Van Calster, T. Van Daele, F. Vanderhaeghe, G. Van Spaendonck, S. Van Hoey, P. Verschelde, T. Westra, C. Wils

Research Institute for Nature and Forest (INBO)



Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

Who are we?

Instituut voor Natuur- en Bosonderzoek (INBO)

The research institute for Nature and Forest (INBO) is the Flemish research and knowledge center for nature and its sustainable management and use.

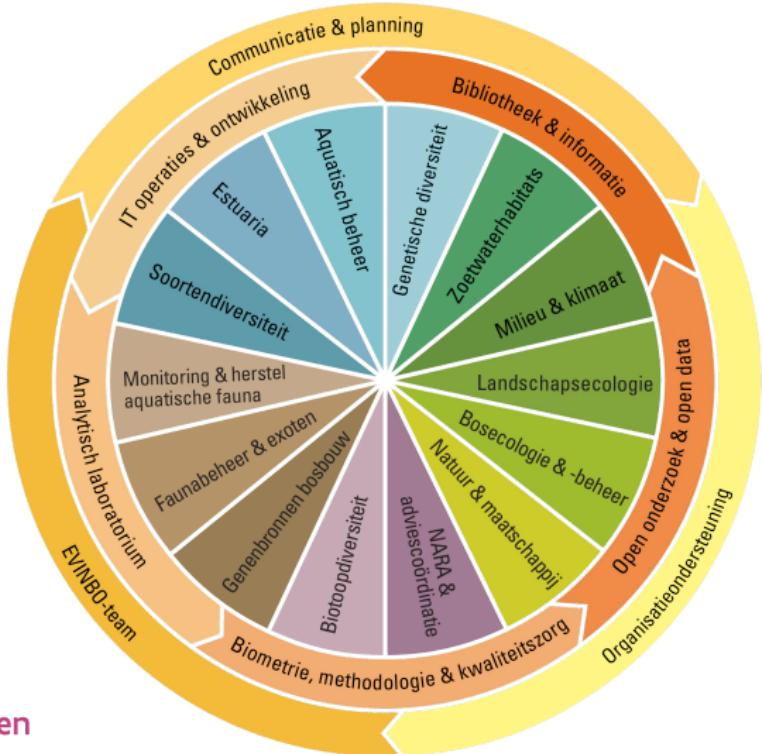
14 teams with scientific staff (191 members)



5 teams in support of scientific staff (40 members)

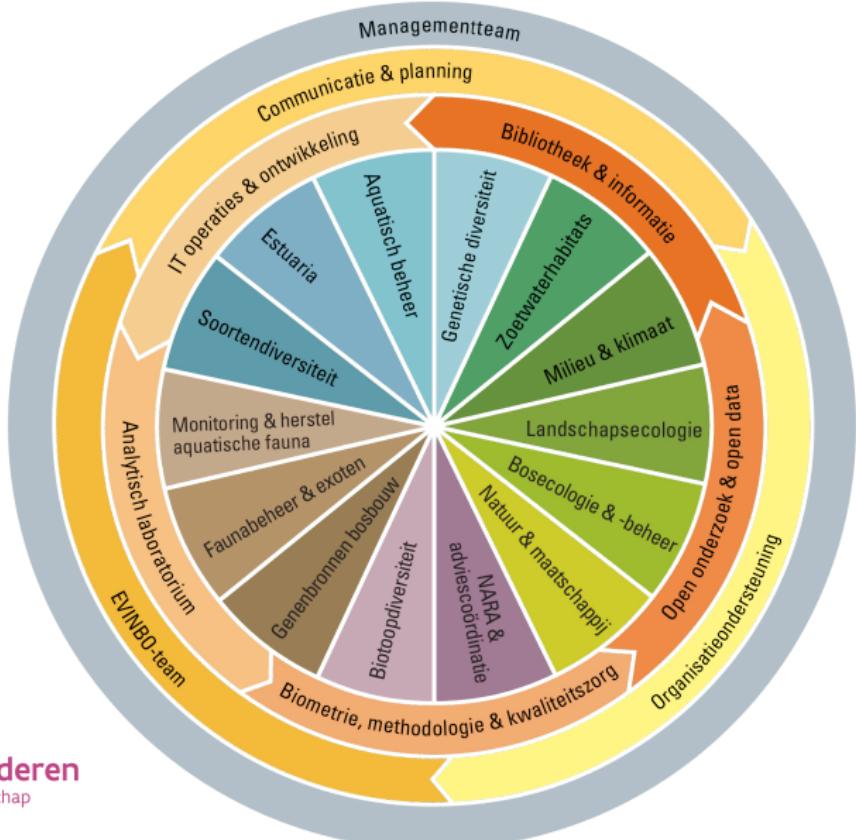


3 administrative teams (22 members)

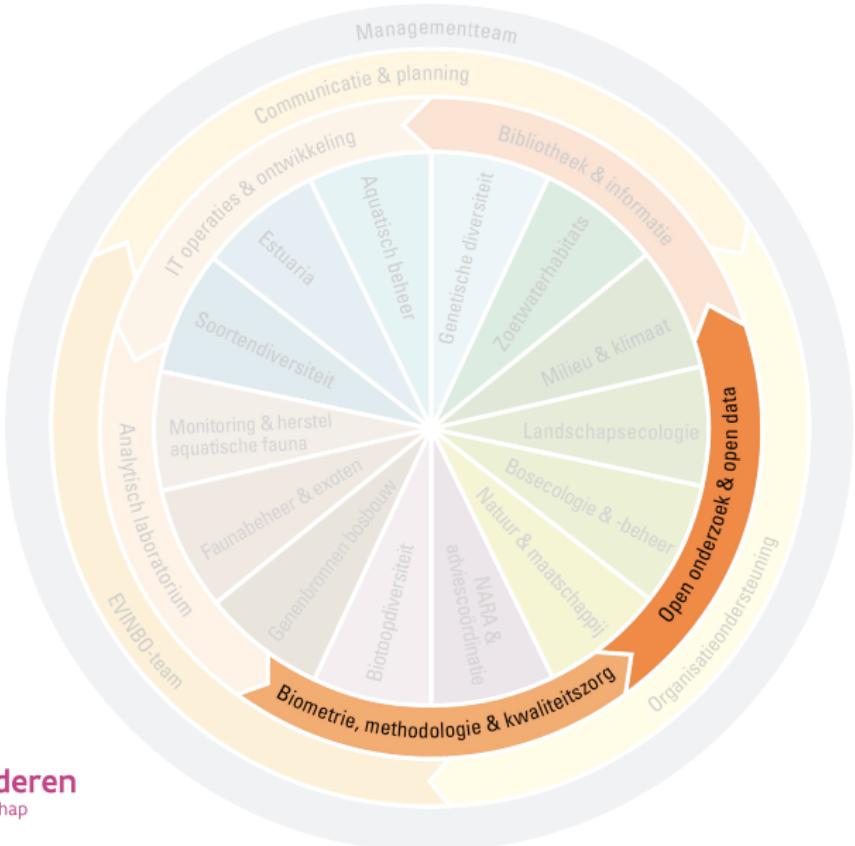


Vlaanderen
is wetenschap

1 management team (9 members)



Vlaanderen
is wetenschap



Peter Diggle paraphrasing Iain Buchan

Informatics seeks to maximize the *utility* of data, whereas statistics seeks to minimize the uncertainty that is associated with the *interpretation* of data.



Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

Open data & open science



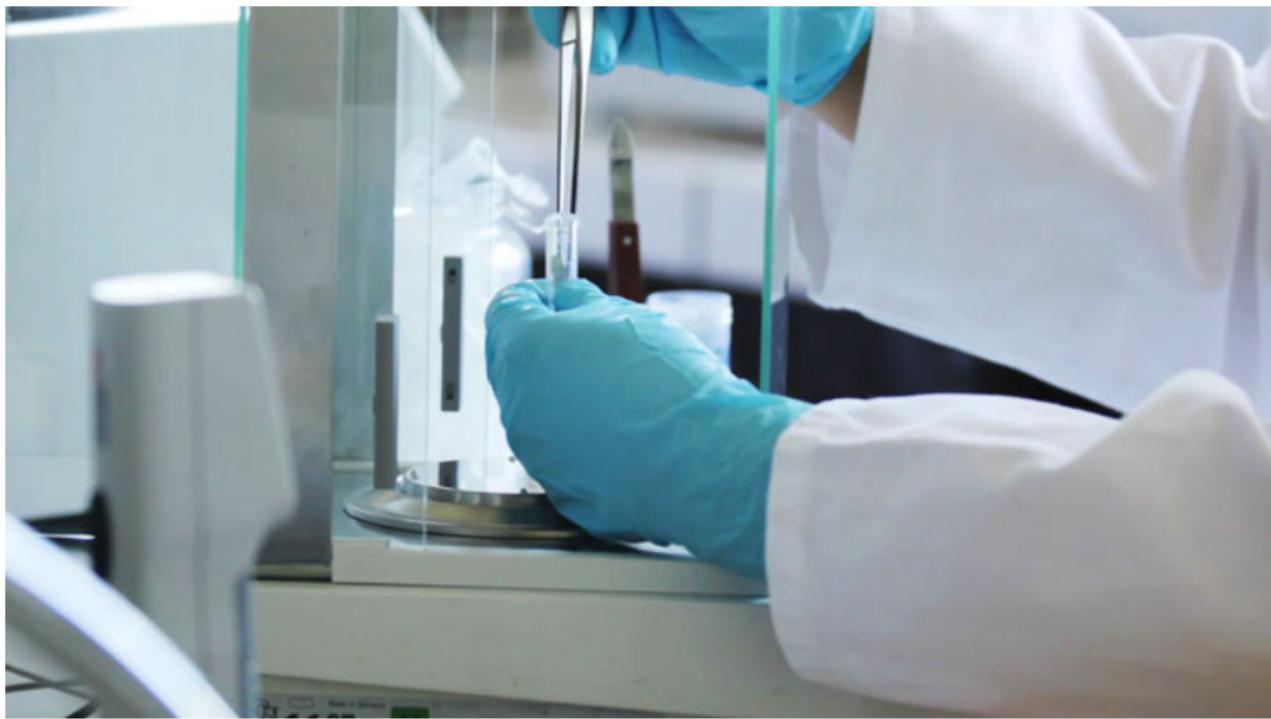
Vlaanderen
is wetenschap



Vlaanderen
is wetenschap



Vlaanderen
is wetenschap





Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

Open data & open science

Support scientific staff in data life cycle

Build and maintain data infrastructure

- ▶ Manage major data systems
- ▶ Setup data warehouses for uniform data access
- ▶ Setup GIS services
- ▶ Design data models for new systems
- ▶ Coordinate development of new data collection tools

Provide research software

- ▶ Data processing tools
- ▶ GIS tools
- ▶ Reporting tools

Provide support

- ▶ Documentation
- ▶ Tutorials
- ▶ Training
- ▶ Ad hoc support



Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

Open data & open science

Open data

INBO is an open data institute

VLAANDEREN

INSTITUUT
NATUUR- EN
BOSONDERZOEK



Route | Nederlands

Search



THEMES ABOUT INBO RESEARCH PUBLICATIONS ALTER-NET EXPERT PORTAL

Home > Open data policy

Open data policy

[Dutch version](#)

Why open data?

In its mission statement, the Research Institute for Nature and Forest (INBO) declares to support decision makers in the development of a sustainable society by providing insight in nature and forest through qualitative research. This mission is expressed as set of strategic goals to support governmental initiatives, partners, and other policy makers, and to implement a European biodiversity strategy.

The operational functioning of the institute aims to accomplish these strategic goals by executing scientific research and monitoring. The outcome is communicated in scientific



The Open Definition

Open data and content can be **freely used, modified, and shared by anyone for any purpose.**

Open data

- 1 Has an open license (legal status)
- 2 Is accessible
- 3 Is machine readable
- 4 Is provided in an open format

Why open data for research?

- ▶ Discoverability

Why open data for research?

- ▶ Discoverability
- ▶ Increase potential use

Why open data for research?

- ▶ Discoverability
- ▶ Increase potential use
- ▶ Collaborations

Why open data for research?

- ▶ Discoverability
- ▶ Increase potential use
- ▶ Collaborations
- ▶ Transparent science

Move towards open data

- ▶ EC: Directive on the re-use of public sector information (2003/98/EG & 2013/37/EU)

Move towards open data

- ▶ EC: Directive on the re-use of public sector information (2003/98/EG & 2013/37/EU)
- ▶ Flemish government: open data should become the norm for the Flemish government (*2011 conceptnota open data*)

Move towards open data

- ▶ EC: Directive on the re-use of public sector information (2003/98/EG & 2013/37/EU)
- ▶ Flemish government: open data should become the norm for the Flemish government (*2011 conceptnota open data*)
- ▶ Funding agencies: e.g. Horizon 2020

Move towards open data

- ▶ EC: Directive on the re-use of public sector information (2003/98/EG & 2013/37/EU)
- ▶ Flemish government: open data should become the norm for the Flemish government (*2011 conceptnota open data*)
- ▶ Funding agencies: e.g. Horizon 2020
- ▶ Research journals: deposit data in an open repository

2011: Publication of the first dataset on GBIF



GBIF

Global Biodiversity
Information Facility

Data • News • Community • About •



Florabank1 - A grid-based database o...

Occurrence dataset published by Research Institute for Nature and Forest (INBO)

3,895,635

Occurrences

[View occurrences](#)

Information

Stats

Activity

Summary

FULL TITLE

Florabank1 - A grid-based database on vascular plant distribution in the northern part of Belgium (Flanders and the Brussels Capital region)

DESCRIPTION

Florabank1 is a database that contains distributional data on the wild flora indigenous species, archeophytes and naturalised aliens) of Flanders and the Brussels Capital Region. It holds about 3 million records of vascular plants, dating from 1800 till present. Furthermore, it includes ecological data on vascular plant species, redlist category information, Ellenberg values, legal status, global distribution, seed bank etc. The database is an initiative of "Flo.Wer" (<http://www.plantenwerkgroep.be>), the Research Institute for Nature and Forest (INBO) (<http://www.inbo.be>) and the National Botanic



DOI doi:10.3897/phytokeys.12.2849

PUBLISHED BY

Research Institute for Nature and
Forest (INBO)

PUBLICATION DATE

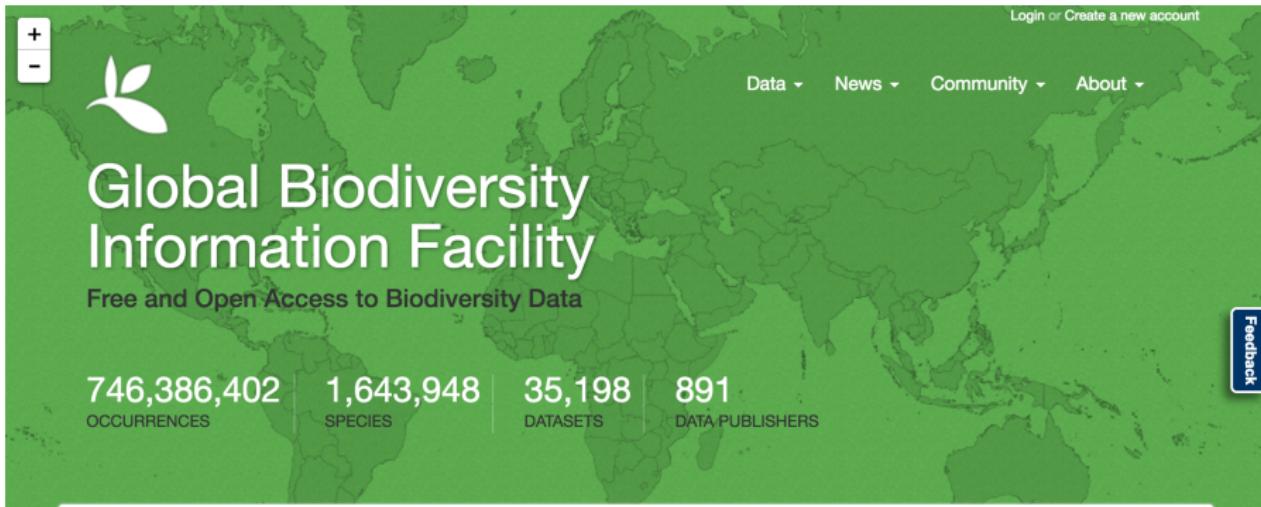
Mar 2 2011



Vlaanderen
is wetenschap

Feedback

Global Biodiversity Information Facility (GBIF)



Sharing biodiversity data for re-use

- [Learn about GBIF](#)
- [Publish your data through GBIF](#)
- [Technical infrastructure](#)

Providing evidence for research and decisions

- [Using data through GBIF](#)
- [Enabling biodiversity science](#)
- [Supporting global targets](#)

Collaborating as a global community

- [Current Participants](#)
- [How GBIF is funded](#)
- [Enhancing capacity](#)



Vlaanderen
is wetenschap

Data publication process

- ▶ Format data to the international Darwin Core standard
- ▶ Document data with metadata (also standardized)
- ▶ Publish dataset online
- ▶ Register dataset with GBIF
- ▶ (Re)publish versioned updates or corrections

2015: INBO open data policy

VLAANDEREN

INSTITUUT
NATUUR- EN
BOSONDERZOEK



Route | Nederlands

Search



THEMES ABOUT INBO RESEARCH PUBLICATIONS ALTER-NET EXPERT PORTAL

Home > Open data policy

Open data policy

[Dutch version](#)

Why open data?

In its mission statement, the Research Institute for Nature and Forest (INBO) declares to support decision makers in the development of a sustainable society by providing insight in nature and forest through qualitative research. This mission is expressed as set of strategic goals to support governmental initiatives, partners, and other policy makers, and to implement a European biodiversity strategy.

The operational functioning of the institute aims to accomplish these strategic goals by executing scientific research and monitoring. The outcome is communicated in scientific



Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions

Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)

Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)
- 3 Aim to publish **result data** (associated with publications)

Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)
- 3 Aim to publish **result data** (associated with publications)
- 4 Publish as **open data** as defined by the Open Definition

Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)
- 3 Aim to publish **result data** (associated with publications)
- 4 Publish as **open data** as defined by the Open Definition
- 5 Publish under **CC0 waiver**

Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)
- 3 Aim to publish **result data** (associated with publications)
- 4 Publish as **open data** as defined by the Open Definition
- 5 Publish under **CC0 waiver**
- 6 Encourage proper data use with **norms**

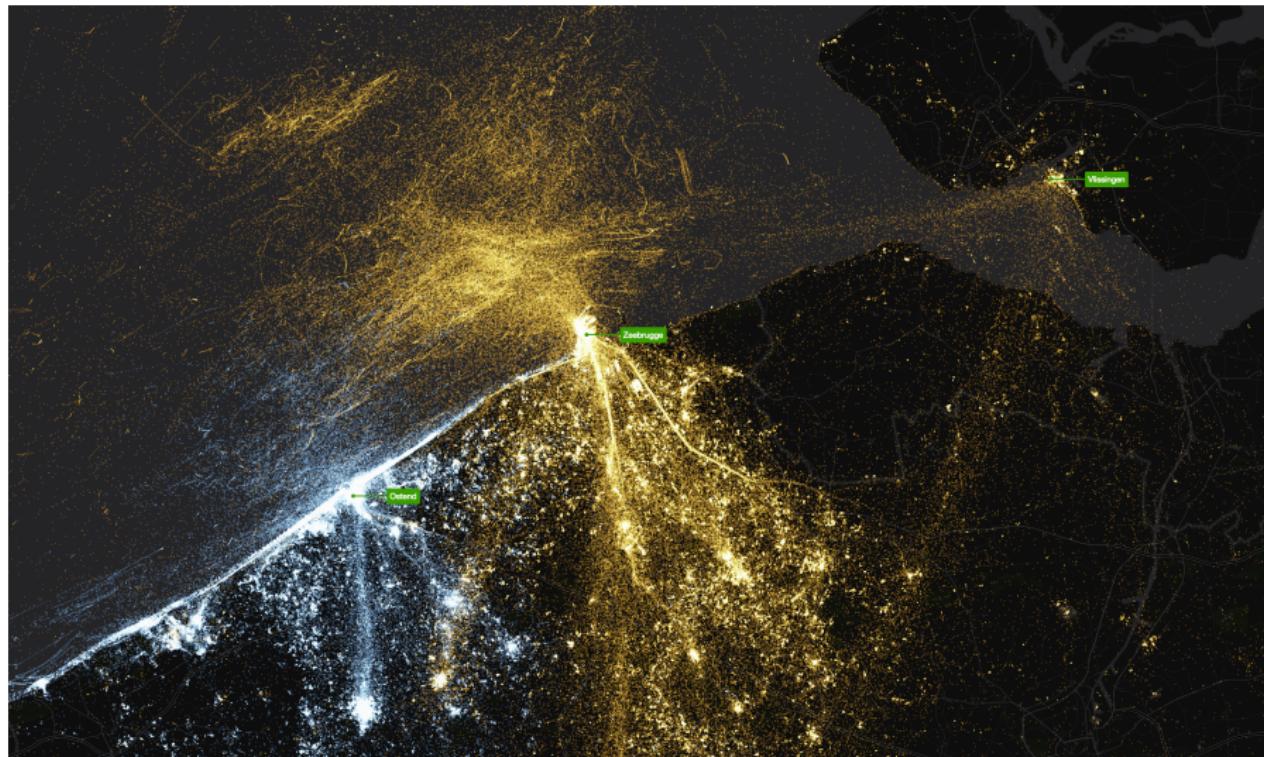
Guidelines of the open data policy

- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)
- 3 Aim to publish **result data** (associated with publications)
- 4 Publish as **open data** as defined by the Open Definition
- 5 Publish under **CC0 waiver**
- 6 Encourage proper data use with **norms**
- 7 Document data with **metadata**

Guidelines of the open data policy

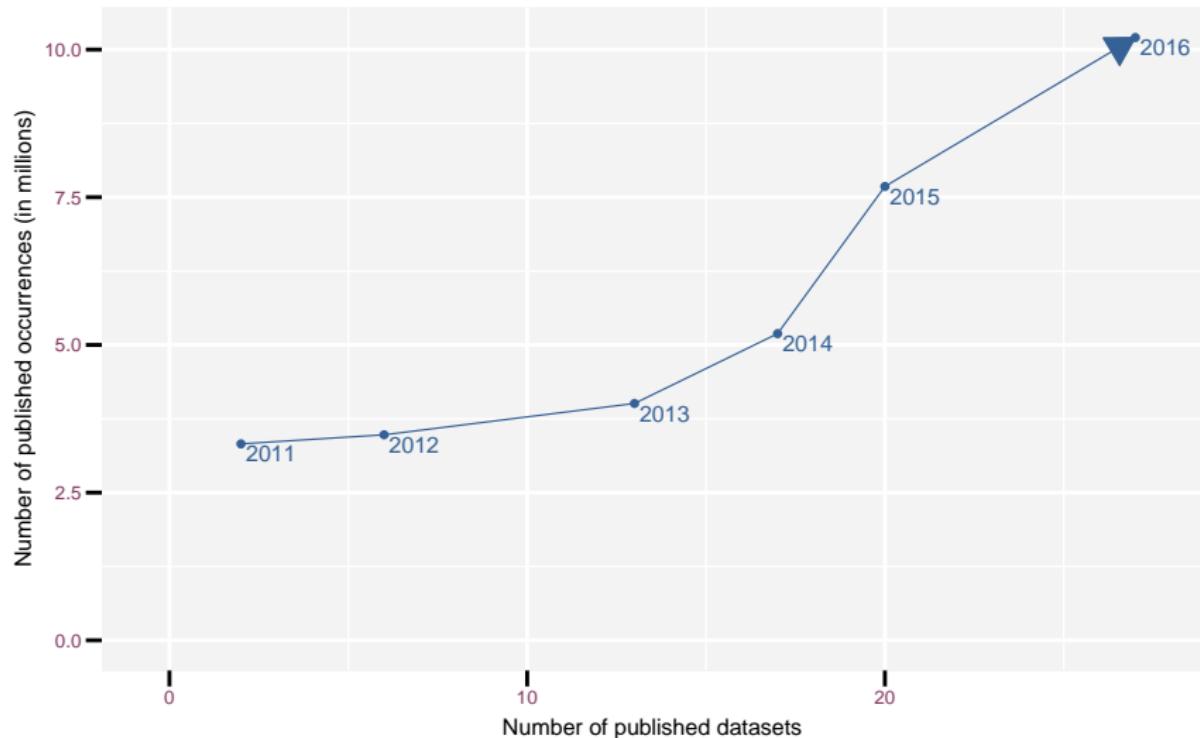
- 1 Policy applies to **data we own** & is bound by legal exceptions
- 2 Aim to publish **collected data** (from major data systems)
- 3 Aim to publish **result data** (associated with publications)
- 4 Publish as **open data** as defined by the Open Definition
- 5 Publish under **CC0 waiver**
- 6 Encourage proper data use with **norms**
- 7 Document data with **metadata**
- 8 Start using **data management plans** (DMP)

We are now one of the major data publishers on GBIF

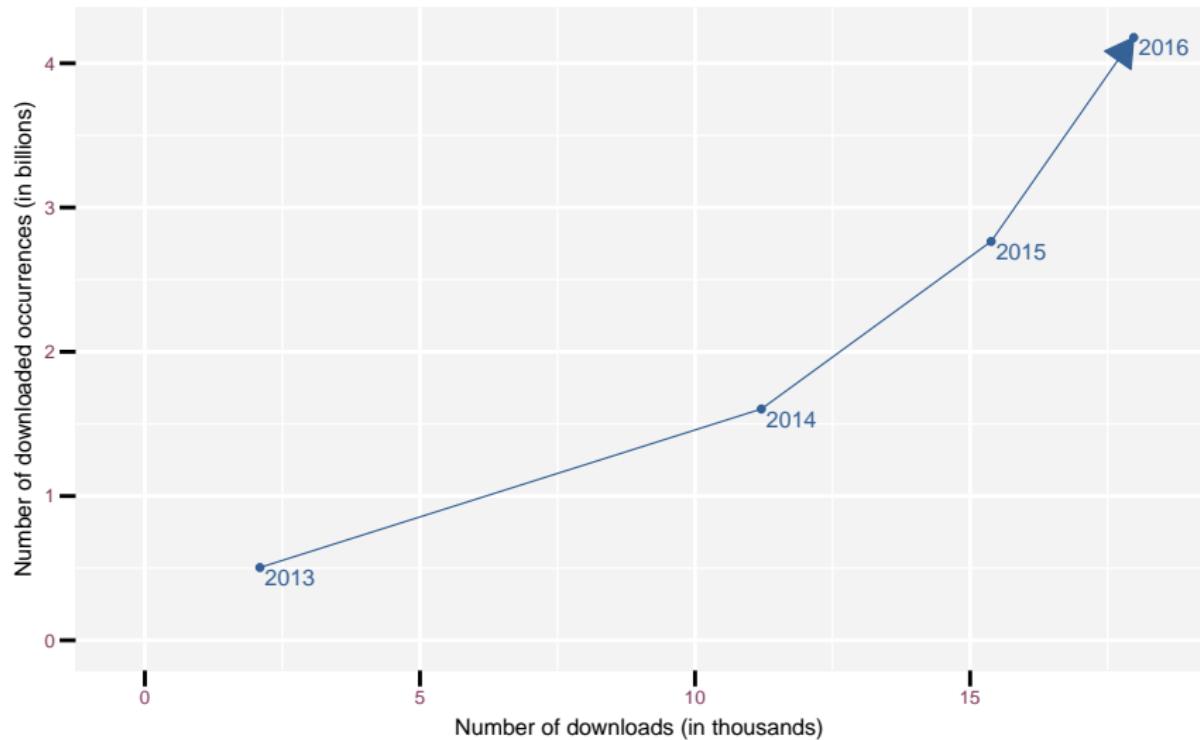


Vlaanderen
is wetenschap

Published data



Use of data



Flemish open data portal

Vlaanderen  Zoek 

Datasets  Activiteiten stroom  Over

Doorzoeken datasets... 

30 datasets gevonden Sorteer op Relevantie 

INBO
Datasets van het Instituut voor Natuur- en Bosonderzoek [lees meer](#)

Volgers 0 Datasets 30

Organisaties 

Groepen

 Vlaanderen is wetenschap

[**Bird tracking - GPS tracking of Lesser Black-backed Gulls and Herring Gulls b...**](#)

[**inbo**](#)

Bird tracking - GPS tracking of Lesser Black-backed Gulls and Herring Gulls breeding at the southern North Sea coast is a species occurrence dataset published by the Research...



[**Loopkevers Grensmaas - Ground beetles near the river Meuse in Flanders, Belgium**](#)

Openly document process on GitHub

inbo / data-publication

Code Issues 66 Pull requests 0 Projects 2 Pulse Graphs

Branch: master data-publication / datasets / bird-tracking-gull-occurrences / Create new file Find file History

README.md

Bird tracking - GPS tracking of Lesser Black-backed Gulls and Herring Gulls breeding at the southern North Sea coast

Bird tracking - GPS tracking of Lesser Black-backed Gulls and Herring Gulls breeding at the southern North Sea coast is a species occurrence dataset published by the Research Institute for Nature and Forest (INBO) at <http://dataset.inbo.be/bird-tracking-gull-occurrences>.

Data publication process

- Metadata (working document)
- Darwin Core mapping (R) from the processed log files
- Darwin Core mapping (SQL) from CartoDB (DEPRECATED)
- Known issues

 **Vlaanderen**
is wetenschap

Challenges

- ▶ International platforms/standards for abiotic data
- ▶ Publish data associated with scientific publications and reports
- ▶ Think about data publication at the start of the project
- ▶ Automate workflows



Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

Biometry, Methodology & Quality assurance

Methodological support

Sir Ronald Aylmer Fisher

To call in the statistician after the experiment is done may be no more than asking him to perform a post-mortem examination: he may be able to say what the experiment died of.



Partner during the entire lifecycle of research

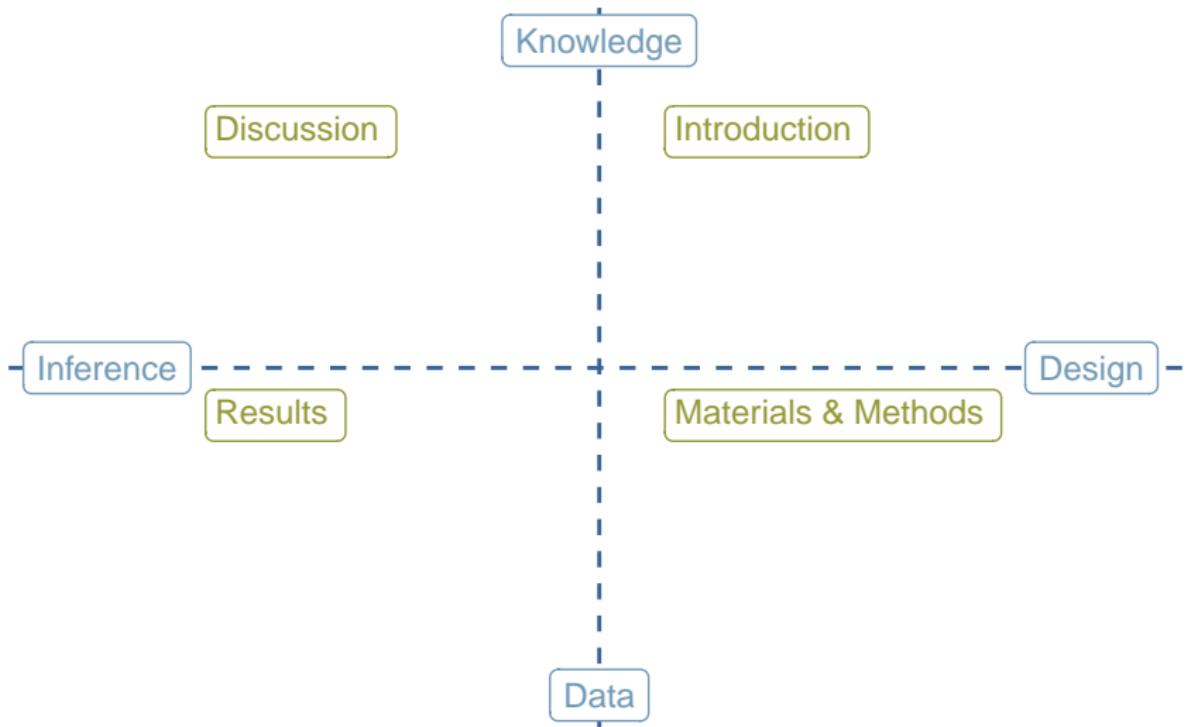
Discussion

Introduction

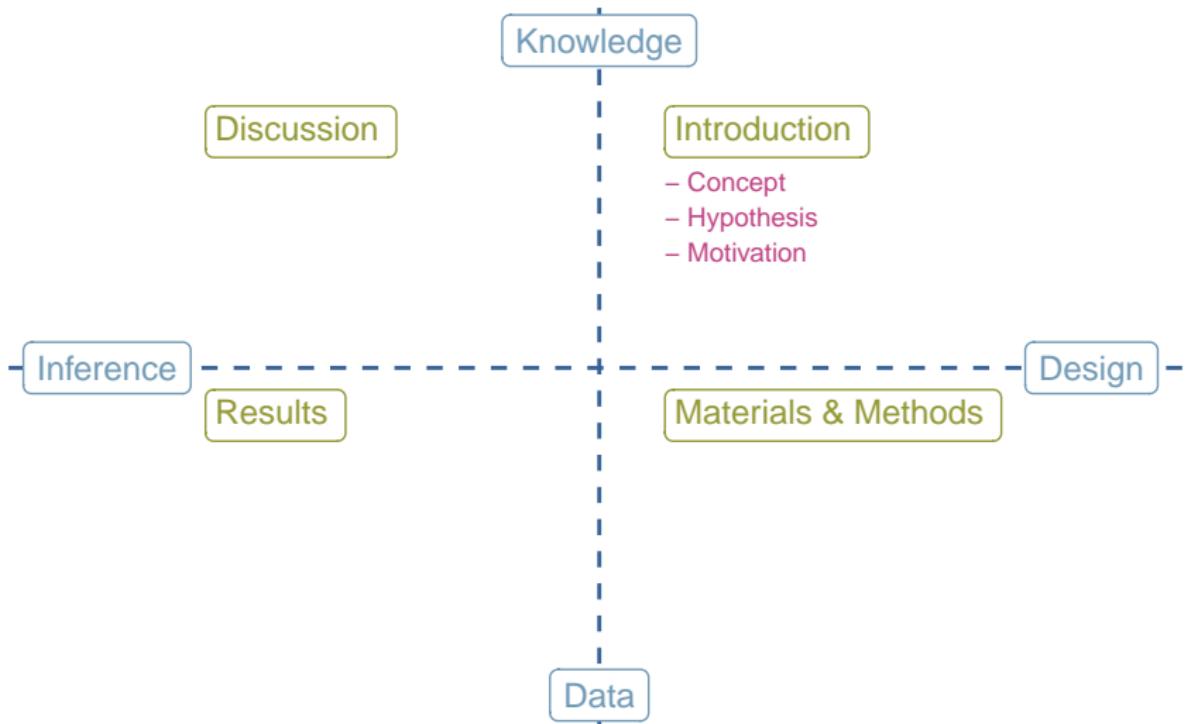
Results

Materials & Methods

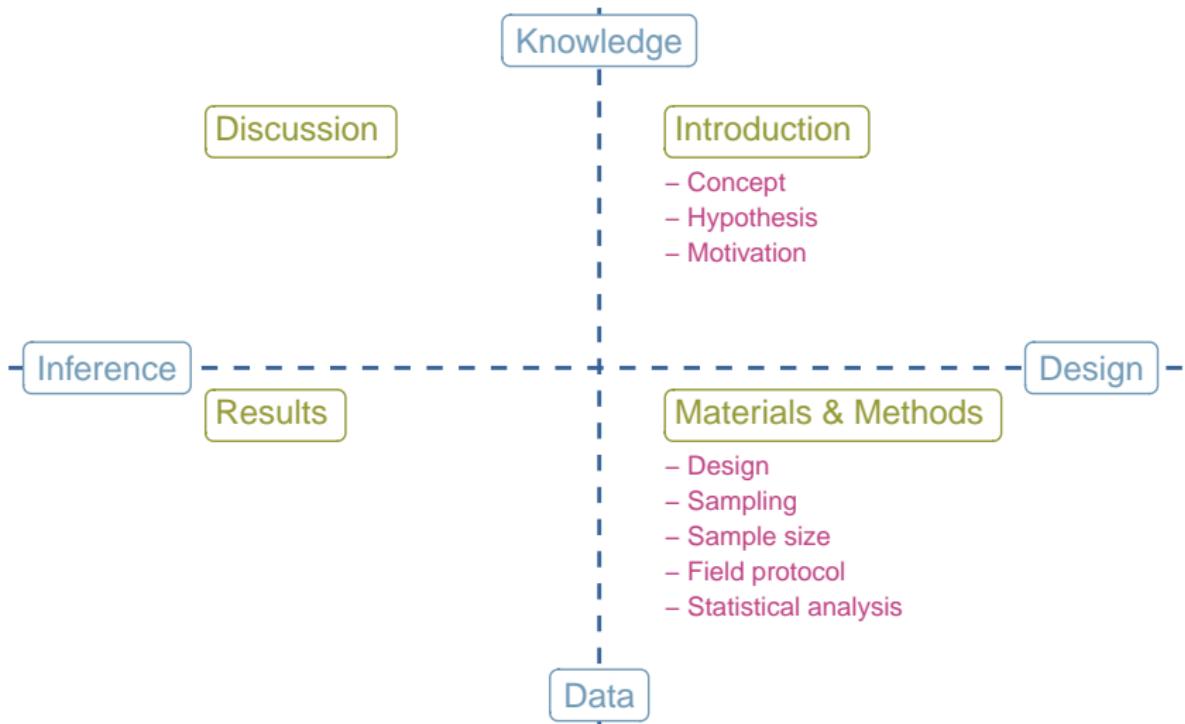
Partner during the entire lifecycle of research



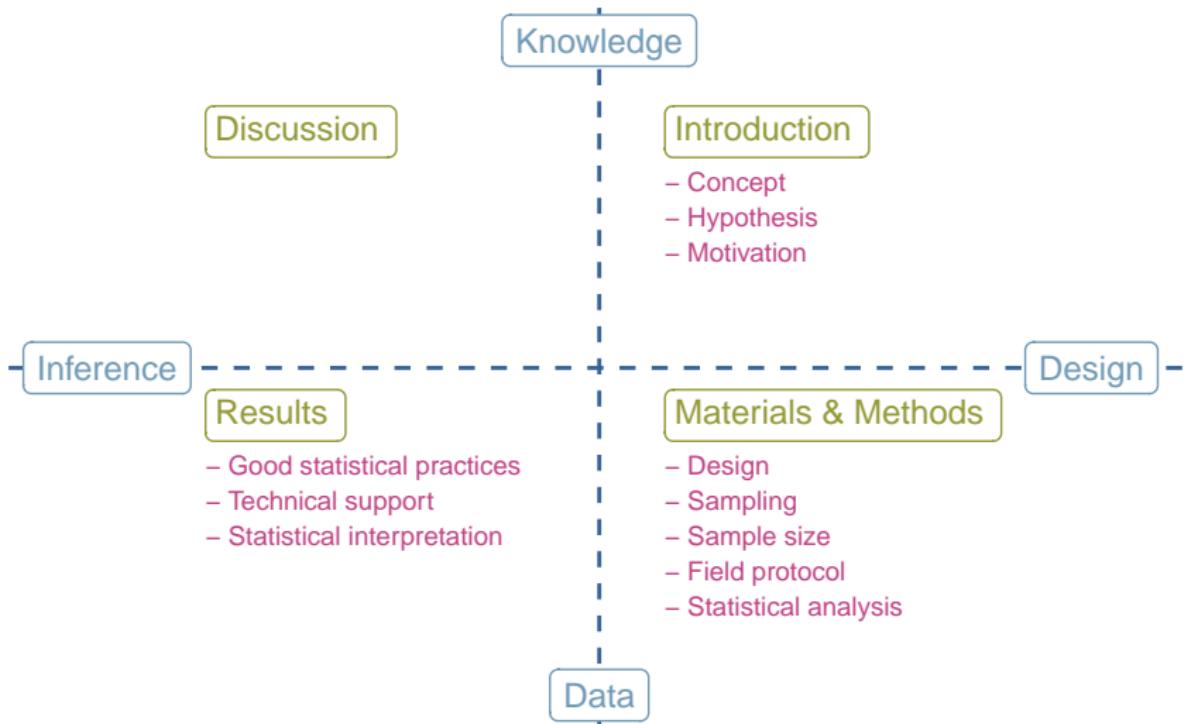
Partner during the entire lifecycle of research



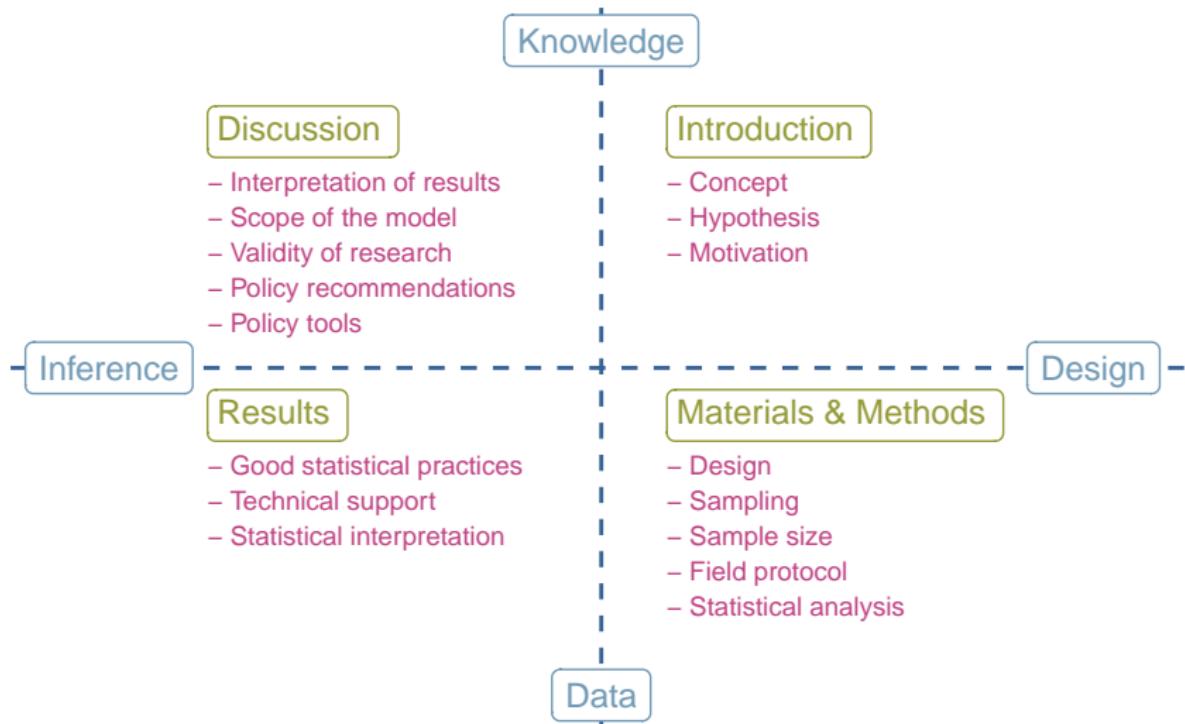
Partner during the entire lifecycle of research



Partner during the entire lifecycle of research



Partner during the entire lifecycle of research



Acquisition, development and maintainance of methodologies

- ▶ Concept for setting up long term ecological monitoring

Acquisition, development and maintenance of methodologies

- ▶ Concept for setting up long term ecological monitoring
- ▶ Field protocols

Acquisition, development and maintenance of methodologies

- ▶ Concept for setting up long term ecological monitoring
- ▶ Field protocols
- ▶ Novel statistical techniques

Acquisition, development and maintenance of methodologies

- ▶ Concept for setting up long term ecological monitoring
- ▶ Field protocols
- ▶ Novel statistical techniques
- ▶ Automation of recurring analyses

Acquisition, development and maintenance of methodologies

- ▶ Concept for setting up long term ecological monitoring
- ▶ Field protocols
- ▶ Novel statistical techniques
- ▶ Automation of recurring analyses
- ▶ Traceability of analyses

Statistical support

Karl Pearson

The unity of all science consists alone in its method, not in its material.

Provide tools to improve open and reproducible statistical analysis

- ▶ R

Provide tools to improve open and reproducible statistical analysis

- ▶ R
- ▶ Scripted analysis

Provide tools to improve open and reproducible statistical analysis

- ▶ R
- ▶ Scripted analysis
- ▶ Version control

Provide tools to improve open and reproducible statistical analysis

- ▶ R
- ▶ Scripted analysis
- ▶ Version control
- ▶ Dynamic documents

Provide tools to improve open and reproducible statistical analysis

- ▶ R
- ▶ Scripted analysis
- ▶ Version control
- ▶ Dynamic documents
- ▶ Corporate identity styles for graphics and dynamic documents

Freeware open source software

A word cloud centered around the letter R, representing freeware open source software. The words are arranged in a circular pattern around the R, with larger words indicating higher frequency or importance. The words include:

- Markdown
- R
- GitHub
- RStudio
- vegan
- testthat
- multcomp
- multimpute
- rgdal
- covr
- geos
- lme4
- ggmap
- readr
- nlme
- lubridate
- mgcv
- devtools
- tidyverse
- INLA
- INBOtheme
- readxl
- leaflet
- knitr
- DBI
- raster
- intr
- gstat
- gstat
- INBOMd
- goglesheets

Improve statistical skills of scientific staff



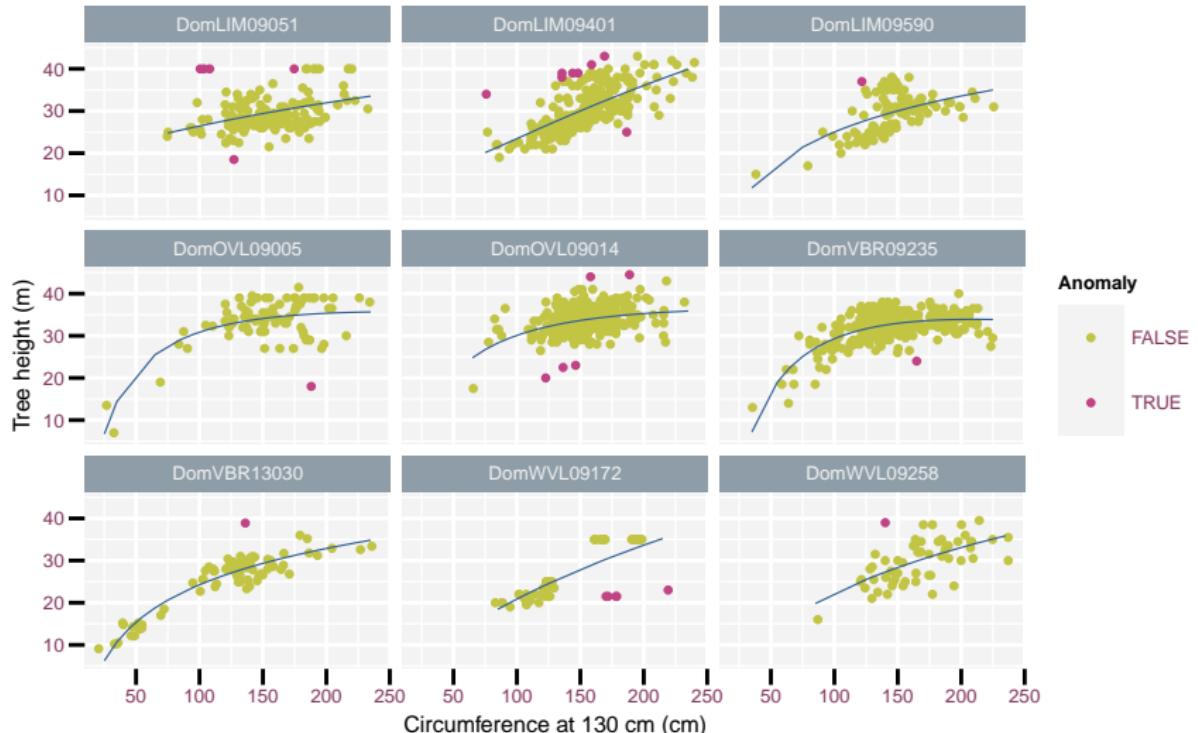
Vlaanderen
is wetenschap

Statistical expertise

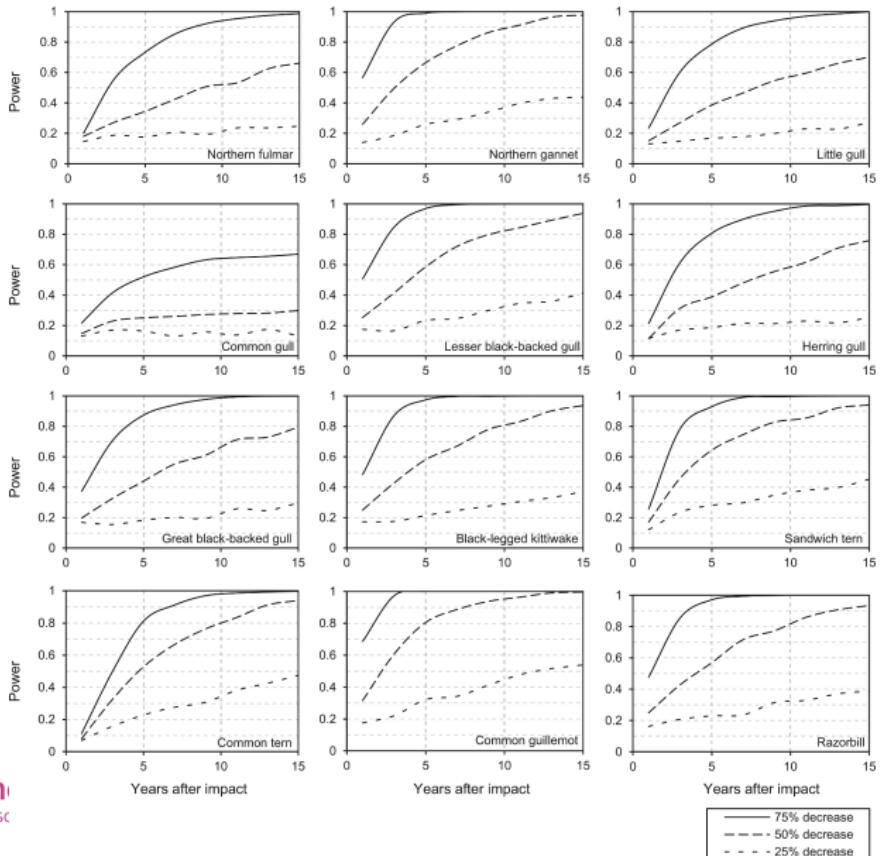
George E. P. Box

All models are wrong, but some are useful.

Data validation through statistical modelling



Power calculation





Vlaanderen
is wetenschap

INSTITUUT
NATUUR- EN BOSONDERZOEK

Towards supporting policy with open reproducible science

Rome wasn't built in a day

- ▶ Gradual change is key for success
- ▶ New procedures should lower the burden in the long run
- ▶ Stimulation and support is vital
- ▶ Focus on early adopters and early majority
- ▶ Allow for a long transition period before enforcing change

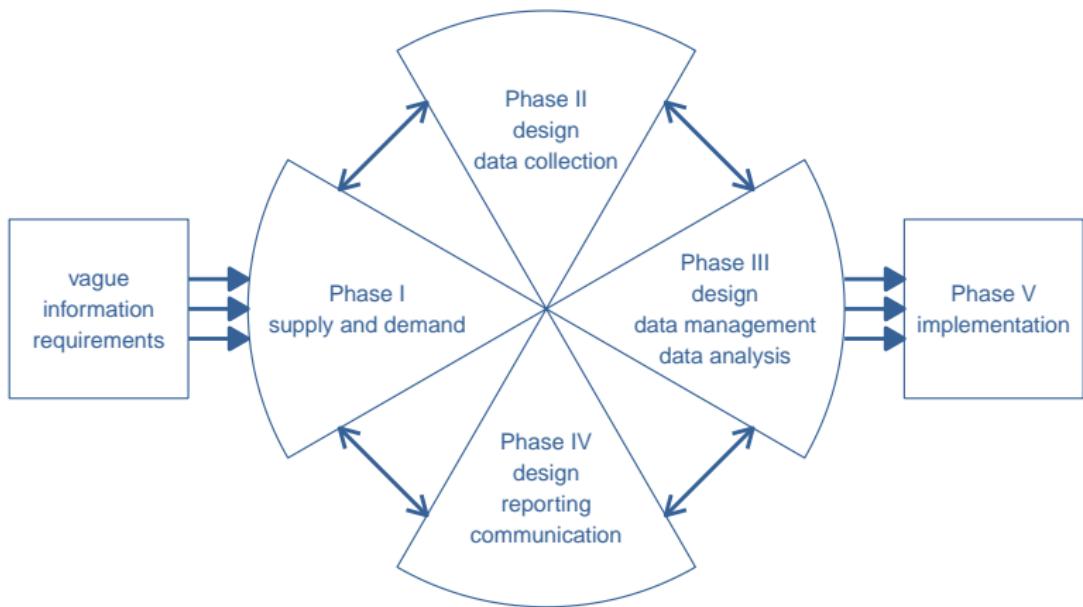
The ‘good’ old days



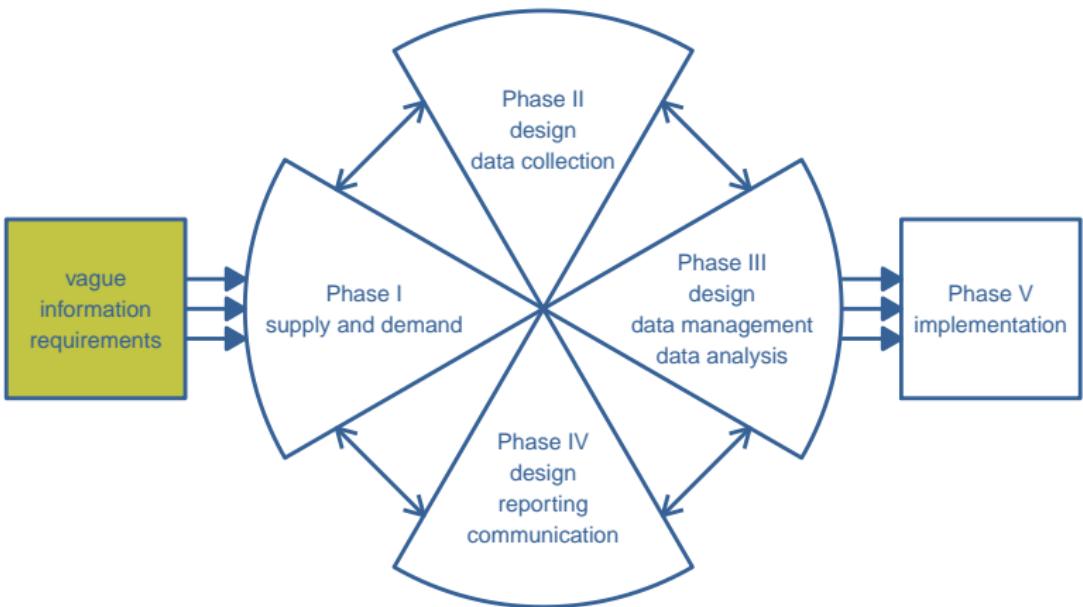
John Tukey

The combination of some data and an aching desire for an answer does not ensure that a reasonable answer can be extracted from a given body of data.

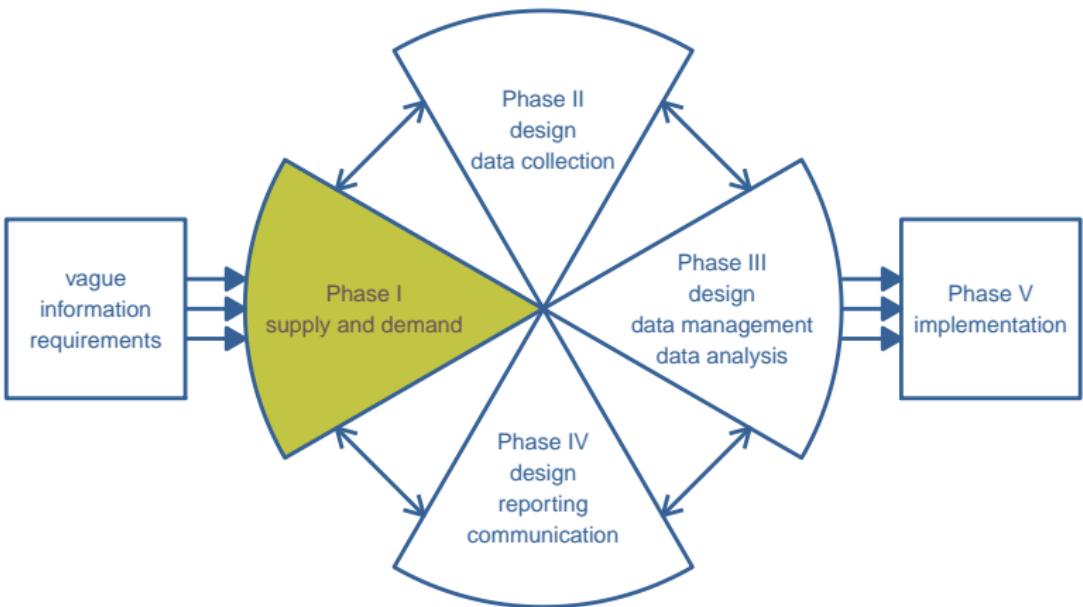
5 stages of preparing long term monitoring



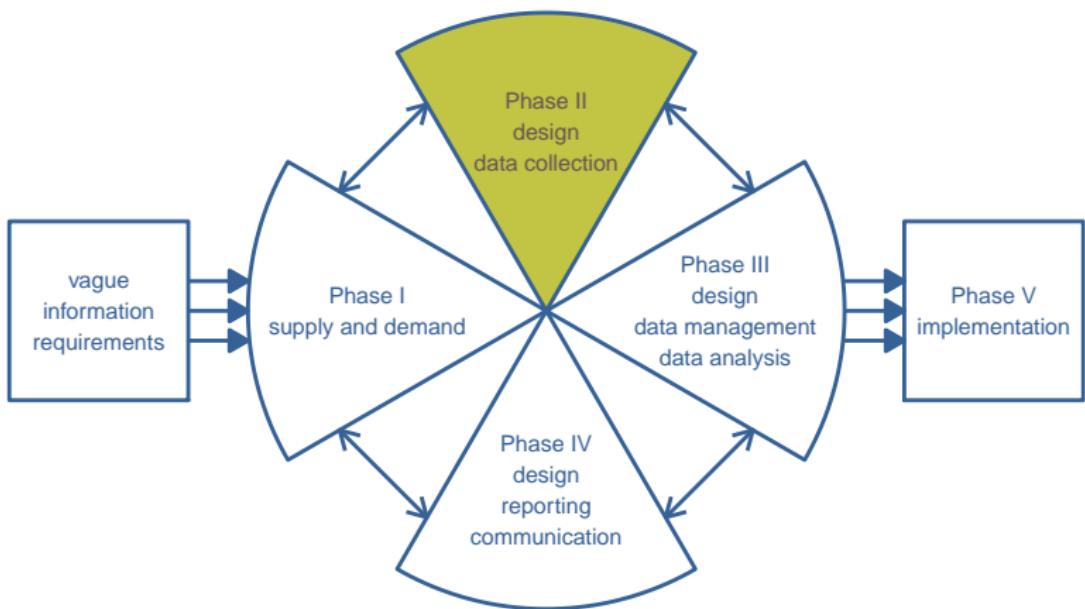
Focus on the relevant questions



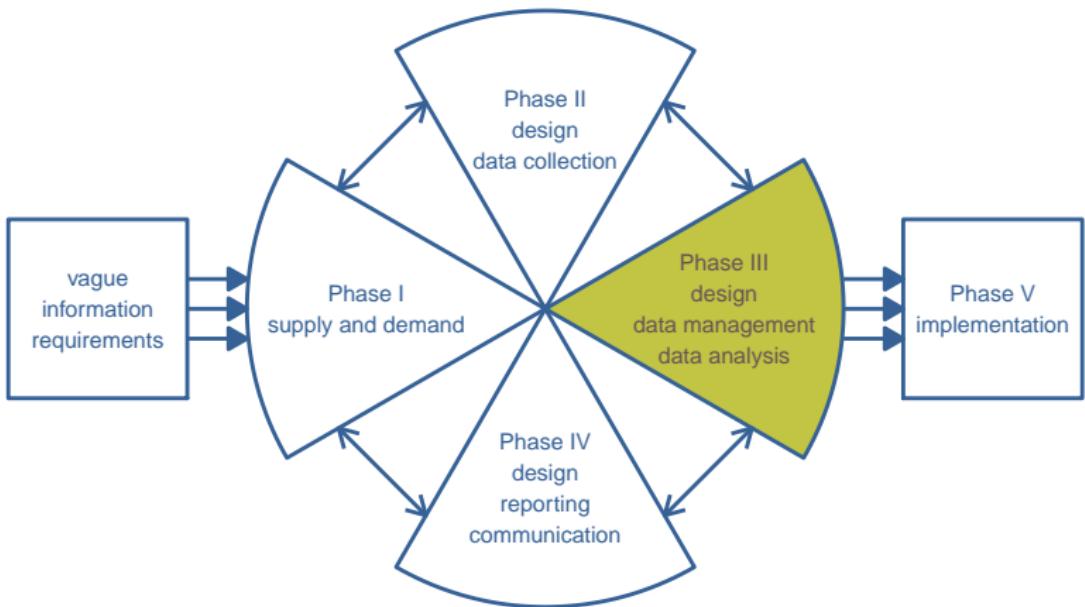
Prioritise the questions



Scenarios based on power and cost

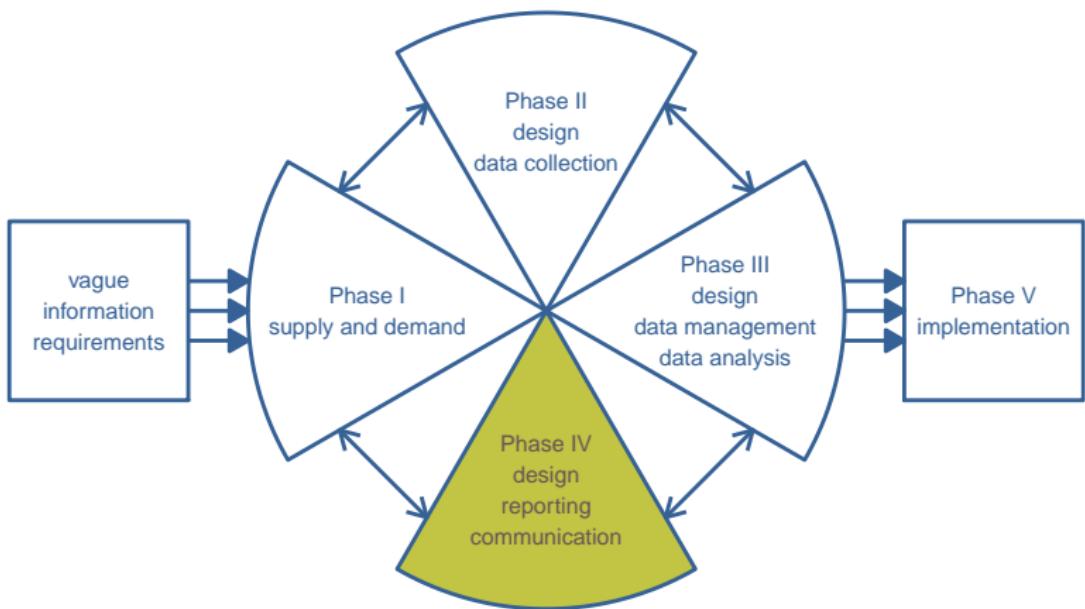


From raw data to final results

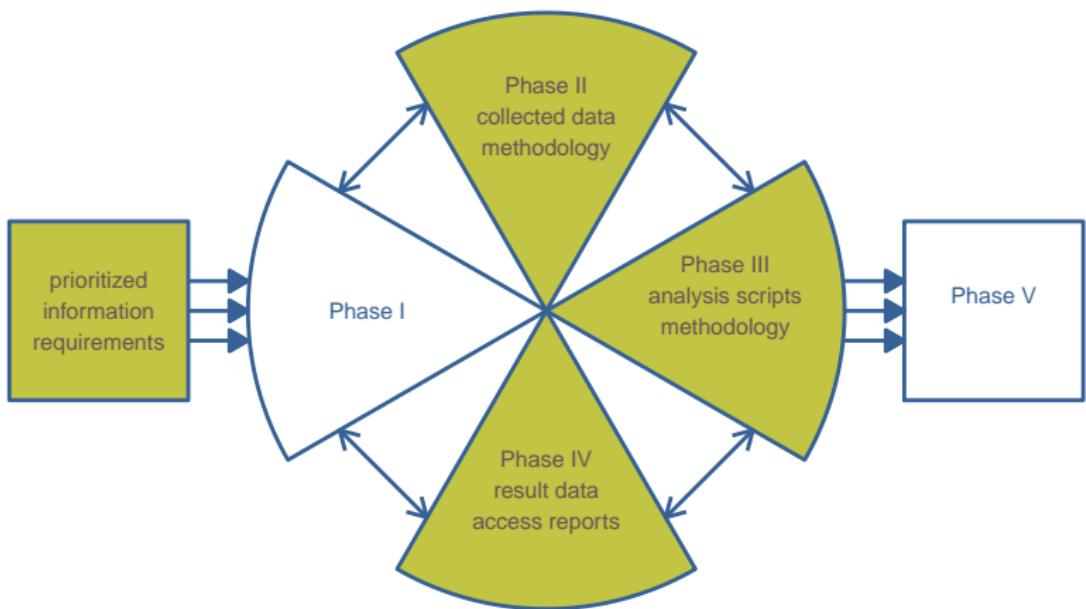


Vlaanderen
is wetenschap

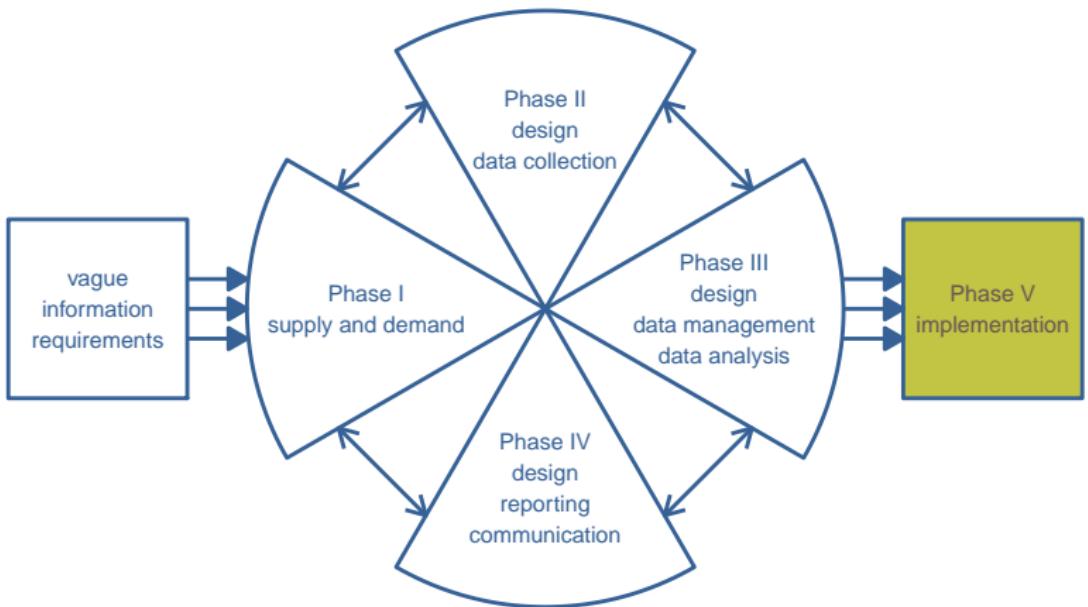
What to communicate



Which aspects can we make open?



We're almost ready to rock'n roll



Vlaanderen
is wetenschap