DataKube

empowering domain experts with scalable knowledge discovery

Kristoffer L Nielbo
`knielbo@sdu.dk`
`knielbo.github.io`

Dept. of History & SDU eScience Center
University of Southern Denmark

Outline

DataKube

Kristoffer L Nielbo
knielbo@sdu.dk
knielbo.github.io

A need for Human
Informatics

SDU eScience Center

SDU DataKube
  organization
  activities

1 A need for Human Informatics

2 SDU eScience Center

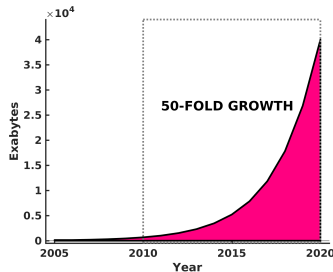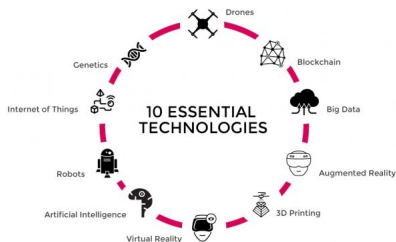3 SDU DataKube

SDU ⬩

A need for Human Informatics

DataKube

Kristoffer L Nielbo
knielbo@sdu.dk
knielbo.github.io

A need for Human
Informatics

SDU eScience Center

SDU DataKube

organization
activities

the data deluge is transforming knowledge discovery and understanding in every domain of human inquiry

a large part of these data are soft and unstructured $\Rightarrow$ to get value from these data, humanities (and social sciences) must utilize automation

human informatics - automatic information processing in the humanities

K. Steinbuch (1957) Informatik: Automatische Informationsverarbeitung [Berlin: SEG-Nachrichten]

SDU

eScience Infrastructure at SDU

DataKube

Kristoffer L Nielbo
knielbo@sdu.dk
knielbo.github.io

A need for Human
Informatics

SDU eScience Center

SDU DataKube
 organization
 activities

14.016 cores, 392 slim nodes, 64 fat nodes, 72 GPU nodes with two Nvidia K40/node

the center is expanding its userbase to include more SSH relevant areas and competencies, e.g., DL, NLP, computer vision

cloud offers a fully GDPR complient environment for cloud computing with a cumulative codebase
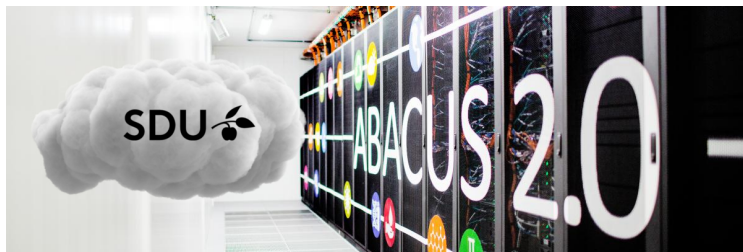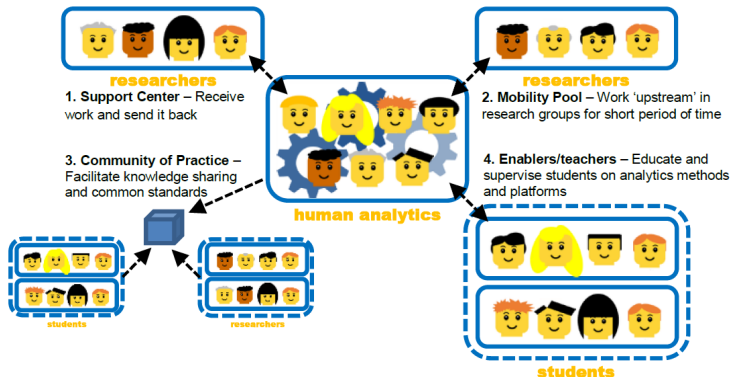
SDU

SDU DataKube

DataKube

Kristoffer L Nielbo
knielbo@sdu.dk
knielbo.github.io

A need for Human
Informatics

SDU eScience Center

SDU DataKube

organization

activities

researchers

1. Support Center – Receive work and send it back

2. Mobility Pool – Work 'upstream' in research groups for short period of time

3. Community of Practice – Facilitate knowledge sharing and common standards

4. Enablers/teachers – Educate and supervise students on analytics methods and platforms

human analytics

researchers

students

students

- Research support through collaboration (pilot, shared funding, shared authorship)

- Recently turned to tool and competency development - software development in the humanities

SDU

stdoutreach & NDHL

DataKube

Kristoffer L Nielbo
knielbo@sdu.dk
knielbo.github.io

A need for Human
Informatics

SDU eScience Center

SDU DataKube
organization
activities

empowering domain experts with scalable knowledge discovery

STDOUTREACH

- standardized resource for competency development text analytics

- `datakuben YouTube`

- InfoVis/visual analytics for Scandinavian Textual Cultural Heritage

NDHL

- modelled over the `NLPL` template, we are proposing a
  Nordic Digital Humanities Laboratory

- "computing allocations of at least one million node*hours
  pr. year for a project directory (for domain-specific
  software and data) of at least a handful of TB, and
  mechanisms to allow NDHL members from Denmark,
  Finland, Norway, and Sweden to obtain user accounts and
  make use of these allocations"

SDU❧