

Sustainability & Data Literacy

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

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

November 15, 2018

Outline

Sustainability & Data
Literacy

Kristoffer L Nielbo
knielbo@sdu.dk
[knielbo.github.io](https://github.com/knielbo)

① Sustainability

Carpentries

CodeRefinery

Humanities Programming

Sustainability

Carpentries

CodeRefinery

Humanities

Programming

② Data Literacy

a need for human informatics
mininmal requirements

Data Literacy

a need for human
informatics
mininmal
requirements

Visualization

value of visualization
limitation

Challenges to data
literacy

③ Visualization

value of visualization
limitation

④ Challenges to data literacy



Sustainability & Data Literacy

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

Sustainability

Carpentries
CodeRefinery
Humanities
Programming

Data Literacy

a need for human informatics
minimal requirements

Visualization

value of visualization
limitation

Challenges to data literacy

sustainability of competency development ~ organizational resilience

- **diversity**, we need to adapt to multiple research environments
- **openness**, we must be inclusive towards the humanities
- **modularity**, we should support a high level of independence
- **slack resources**, we have to allow for “wasteful” development
- **matching cycles**, we should not be focused on rapid growth

Sustainability

Carpentries

CodeRefinery

Humanities

Programming

Data Literacy

a need for human
informatics
minimal
requirements

Visualization

value of visualization
limitation

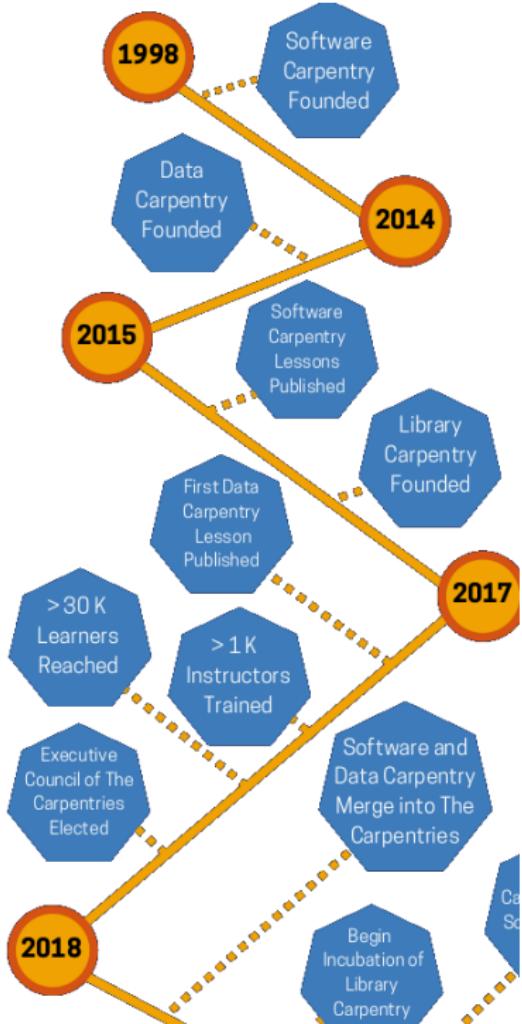
Challenges to data
literacy

goals

- “builds global capacity in essential data and computational skills for conducting efficient, open, and reproducible research.”
- teaching basic *scientific computing*

model

- strong brand, crowd-sourced development
- peer-to-peer, collaborative learning and knowledge transfer, train-the-trainer program
- external funding and membership tiers

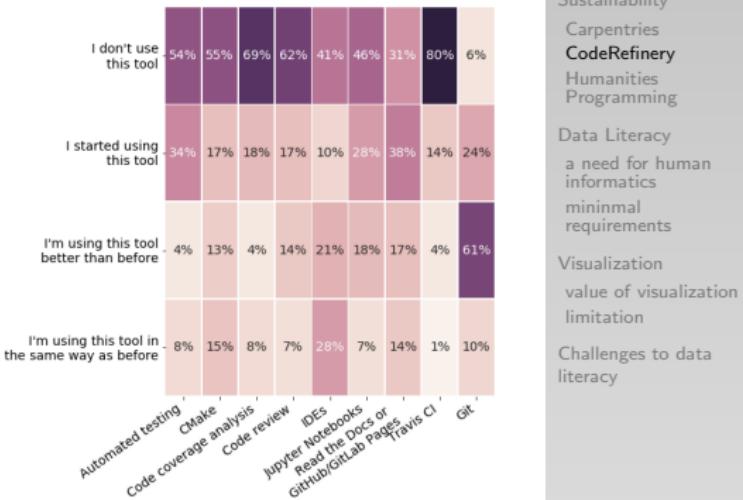


goals

- increased competence in software development
- better science through reproducibility and reusable code
- strengthening Nordic communities

model

- peer-to-peer, collaborative learning and knowledge transfer, train-the-trainer program
- phase 1 and phase 2 have been supported by NeIC
- future is either partnership model or embedded in the Carpentries



Sustainability

Carpentries

CodeRefinery

Humanities

Programming

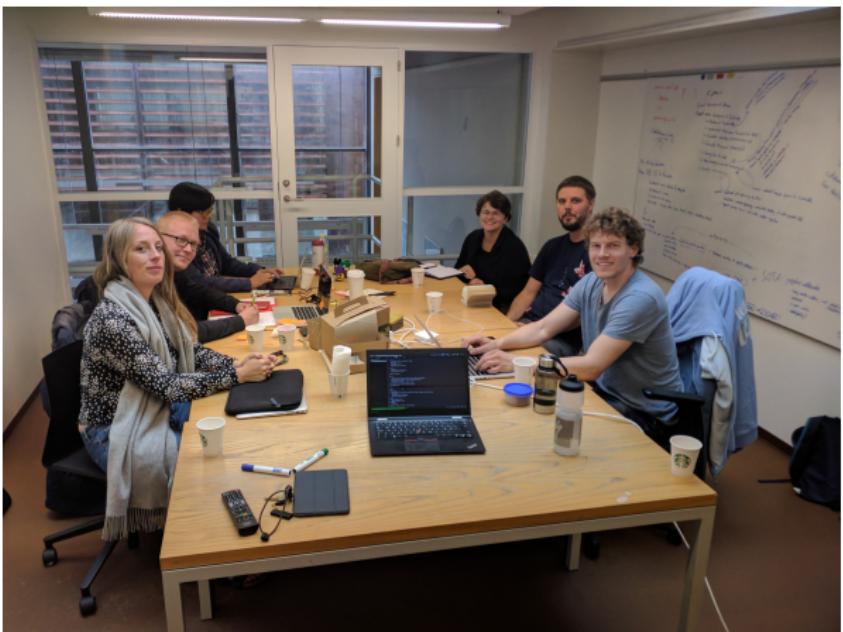
Data Literacy

a need for human
informatics
minimal
requirements

Visualization

value of visualization
limitation

Challenges to data
literacy



Sustainability
Carpentries
CodeRefinery
Humanities
Programming

Data Literacy
a need for human
informatics
minimal
requirements

Visualization
value of visualization
limitation

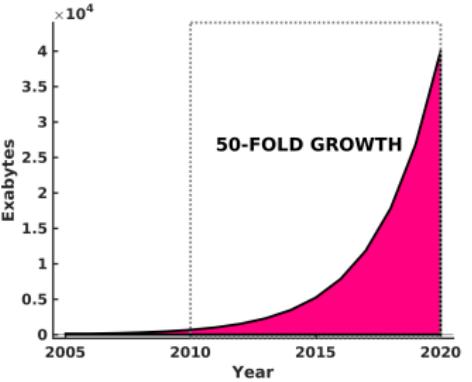
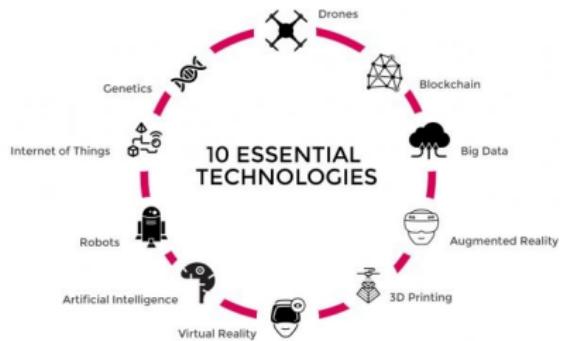
Challenges to data
literacy

- dissatisfaction with “example mapping” exercises, we started designing state-of-the-art programming for humanities
- redefining decomposition, pattern recognition, and algorithmic design
- debate physical and visual programming

A need for human informatics

Sustainability & Data Literacy

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



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 ⇒ to get value from these data, humanities (and social sciences) must utilize automation

human informatics - automatic information processing in the humanities

Sustainability
Carpentries
CodeRefinery
Humanities
Programming

Data Literacy
a need for human informatics
minimal requirements

Visualization
value of visualization
limitation

Challenges to data literacy

minimal requirements

Sustainability & Data Literacy

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

digital literacy

- knowledge of research and methods in the humanities that rely heavily on digital technology

Sustainability

Carpentries

CodeRefinery

Humanities

Programming

Data Literacy

a need for human informatics

minimal requirements

Visualization

value of visualization
limitation

Challenges to data literacy

computational literacy

- basic practice understanding and implementation of algorithmic problem solving (, & ...)

data literacy

- basic practice understanding and implementation of statistical analysis and visualization (, & ...)

	Set A		Set B		Set C		Set D	
	X	Y	X	Y	X	Y	X	Y
0	10	8.04	10	9.14	10	7.46	8	6.58
1	8	6.95	8	8.14	8	6.77	8	5.76
2	13	7.58	13	8.74	13	12.74	8	7.71
3	9	8.81	9	8.77	9	7.11	8	8.84
4	11	8.33	11	9.26	11	7.81	8	8.47
5	14	9.96	14	8.10	14	8.84	8	7.04
6	6	7.24	6	6.13	6	6.08	8	5.25
7	4	4.26	4	3.10	4	5.39	19	12.50
8	12	10.84	12	9.13	12	8.15	8	5.56
9	7	4.82	7	7.26	7	6.42	8	7.91
10	5	5.68	5	4.74	5	5.73	8	6.89
mean	9.00	7.50	9.00	7.50	9.00	7.50	9.00	7.50
std	3.32	2.03	3.32	2.03	3.32	2.03	3.32	2.03
corr	0.82		0.82		0.82		0.82	
lin. reg.	$y = 3.00 + 0.500x$		$y = 3.00 + 0.500x$		$y = 3.00 + 0.500x$		$y = 3.00 + 0.500x$	

“visualizations are great, it means I have to read less”

Sustainability

Carpentries

CodeRefinery

Humanities

Programming

Data Literacy

a need for human
informatics
minimal
requirements

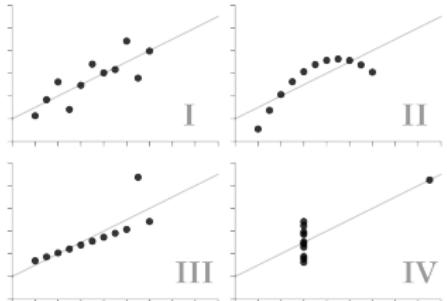
Visualization

value of visualization
limitation

Challenges to data
literacy

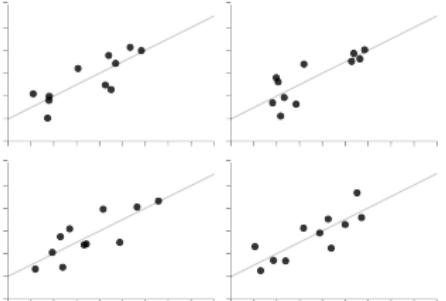
✓ Anscombe's Quartet

Each dataset has the same summary statistics (mean, standard deviation, correlation), and the datasets are *clearly different*, and visually *distinct*.



✗ Unstructured Quartet

Each dataset here also has the same summary statistics. However, they are not *clearly different* or *visually distinct*.



challenges to data literacy

Sustainability & Data Literacy

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

example

- DL2018 at Aarhus University combines group training with peer-to-peer and centralized support

challenges

- design principles and standards (e.g., Carpentries)
- implementation
- retain the “soft human core”
- cost of facilitating creativity

Sustainability

Carpentries

CodeRefinery

Humanities

Programming

Data Literacy

a need for human informatics
minimal requirements

Visualization

value of visualization
limitation

Challenges to data literacy

THANK YOU

knielbo@sdu.dk

knielbo.github.io

slides: http://knielbo.github.io/files/kln_neic_dhl.pdf

& credits to

Max R. Echardt and Katrine F. Baunvig, datakube, University of Southern Denmark, DK
Mads Rosendahl Thomsen, Comparative Literature, School of Communication and
Culture, Aarhus University, DK