

DH Benelux 2022

Developing Data Stories as Enhanced Publications in Digital Humanities

1 June 2022

Willemien Sanders, Roeland Ordelman, Mari Wigham, Rana Klein,
Jasmijn Van Gorp, Julia Noordegraaf

CLARIAH Media Suite

- . multimodal collections
- . 91 datasets from 62 organizations
- . bookmarking, annotating and analysing
- . Media Suite Data Stories: publications of quantitative research
- . underlying data accessible for scrutiny



Context

- . 'storyfinding' (D'Ignazio and Klein 2020)
 - . digital hermeneutics (Fickers and Clavert 2021)
 - . protocols for ... self-reflexive position (Fickers and Clavert 2021)
 - . need for contextualization data (D'Ignazio and Klein 2020)
 - . interplay technology and research (Condie and Costa 2018)
- > demands and practice of developing academic data stories

Our paper:

- . data stories
- . data stories in journalism
- . data stories in the humanities
- . Media Suite Data Stories – protocol
- . conclusion and future work

Data stories

- . Journalism: production / academia: potential (Riche et al. 2018)
- . variety of narratives involving quantitative data and, usually, visualizations (Weber 2020)
- . common practice: data collection, automated analysis, visualization, narration (Gray and Bounegru 2021, Weber, Engebretsen, and Kennedy 2018)
- . also exploration needed to understand context (D'Ignazio and Klein 2020)

Data stories

- . story / narrative: events, time, causality (Weber, Engebretsen, and Kennedy 2018)
- . visualizations: make data/analysis understandable
(Duangphummet, Apiwan, and Puripant Ruchikachorn 2021; Riche et al. 2018; Weber, Engebretsen, and Kennedy 2018)
- . output: multimodal and hybrid ; multimodal interactive visualizations (Weber, Engebretsen, and Kennedy 2018)

Data stories

Our definition:

Data stories are the output of an iterative process of data collection, data exploration, data preparation, data analysis, data visualization, interpretation and narration. The result is a narrative that is based on the analysis of quantitative data and includes visualisations of these data with explanations of their meaning and their context.

Data stories in journalism

- . use of data started in economic journalism in C17 (Arrese 2022)
- . use of public (gov) data -> lack of context (D'Ignazio and Klein 2020)
- . transparency with respect to journalistic data visualisations needed (Weber, Engebretsen, and Kennedy 2018)

Data stories in journalism

Taxonomies

- author-driven versus reader-driven

- magazine style, annotated

- chart, partitioned poster, flow

- chart, comic strip, slide show,

- and film/video/animation (Segel

- and Heer 2010)

user, and the graphical elements play an important role in making this possible. ... of the ... merely an ...

infogram Examples Features Blog Customers Pricing Use Case Sign in Get started

Introduction Nation-building Security Counter-narcotics

2007

Drug dealer to see change over time

Click on each country to see nation-building activities and related PDF locations.

Opium cultivation (hectares)

- 25,000+
- 15,000 - 24,999
- 10,000 - 14,999
- 5,000 - 9,999
- 1,000 - 4,999
- 500 - 999
- 100 - 499
- 0 - 99

US (troops: 45,780)

In the fiscal year 2009, the US Department of Defense spent \$1.25B on counter-narcotics operations in Afghanistan (with \$161M requested for 2010), while the Department of Justice's Drug Enforcement Administration spent \$1.9B.

In the same year, the US Department of State's International Narcotics Control and Law Enforcement Fund contributed \$257M to promoting alternative livelihoods for opium producers (with \$275M requested for 2010), and \$200M to the eradication of plantations (with \$100M requested for 2010).

Source: US Office on Drugs and Crime

FINANCIAL TIMES

infogram Blog

Email Address*

Subscribe

Minnesota Quarterly Employment (by Industry) 2000-2007

All Minnesota Industries

93,792 (+3.7%) jobs added between 2000-2007



Share Your Insights (1362 Comments)

These seem to be three of the areas hardest hit by the slowdown.

by Andrew H. on Jul 16

Health care has accounted for more than half of the jobs added since 2000.

by Bryan C. on Jul 18 - 1 reply

What happened to 'information' jobs?? 25K+ lost in one quarter? That doesn't seem possible. Or did somebody

move out of state that I've overlooked?

by brian s. on Jul 18 - 1 reply

Minn Dept of Econ Dev attributed 2000-2001

unemployment to the dot com bust

(<http://www.deed.state.mn.us/mjpublications/trends/0701/layoff.htm>). It's interesting that the initial drop in

Arts employment preceded the start in Managerial

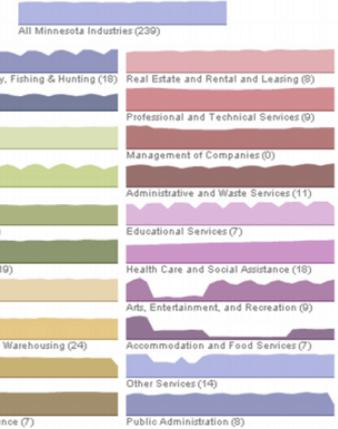
employment by 1 qtr. And Arts also bottomed out 3

qtrs before Management. Is Arts, Entertainment, and

Recreation employment an indicator of future trends in

White-Collar employment?

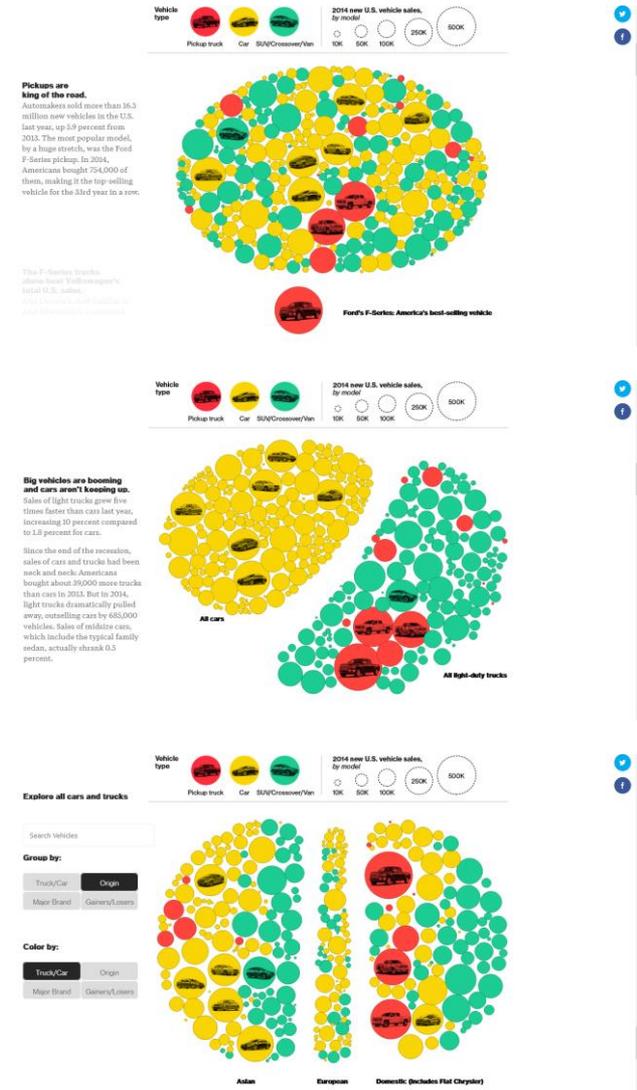
by Brad S. on Jul 19



Data stories in journalism

Taxonomies

- . communicating narrative and explaining data; linking separated story elements; enhancing structure and navigation; and providing controlled exploration (Stolper et al. 2016)



Data stories in journalism

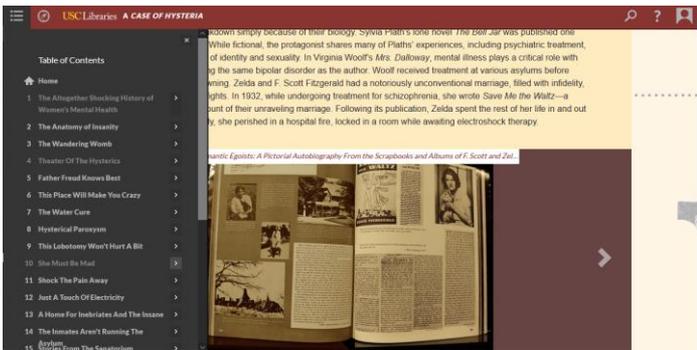
- . visualization -> narrator, sequentiality, time and “tellability” -> narrative (Weber 2020)

Journalism: storytelling *through* visualizations

Humanities: storytelling *with* visualization

Data stories in humanities

. initiatives:



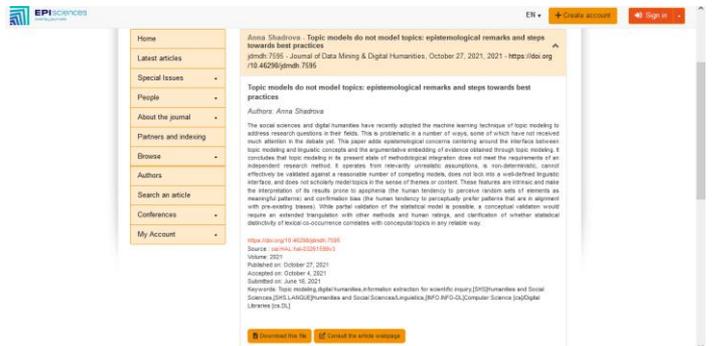
<- Scalar
J. of Digital
History ->



J. of Open
Humanities
<- Data



J. of Data
Mining and
Digital ->
Humanities



Data stories in Humanities

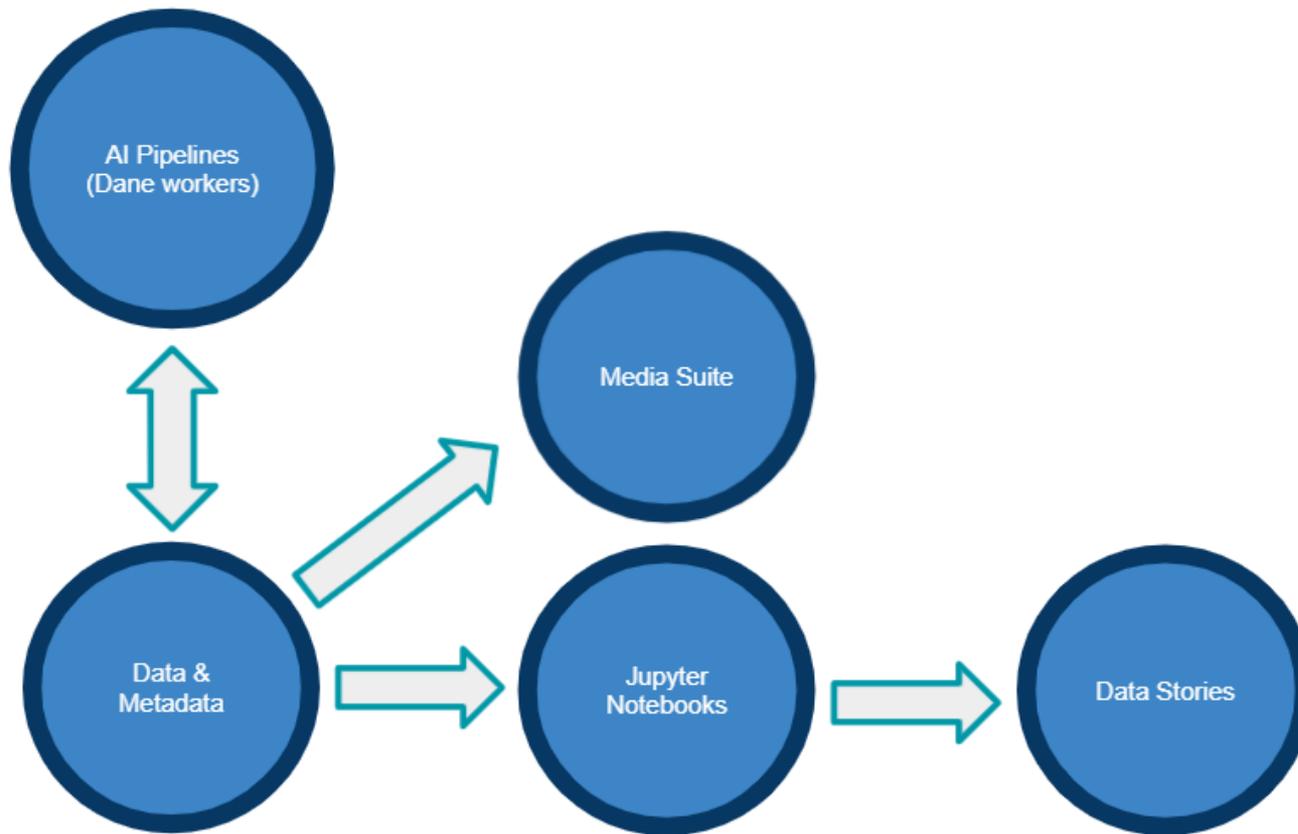
Journalism: uncovering hidden truths (Parasie 2015)

Humanities: situated knowledge (Haraway 1998)

-> how to best account for methodology and research practice?

- . guidelines (Figueiras and Vizoso 2022)
- . protocol based on experience (Duangphummet and Ruchikachorn 2021)

Media Suite Data Stories



Media Suite Data Stories



- . “author defined” (Stolper et al. 2016)
- . Domain specific knowledge + critical approach data & tool
- . Transparency with respect to analysis
- . Accessibility of underlying data
- . Scrollytelling (Weber 2020)

Media Suite Data Stories



- . Coalitiemannen spreken over (internationale) veiligheid
- . 'Ladies and gentlemen...'. Female and male presence in Dutch popular media
- . Factual divergences in Dutch television news: comparisons between AIDS/HIV, SARS and COVID-19 news reportings
- . 15 years of the popular Dutch chat show 'DWDD' - in data

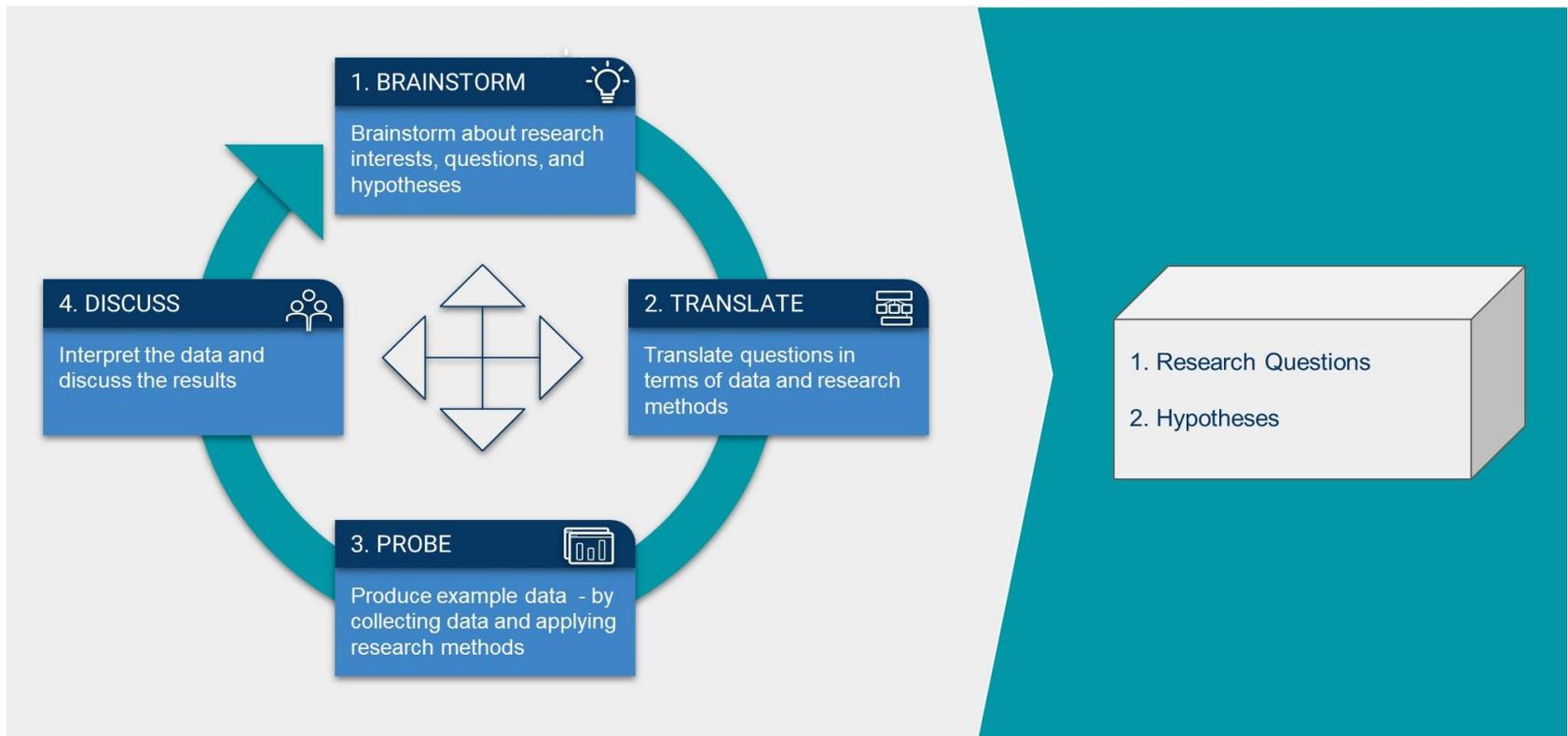
Media Suite Data Stories



- . Context: data registry
 - > metadata NPO staff
- . Editorial protocol: four iterative stages:
 - exploration
 - research
 - review
 - publication

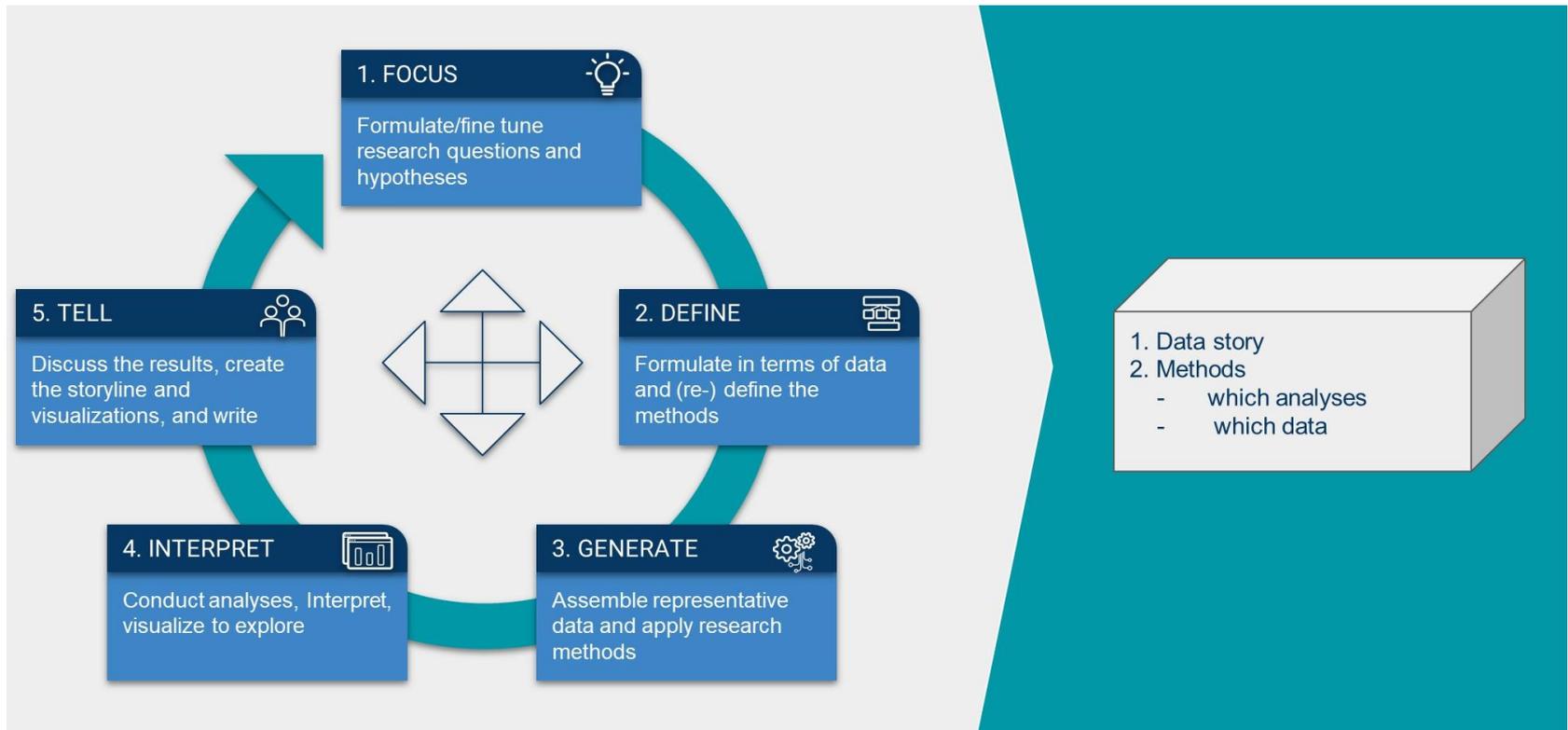
Media Suite Data Stories

. exploration



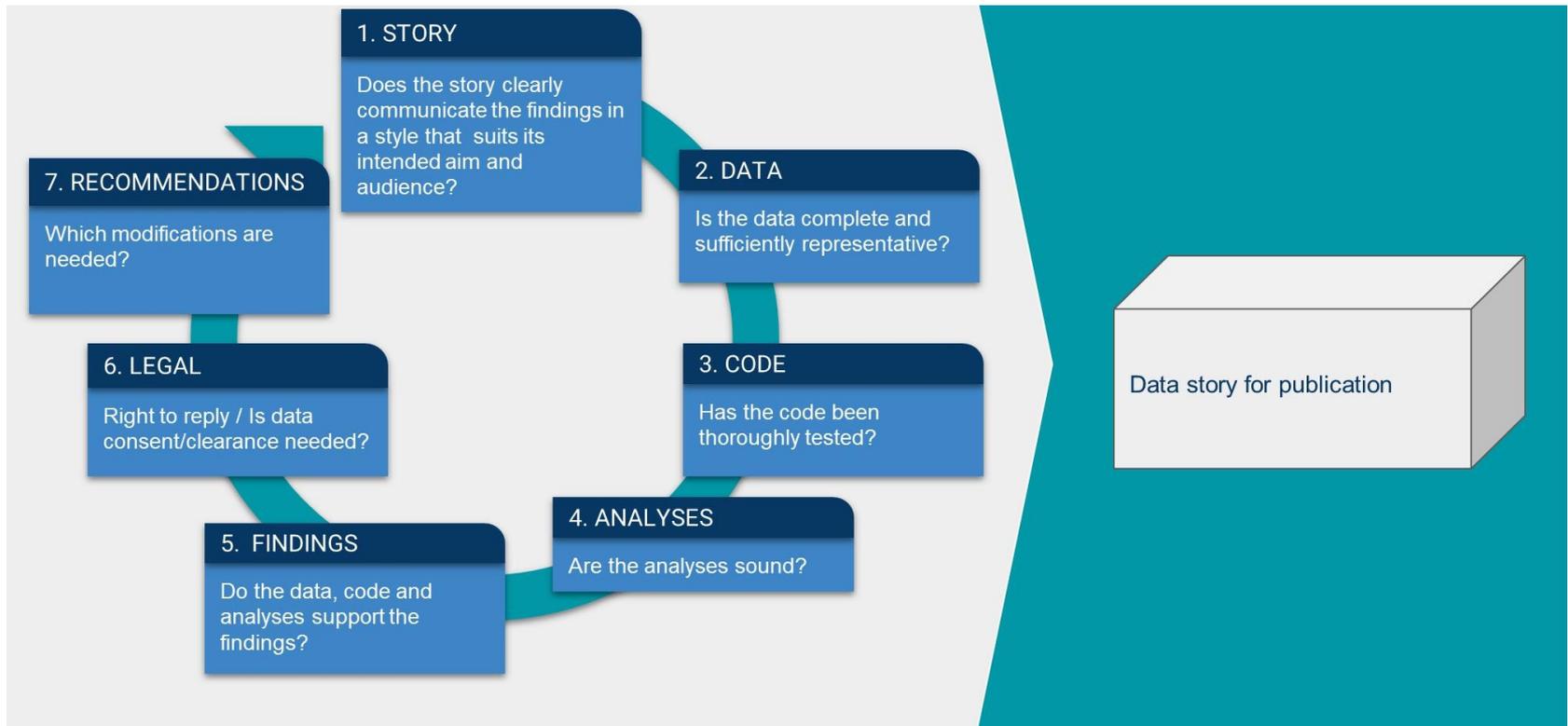
Media Suite Data Stories

. research



Media Suite Data Stories

. review



Media Suite Data Stories

. publication



Media Suite Data Stories

Conclusion and future work

- . Protocol as framework for Media Suite Data Stories ->
update if needed
- . Make developing Media Suite Data Stories more appealing ->
lowering threshold
- . Developing an academic Media Suite Data Story in
collaboration with a journal

Media Suite Data Stories

Conclusion and future work

- . Lowering threshold
 - . Modules / building blocks
 - . E.g., query for all 8PM newscasts
 - . E.g., searching for words in a dataset
 - . Making Jupyter Notebooks available for researchers
 - . Focusing on teams of researchers and data scientists

DH Benelux 2022

Developing Data Stories as Enhanced Publications in Digital Humanities

Thank you!

Willemien Sanders, Roeland Ordelman, Mari Wigham, Rana Klein,
Jasmijn Van Gorp, Julia Noordegraaf