DSI Approach: How to Define a Good Research Question for Digital Unstructured Data
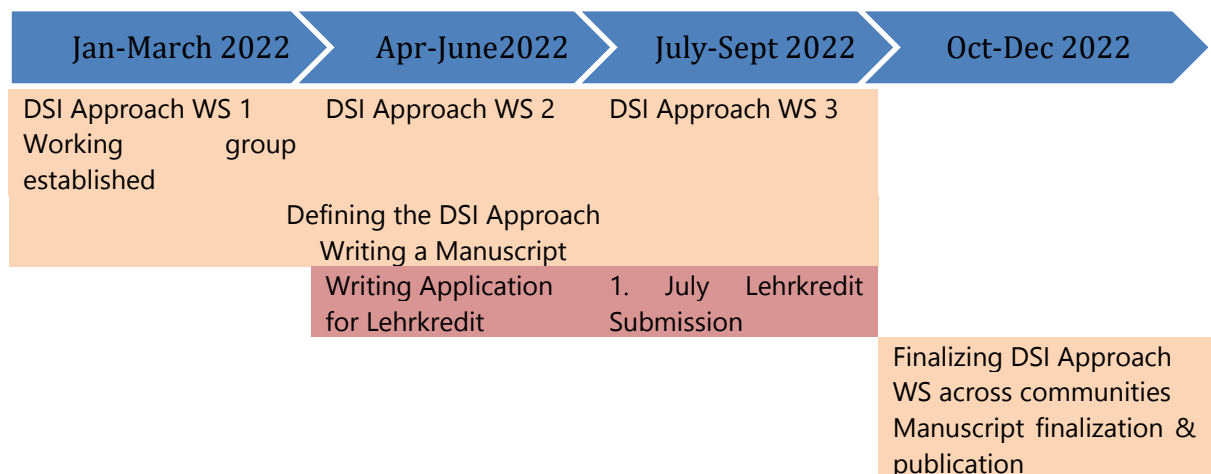
Goals: to define a "DSI Approach: How to Define a Good Research Question for Digital   Unstructured Data"
to create teaching tools
to develop a checklist/framework that will guide applied researchers in finding a feasible, relevant study question.

Milestones: 3 Workshops, Publication, Submission of a Proposal for Lehrkredit

Project Lead: Viktor von Wyl

Involved members: Viktor von Wyl, Jana Sedlakova, Dominik Ettlin, Thomas Friemel (Communication & Health Community), Andrea Horn, Anke Maatz, Gerold Schneider, Markus Wolf, Felix Gille (Ethics Community), Christina Haag, Chloé Sieber, Aleksandra Zumbrunn, Gero Daniel

Timeline (overview):

Jan-March 2022	Apr-June2022	July-Sept 2022	Oct-Dec 2022
DSI Approach WS 1 Working group established	DSI Approach WS 2	DSI Approach WS 3	
	Defining the DSI Approach Writing a Manuscript		
	Writing Application for Lehrkredit	1. July Lehrkredit Submission	
			Finalizing DSI Approach WS across communities Manuscript finalization & publication

Motivation

Both the literature review and the survey conducted in the DSI-Health community in 2021 have shown that one of the most crucial but at the same time very underappreciated steps in research is to define a good hypothesis and research question. This task becomes even more challenging when unstructured data are intended to be used. Unstructured data are important sources of health information that are generated by digital technologies. These digital technologies open up many opportunities which might be missed when there is not enough guidance and understanding of how to use unstructured data in an effective, reliable and ethical way. In other words, without frameworks of interpretation of the data output, no useful knowledge can be derived.

Furthermore, there appears to exist very limited guidance (for early-career scientists and in general) on how to approach the formulation of a good research question and hypothesis in this area. Current concepts and theories that can inform the formulation of new hypotheses and questions need to be adjusted to the new universe of data at hand. These aspects are particularly important since unlike other issues such as standardization problems, the problem of alignment with a good hypothesis and research question can be solved in the research community.
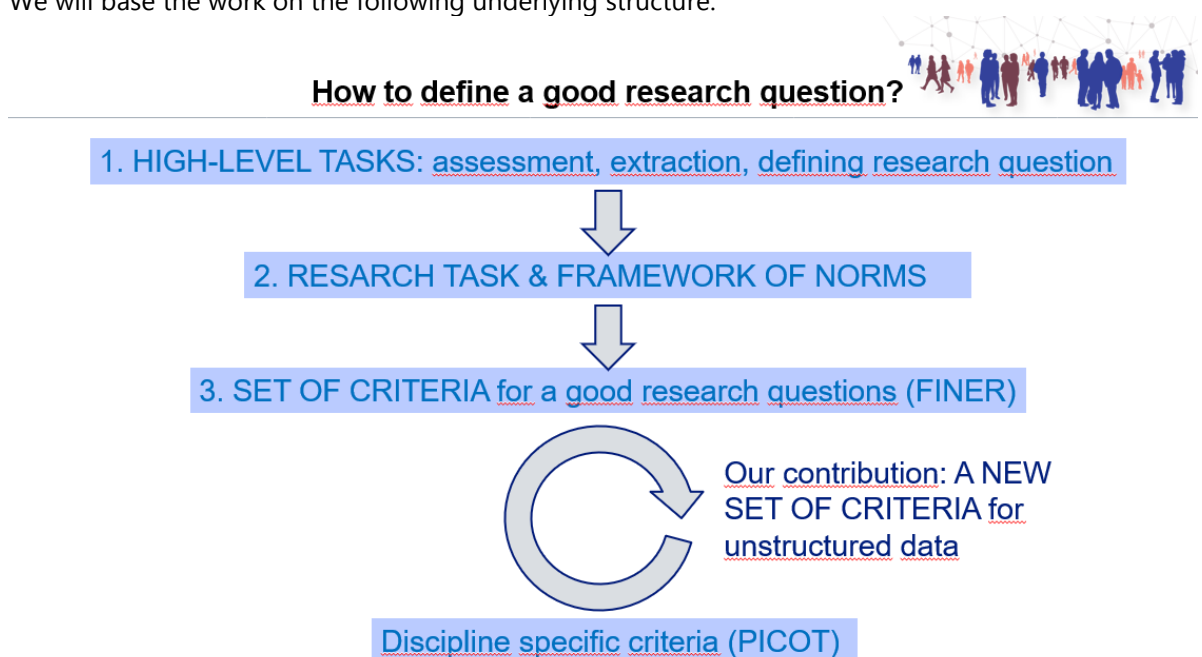
We believe that the topic of defining a good research question is generic, but to our knowledge only very limited teaching materials and courses are available for the health area appropriate for the digital revolution going on. Given the diversity of our and other DSI communities, as well as the excellent teaching platforms provided by DSI (Studium Digitale, DSI Excellence Program for PhDs) we believe that there is not only a strong need for such teaching materials but also a high chance of success for launching a sustainable course offering under the DSI-banner (e.g., online, as a MOOC), thereby also promoting the name and reputation of DSI at large.

With this extension project, we hope to make a long-lasting contribution to science education and capacity building at DSI and the UZH (as well as to promoting the work and expertise of the DSI health community). The project reflects the needs and interests within the Health Community. Furthermore, the project has the potential to foster the collaboration between Health Community and other DSI Communities such as Communication and Ethics Community.

Description and Roadmap

We will conduct a series of workshops to set the ground for defining a "**DSI Approach: How to Define a Good Research Question for Digital Unstructured Data**" and to create teaching tools. To this end, we envision initiating an interdisciplinary community effort in 2022 to develop a **checklist/framework** that will guide applied researchers in finding a feasible, relevant study question. The effort will result in a white paper / methods paper to be published in a peer-reviewed journal. Besides, the workshops will serve as a capacity building for researchers working with unstructured data and to allow an exchange between junior and senior researchers and the different health-related disciplines.

We will base the work on the following underlying structure:

How to define a good research question?

1. HIGH-LEVEL TASKS: assessment, extraction, defining research question

2. RESARCH TASK & FRAMEWORK OF NORMS

3. SET OF CRITERIA for a good research questions (FINER)

Our contribution: A NEW SET OF CRITERIA for unstructured data

Discipline specific criteria (PICOT)

Our contribution:

An interdisciplinary view on how to define research questions.
Achieving clarity on the feasibility and added value of inclusion of unstructured data
The following three main domains will be reflected in the DSI checklist:
1.   Criteria for suitable unstructured data/ how to evaluate suitability?
2.   Criteria for how to create a good match between research idea and unstructured data?
3.   Criteria for assessment of existing skills and resources

1 Workshops Overview

To define the checklist for finding a feasible, relevant study question, we envision integrating the following aspects and viewpoints from different scientific disciplines:
Wkshp 1: Terminology & Frameworks
Wkshp 2: UDE-FINER: Complementing the Checklist
Wkshp 3: Finalizing White Paper
Timeline (detailed)

January – Feb
- The core of the working group is established
- End of February & Beginning of March
  - **Wkshp 1**: **Terminology & Frameworks**

March – April
- Implementing feedbacks from WS 1 & feedback round
- Writing the White Paper's relevant part

May - June
- Beginning of May: **Wkshp 2: UDE-FINER: Complementing the Checklist**
- Implementing feedbacks from WS 2 & feedback round
- Writing the White Paper's relevant part
- Writing project proposal Lehrkredit & feedback round
- 30. June **submission Lehrkredit**

July – Sept
- Writing White Paper
- Preparation WS 3
- Sept: **Wkshp 3**: **UDE-FINER: Complementing the Checklist + Finalizing White Paper**

October – December
- Implementing Feedback WS 3 & feedback round
- Workshop with the DSI-Health and other DSI-Communities: **Ethical Issues (?)**
- Open Test/Feedback Call (?)
- Finishing White Paper and Publication

How will we proceed?

Once we establish the working group, we will define dates for the workshops. We expect from everyone to actively participate in at least one workshop. Before each workshop, we will ask participants to provide inputs and feedback for the relevant focus topic. We will work on the checklist and white paper collectively. All documents will be shared on Google Drive. Each active participant can be a co-author of the publication.
To sign up to the working group please use this document:
https://docs.google.com/spreadsheets/d/1OLiRaCLqbwuWUVXQZ9henM8HHusqOXmo/edit?usp=sharing&ouid=109036828346874317817&rtpof=true&sd=true