BUILDING A DIGITAL PLATFORM FOR TRACEABLE HISTORICAL RESEARCH

INTRODUCING KIARA

SEAN TAKATS & LAUREN COETZEE

@DHARPAproject
dharpa@uni.lu
dharpa.org

DIGITAL HISTORICAL SOURCES

A critical analysis of the evolution of these digital sources and how we can safeguard data authenticity, transparency and reproducibility

BUILDING A DIGITAL PLATFORM

Intro to Kiara
Aims of Project
Team Overview
Development process

The DHARPA Project

The Digital History Advanced Research Projects Accelerator

An interdisciplinary team working to enable critical DH practice.

One of the main goals is pulling the lid off "black box" technologies, creating and fostering more transparent, traceable and reproducible research.

Our team advocates for the agency and role of the researcher in the research process, and part of understanding this role in knowledge production is also creating tools to assist in making this process easier and the standard for historical practice in the digital age.

AIMS OF DHARPA

SOFTWARE & METHODOLOGIES

USER-FRIENDLY TOOLS TO OPERATIONALISE SELF-REFLEXIVE, REPRODUCIBLE RESEARCH

COLLABORATION

A RESPONSIBE LAB TO RAPIDLY PROTOTYPE & LAUNCH APPLIED RESEARCH SOLUTIONS

OUTREACH & INFRASTRUCTURE

TRAINING & PLATFORMS TO HELP START, SUSTAIN AND MAINTAIN DH RESEARCH PROJECTS

DHARPA TEAM

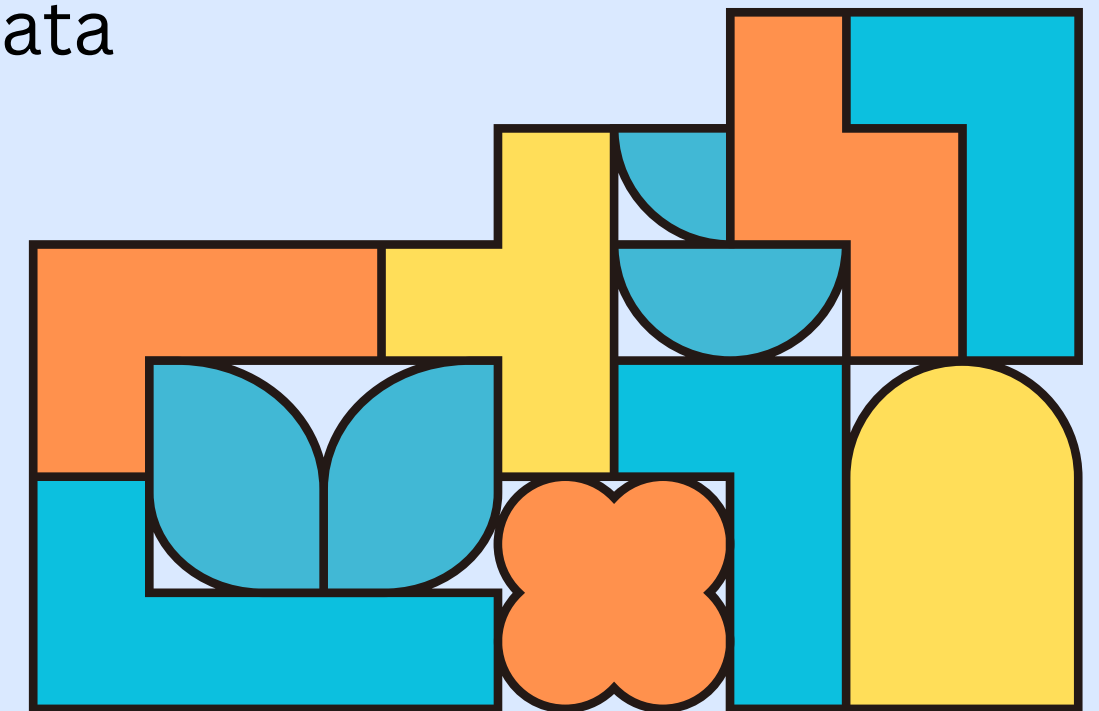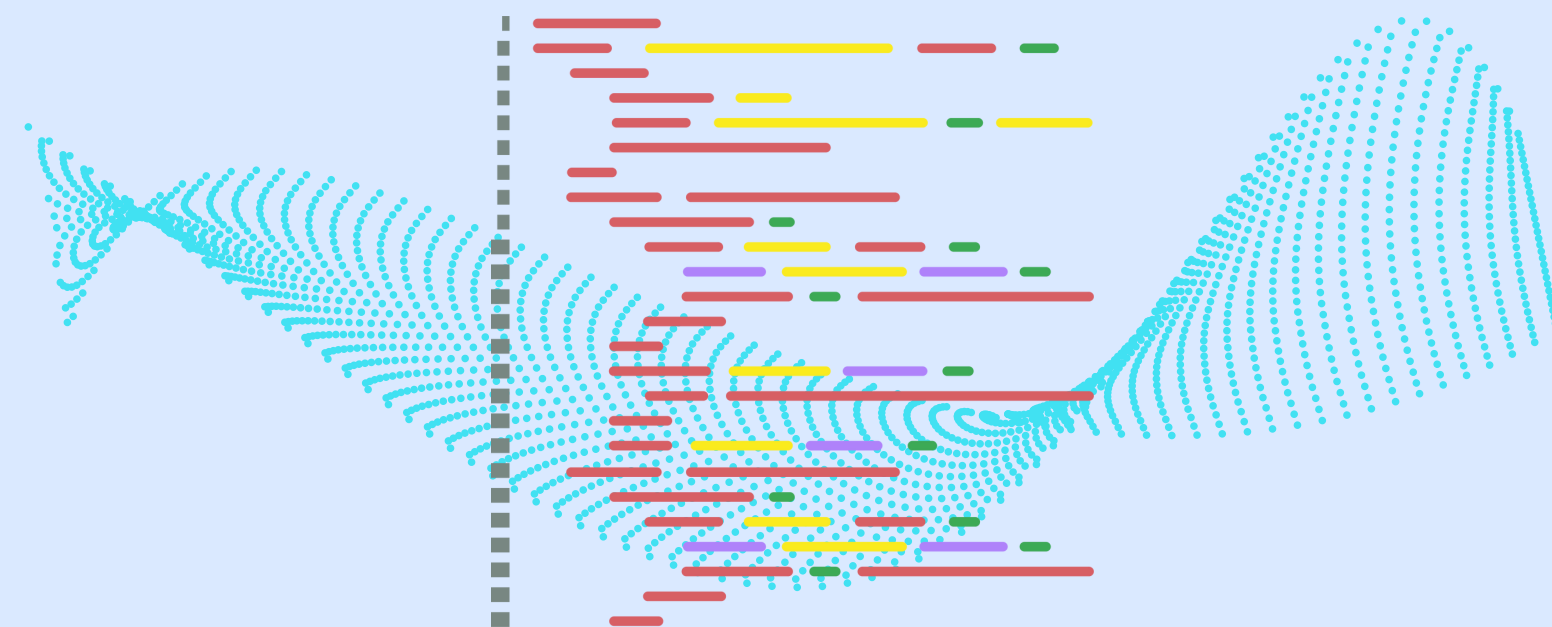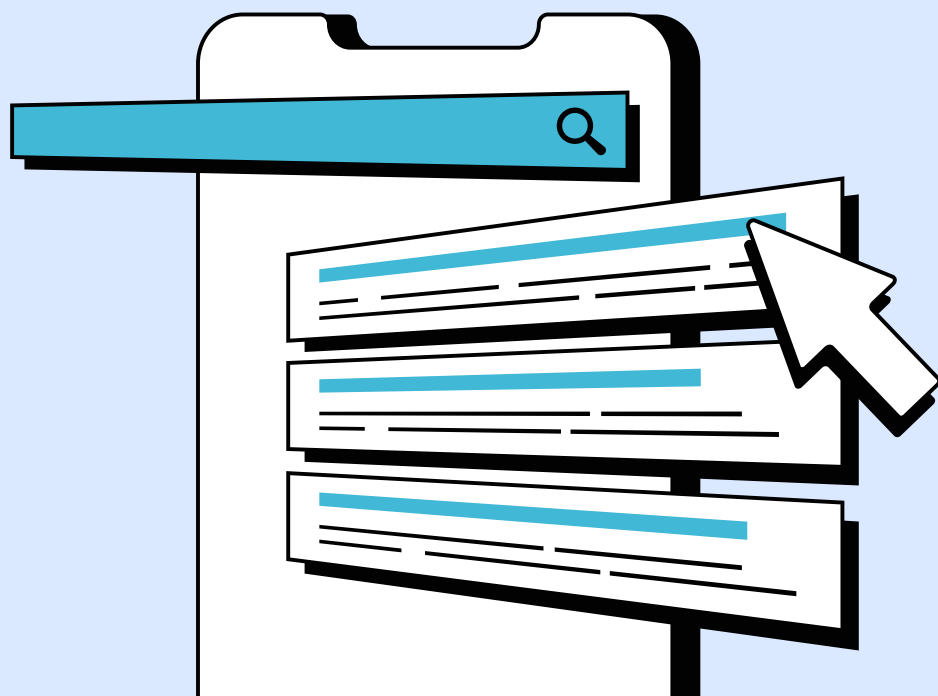|	
Sean Takats	Principal Investigator
Caitlin Burge	Co-Principal Investigator
Helena Jaskov	Co-Principal Investigator

Sandra Sabelus	Administrative Assistant
Caroline Appleby	Developer
Markus Binsteiner	Developer
Mariella de Crouy Chanel	Developer
Finola Finn	Post-doctoral Researcher
Luca Federico Cerra	PhD Candidate
Eliane Schmid	PhD Candidate
Lauren Coetzee	PhD Candidate

HELLO KIARA

A DATA ORCHESTRATION ENGINE

Using a modular approach, kiara lets users re-use tried and tested data orchestration pipelines, as well as create new ones from existing building blocks.

It also helps you manage your research data, and augment it with automatically-, semi-automatically-, and manually- created metadata

MODULES

LINEAGE

[6]: kiara.retrieve_operation_info('create.table.from.file')

[6]: Author(s)
              Markus Binsteiner   markus@frkl.io

     Context
              Tags   tabular
              Labels   package: kiara_plugin.tabular
              References   source_repo: https://github.com/DHARPA-Project/kiara_plugin.tabular
                            documentation: https://DHARPA-Project.github.io/kiara_plugin.tabular/

     Module type   create.table
     Module config
              "source_type"   "file"
              "target_type"   "table"

     Operation details
              Documentation   Create a table from a file, trying to auto-determine the format of said file.

                               Currently supported input file types:

                                 • csv
                                 • parquet

              Inputs

field name	type	description	Required	Default
file	file	The source value (of type 'file').	yes	-- no defau
first_row_is_header	boolean	Whether the first row of a (csv) file is a header row. If not provided, kiara will try to auto-determine. Ignored if not a csv file.	no	-- no defau

Outputs

field name	type	description
table	table	The result value (of type 'table').

kiara ia a data-orchestration framework; this is the command-line frontend for it.
For more information, visit the kiara homepage:
https://dharpa.org/kiara.documentation .

— Options —

--config   -cnf   TEXT   A kiara config file (or folder containing one named 'kiara.config').

--context   -ctx,-c TEXT   The name of the kiara context to use (or the path to a context file).

--pipelines   -p   TEXT   File(s) and extra pipel

--plugin   -P   TEXT   Ensure the package(s) virtual env

--use-background-service -b   Always use (start if n

--version   -v   Show the version of kiara and installed plugins, then exit.

--runtime-info   -ri   Show current environment information, then exit.

--help   -h   Show this message and exit.

— Commands —

archive   Kiara archive related sub-commands.
context   Kiara context related sub-commands.
data   Data-related sub-commands.
data-type   Information about available data types.
info   Kiara config related sub-commands.
module   Module-related sub-commands.
operation   Operation-related sub-commands.
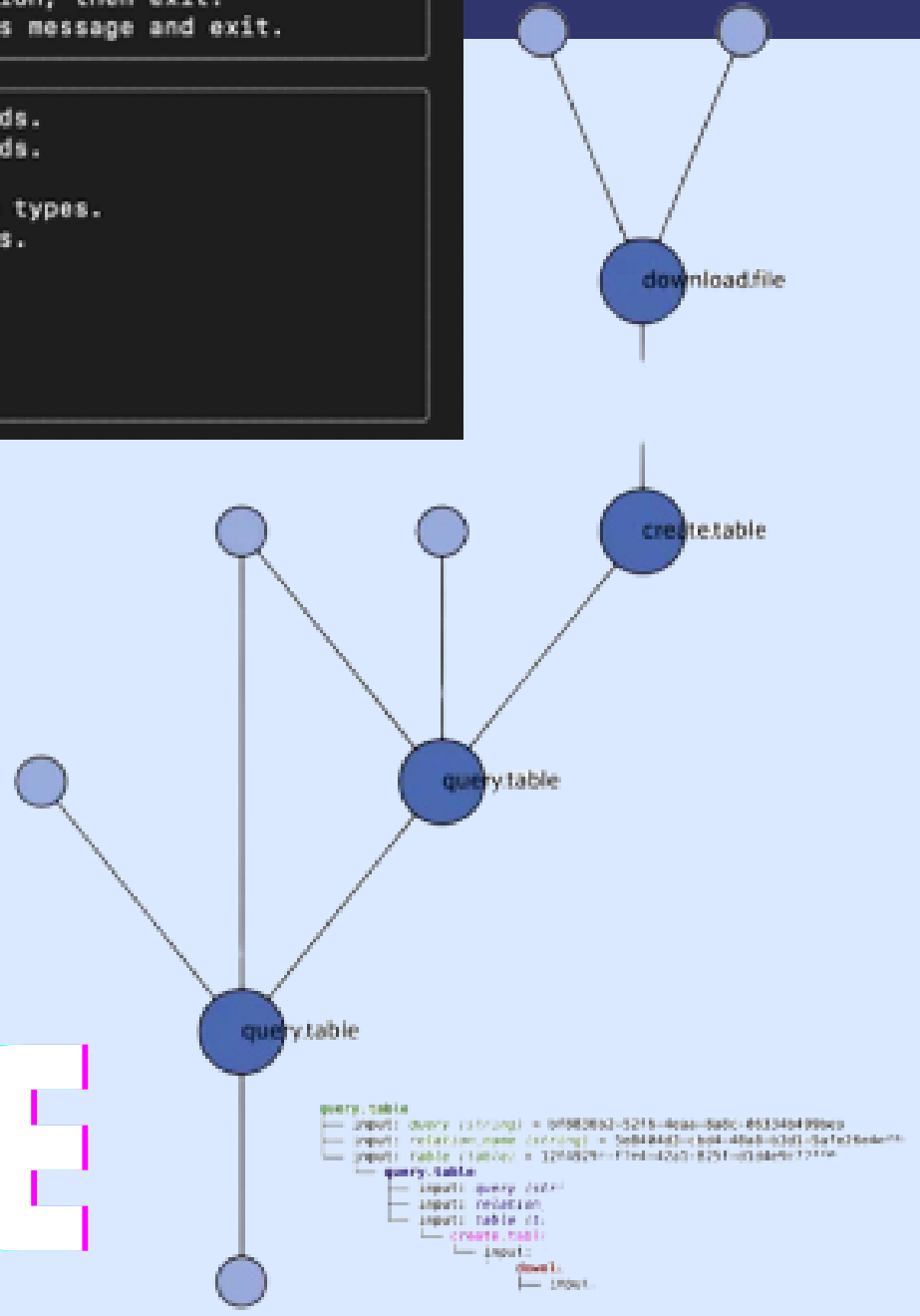pipeline   Pipeline-related sub-commands.
render   Render-related sub-commands.
run   Run a kiara operation.

download.file
create.table
query.table
query.table

[4]: GephiFile = kiara.run_job('import.local.file', inputs={'path':'/Users/eliane.schmid/Desktop/Gephi_Kiara.csv'}, comment = 'upload file')

[5]: GephiFile = GephiFile['file']
     GephiFile

[5]: value_id   8642cc8b-46c2-416d-a4db-3dd586cb6a18
     kiara_id   7ebe6f6f-1d7c-4172-acb6-058d16db9b00

     data_type_info
              data_type_name   file
              data_type_config   {
                                    "content_type": null
                                 }
              characteristics   {
                                    "is_scalar": false,
                                    "is_json_serializable": false
                                 }
              data_type_class
                                 python_class_name   FileValueType
                                 python_module_name   kiara.data_types.included_core_types.filesystem
                                 full_name   kiara.data_types.included_core_types.filesystem.FileValueType

     destiny_backlinks {}
     job_id   4140d43c-1886-492e-923c-4f93b1d7eb47
     property_links   {
                          "metadata.file": "b751e472-aa31-4e85-8e3d-4ae97aa5a027",
                          "metadata.python_class": "46031653-dbab-4492-88f3-566bb7eaecd4"
                        }
     value_created   2024-06-17 13:52:21.393082+02:00
     value_hash   zdpuAvM6pGa4fYHCUAqwtHtgstPLe5U8f8DDhKqhgg9FL6Sed
     value_schema
                        type   file
                        type_config {}
                        default   __not_set__
                        optional   False
                        is_constant False
                        doc   The loaded files.

     value_size   3.79 KB
     value_status   -- set --

digital humanities Benelux 8 — Brussels 2023

Welcome!

We're pleased to announce that the 10th edition of the DHBENELUX Conference will take place in Belgium at the Royal Library of Belgium / KBR.

The annual DHBENELUX Conference serves as a platform for the community of interdisciplinary Digital Humanities researchers to meet, present and discuss their latest research findings [...] demonstrate tools and [...]

Date
23rd of May – 2nd of June [...]

Location
Royal Library of Belgium [...]

Royal Library of Belgium / KBR

SEARCH

Call for sponsors!

DH Benelux is still looking for generous sponsors to support this 10th edition.

digital HISTORY 2023
24.–26.05.2023 - Humboldt-Universität zu Berlin

DIGITAL HISTORY TAGUNG 2023

Digitale Methoden in der geschichtswissenschaftlichen Praxis
Fachliche Transformationen und ihre epistemologischen Konsequenzen

Dies ist die zweite Tagung einer Konferenzreihe, die künftig alle zwei Jahre von der AG Digitale Geschichtswissenschaft im VHD mit [...]

Digital History: The Story So Far

As the field of Digital History continues to grow, so too does the number of tools, software, and coding packages built to support and advance digital history in practice. The range of this is at times staggering: from applications suitable for the most novice of digital historians, to coding guides and tools for those working to more nuanced and specific end-goals, researchers have an ability to engage with their materials in digital, quantitative ways on a never before seen level. Often we focus primarily on the new *findings* that come of out this new way of approaching research – but what about the ways we get to those findings?

Regardless of the type of digital analysis being performed or even the software being used, the process is normally the same: input some data, click some buttons or run some code (perhaps a couple of times over to edit the code and adjust the outcomes), and get your end result.

You've got an outcome – but do you know how you've got from a to b? It's likely that variables have been written over several times along the way, and the data has changed from one type to another, been filtered or added to, and decision after decision has been made without necessarily knowing it. Each little adjustment or re-run of the code has contributed to the research process and is critical to the end output or findings.

But how do we keep track?

Hello *kiara*.

Introducing *kiara*, a new data orchestration tool.

This new tool incorporates a number of different digital research approaches, but most importantly documents and encourages users to critically reflect on the process and use of DH tools. In doing so, the software opens up the black box of digital research, moving away from button-clicking software and making digital research more transparent and open to commentary, replicability, and criticism. It not only makes the research process itself more open, allowing users to visualise and examine the individual steps from start to finish, but also allows them to track changes to the data itself, something that is either imperceptible or, perhaps more importantly, *forgotten* about in traditional digital history methods and tools. kiara therefore acts as a 'wrapper' to this digital research process, tracking and documenting the steps and changes to the data, producing a veritable map of the journey that can be reflected upon and shared.

This tutorial will walk you through installation of *kiara* in Jupyter Notebooks, and some basic but essential functions that can be built on in further notebooks. At the end, it will showcase **the data lineage**, having tracked the research process and changes to the data from start to finish.

This tutorial requires you to know **python** and **SQL**.

Installation

Before running this notebook, you need to install Kiara and its dependencies in a virtual environment (such as Conda) by running the following command in your terminal:

DOCUMENTATION & WORKSHOPPING

REFLECTIONS AS A HISTORIAN

TRANSITION FROM TRADITIONAL TO DIGITAL HISTORICAL RESEARCH

CREATING AND INVESTING IN RESPONSIBLE RESEARCH

MAKE DIGITAL RESEARCH METHODS EASY TO UNDERSTAND AND SHARE

WHAT TOOLS SO WE NEED AS DIGITAL HISTORIANS?

ENCOURAGE THE CHOICES AND AGENCY OF RESEARCHERS

METHODS THAT ADD VALUE IN THE LONG RUN

GOING FORWARD

- Atomising research methods into reusable, stable components
- No more descriptivist analytical workflows with ever-smaller, discrete tasks
- Evade descriptive research in favour of reflective research
- Evoke chiaroscuro, the play of light and shadow, to shed light on the dark spaces of digital research

REFERENCES LIBRARY