

- Exploring plain text files
  - In-class 2: Explore a new book from [Project Gutenberg](#)
- Stages of data
- Downloading Anaconda and Python
  - Anatomy of a Python Script

# Thinking about data

What makes up research data?

- Research data can be understood as **materials or information necessary to come to your conclusion** but what they are depend on your project

# Thinking about data

Object Number	Is Highlight	Is Timeline View	Is Public Domain	Object ID	Department	Accession Number	Object Name	Title	Culture	Period	Dynasty	Reign	Portfolio	Artist Role
1	1981.1005.40	FALSE	FALSE	382014	Drawings and	1981	Drawing	Designs for a	NA	NA	NA	NA	NA	Architect
2	51.644.78(44)	FALSE	FALSE	840183	Drawings and	1951	Photographs	Harold Brown	NA	NA	NA	NA	NA	Architect/Artist
3	51.644.98	FALSE	FALSE	835020	Drawings and	1951	Photographs	Photograph of	NA	NA	NA	NA	NA	Architect/Photographer
4	205.234.3	FALSE	FALSE	32425	Arms and Arm	1828	Book	German, P	NA	NA	NA	NA	NA	Armorer/Engraver
5	63.712.76	FALSE	FALSE	388084	Drawings and	1963	Drawing	Design for a	NA	NA	NA	NA	NA	Artist
6	14.42	FALSE	FALSE	334912	Drawings and	1914	Drawing	Portrait of B	NA	NA	NA	NA	NA	Artist
7	47.37.15	FALSE	FALSE	209133	Photographs	1947	Photograph	Painting of O	NA	NA	NA	NA	NA	Artist
8	2015.336	FALSE	FALSE	747263	Photographs	2015	Photograph	Studio Portra	NA	NA	NA	NA	NA	Artist
9	2018.502.9	FALSE	FALSE	771843	Photographs	2018	NA	Donna de	NA	NA	NA	NA	NA	Artist
10	1P318	FALSE	FALSE	38780	Asian Art	1918	Print	NA	Japan	NA	NA	NA	NA	Artist
11	62.650.163	FALSE	FALSE	385970	Drawings and	1962	Print	In the Provin	NA	NA	NA	NA	NA	Artist
12	2019.67	FALSE	FALSE	820845	Drawings and	2019	Drawing	In the Park, I	NA	NA	NA	NA	NA	Artist
13	2017.276	FALSE	FALSE	738855	Modern and O	2017	Unlimited	Unlimited	NA	NA	NA	NA	NA	Artist
14	1986.287.465	FALSE	FALSE	36394	Asian Art	1986	Folding fan	现代 康博	China	NA	NA	NA	NA	Artist
15	21.10.1(3)	FALSE	FALSE	727538	Drawings and	1921	Print	Grotesque P	NA	NA	NA	NA	NA	Artist
16	28.27.3	FALSE	FALSE	428919	Drawings and	1926	Print	Anne of Aust	NA	NA	NA	NA	NA	Artist
17	1985.1123.1	FALSE	FALSE	264364	Photographs	1985	Carte-de-visi	Man with Ba	NA	NA	NA	NA	NA	Artist
18	19.76.127	FALSE	FALSE	725391	Drawings and	1919	Print	Le Bastien	NA	NA	NA	NA	NA	Artist
19	50.575.32	FALSE	FALSE	429959	Drawings and	1950	Print	Solène Sime	NA	NA	NA	NA	NA	Artist
20	87.539.106	FALSE	FALSE	838778	Drawings and	1967	Print	La Gilette	NA	NA	NA	NA	NA	Artist
21	87.762.101	FALSE	FALSE	348243	Drawings and	1967	Drawing	Mes Cogna	NA	NA	NA	NA	NA	Artist
22	49.97.372	FALSE	FALSE	370509	Drawings and	1949	Print	Sloth	NA	NA	NA	NA	NA	Artist
23	1978.492.1a	FALSE	FALSE	481944	Modern and C	1978	Table	Table	NA	NA	NA	NA	NA	Artist
24	52.570.200	FALSE	FALSE	346400	Drawings and	1952	Drawing	Frontal View	NA	NA	NA	NA	NA	Artist
25	32.35(139)	FALSE	FALSE	400044	Drawings and	1932	Print	The Second	NA	NA	NA	NA	NA	Artist
26	1993.253	FALSE	FALSE	335534	Drawings and	1993	Drawing	The Soldier	NA	NA	NA	NA	NA	Artist
27	1981.1229.57	FALSE	FALSE	289091	Photographs	1981	Carte-de-visi	Portrait of M	NA	NA	NA	NA	NA	Artist
28	1981.1158.42	FALSE	FALSE	284078	Photographs	1981	Photograph	[Man]	NA	NA	NA	NA	NA	Artist
29	2007.486.4	FALSE	FALSE	362299	Drawings and	2007	Print	Carnival	NA	NA	NA	NA	NA	Artist
30	69.95.2622	FALSE	FALSE	334134	Drawings and	1940	Print	Christ on the	NA	NA	NA	NA	NA	Artist
31	53.652.11	FALSE	FALSE	362814	Drawings and	1953	Drawing	Design and p	NA	NA	NA	NA	NA	Artist
32	1975.535.2	FALSE	FALSE	261257	Photographs	1975	Photograph	Doll Head II	NA	NA	NA	NA	NA	Artist
33	1994.254.241	FALSE	FALSE	273572	Photographs	1994	Negative	[Peggy Hoff	NA	NA	NA	NA	NA	Artist
34	61.53	FALSE	FALSE	341618	Drawings and	1961	Drawing	Scene from I	NA	NA	NA	NA	NA	Artist
35	1987.1194.39	FALSE	FALSE	265862	Photographs	1987	Photograph	[Munir Rakey	NA	NA	NA	NA	NA	Artist
36	58.504.10	FALSE	FALSE	295959	Photographs	1950	Photographs	[Woman in P	NA	NA	NA	NA	NA	Artist
37	1994.258.473	FALSE	FALSE	275929	Photographs	1994	Negative	[Carved Esq	NA	NA	NA	NA	NA	Artist
38	33.167	FALSE	FALSE	51584	Asian Art	1833	Handscroll	图卷	China	NA	NA	NA	NA	Artist

STRUCTURED

The Souls of Black Folk

by W. E. B. Du Bois

Herein is Written

- The Forethought
- I. Of Our Spiritual Freedom
- II. Of the Dawn of Freedom
- III. Of Mr. Booker T. Washington
- IV. Of the Meaning of Progress
- V. Of the Wings of Atalanta
- VI. Of the Training of Black Men
- VII. Of the Black Belt
- VIII. Of the Quest of the Golden Age
- IX. Of the Sons of Master and Slave
- X. Of the Faith of the Fatherless
- XI. Of the Passing of the First Birth
- XII. Of Alexander Crummell
- XIII. Of the Coming of John
- XIV. Of the Sorrow Songs
- The Afterthought

To  
Burghardt and Yolande  
The Lost and the Found

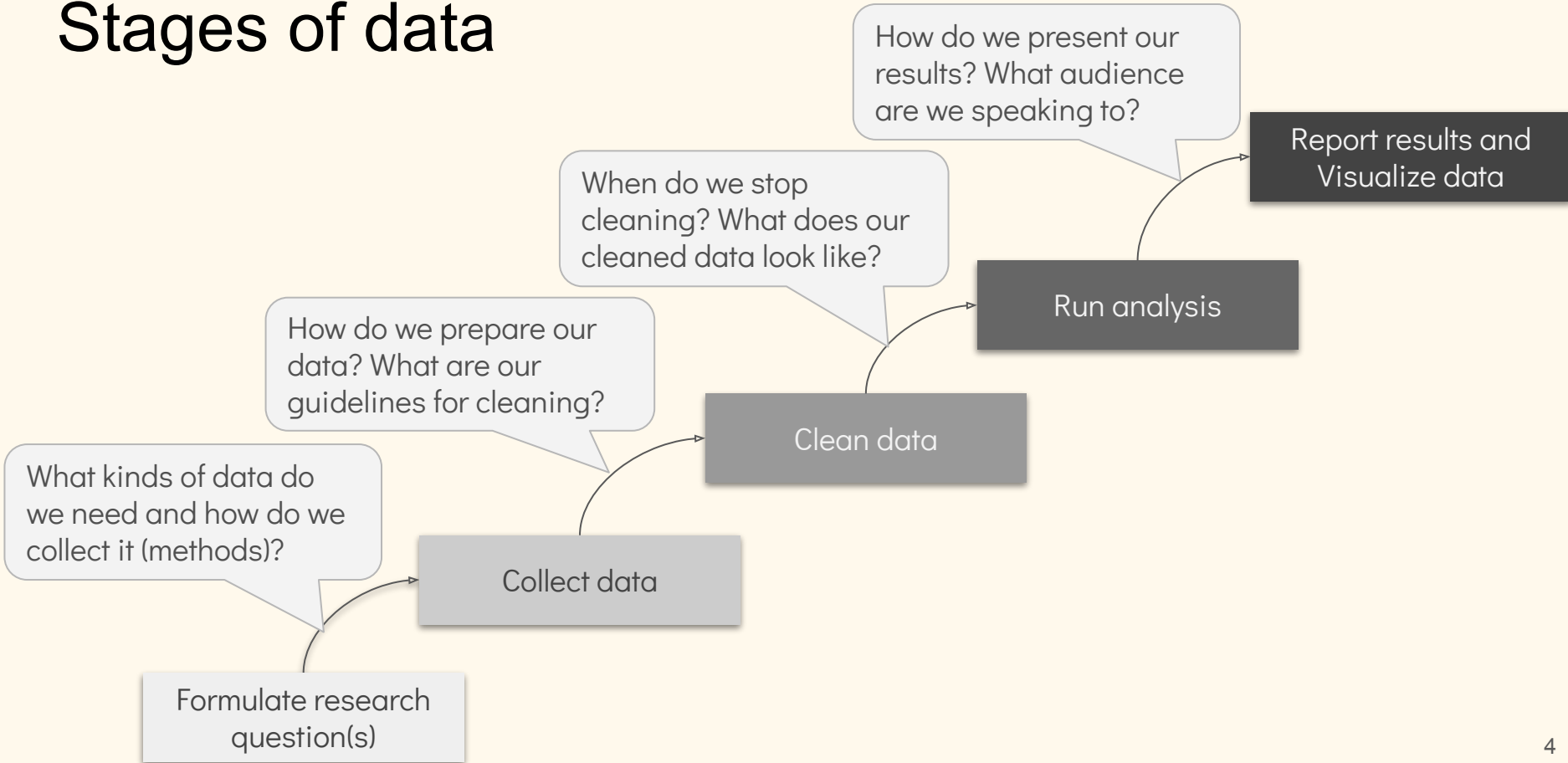
The Forethought

Herein lie buried many things which if read with patience may show the strange meaning of being black here at the dawning of the Twentieth Century. This meaning is not without interest to you, Gentle Reader; for the problem of the Twentieth Century is the problem of the color line. I pray you, then, receive my little book in all charity, studying my words with me, forgiving mistake and foible for sake of the faith and passion that is in me, and seeking the grain of truth hidden there.



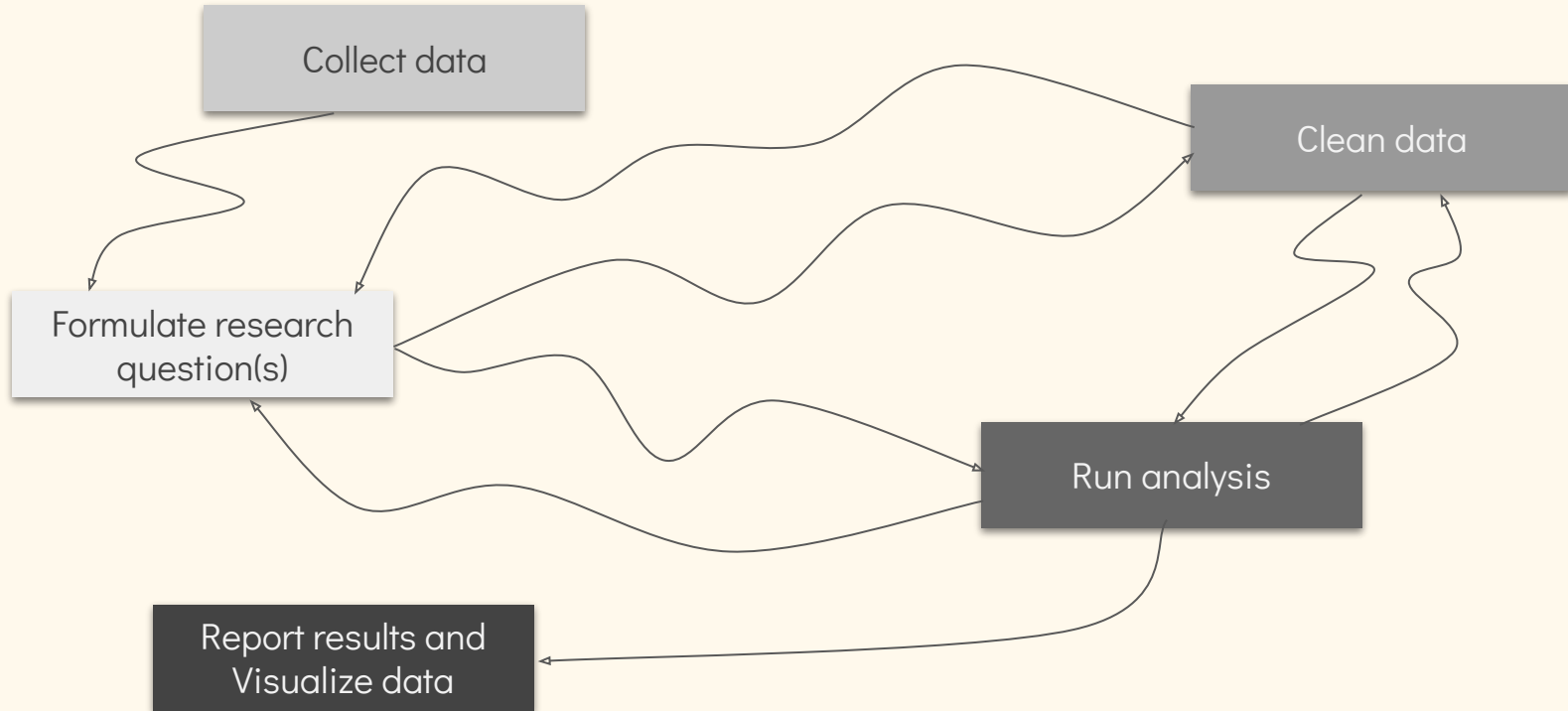
UNSTRUCTURED

# Stages of data





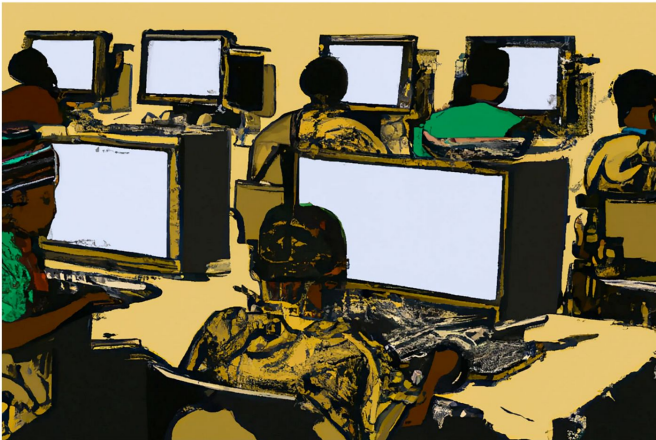
# Stages of data



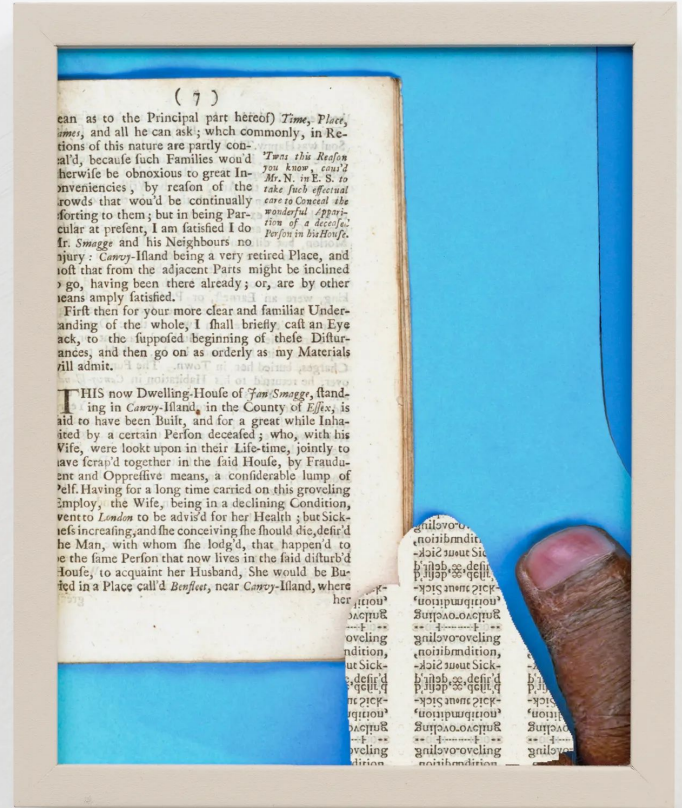
# Data and labor

## Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less

15 MINUTE READ



This image was generated by OpenAI's image-generation software, Dall-E 2. The prompt was: "A seemingly endless view of African workers at desks in front of computer screens in a printmaking style." TIME does not typically use AI-generated art to illustrate its stories, but chose to in this instance in order to draw attention to the power of OpenAI's technology and shed light on the labor that makes it possible. Image generated by Dall-E 2/OpenAI



Data is *people*...

...and there isn't a  
risk-free approach