

Digital Humanities Across Borders

Class 13: Creating and Cleaning Structured Data

Shameless promo



The image shows the Stanford Libraries website header. It features the Stanford Libraries logo (a circular emblem with a floral design) and the text "Stanford LIBRARIES". To the right is a search bar with a magnifying glass icon.

Navigation menu: About, Libraries, Using the libraries, Collections, Research support, Ask us.

Breadcrumbs: Home » About » Workshops » Documenting Workflow and Data Management

Documenting Workflow and Data Management

RSS DATE AND TIME February 26, 2019
LOCATION Catherine Tierney Seminar Room (Green Library 121A)
ADMISSION Free and open to all Stanford affiliates. Registration is required and opens February 12: [Data Mgmt](#)
AUDIENCE Faculty/Staff
Students

EVENT SPONSOR Stanford University Libraries
CONTACT contact-cidr@stanford.edu

MORE INFORMATION <http://cidr.stanford.edu>

February 2019						
Su	Mo	Tu	We	Th	Fr	Sa
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28		

Shameless promo

My Account Feedback

 **Stanford LIBRARIES**



[About](#) [Libraries](#) [Using the libraries](#) [Collections](#) [Research support](#) [Ask us](#)

[Home](#) » [About](#) » [Workshops](#) » Data Modeling: From Research Materials to Data

Data Modeling: From Research Materials to Data

RSS

DATE AND TIME March 5, 2019
3:00pm to 5:00pm



LOCATION Catherine Tierney Seminar Room (Green Library 121A)

ADMISSION Free and open to all Stanford affiliates. Registration is required and opens February 19: [Data Modeling](#)

AUDIENCE Faculty/Staff
Students

EVENT SPONSOR Stanford University Libraries

CONTACT contact-cidr@stanford.edu

MORE INFORMATION <http://cidr.stanford.edu>

Su	Mo	Tu	We	Th	Fr	Sa
						1 2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28		

Data is like fruit



And code is like a child



Feeding fruit to children



Feeding fruit to children



You have to wash, peel & cut it up



This can be necessary when combining multiple tools

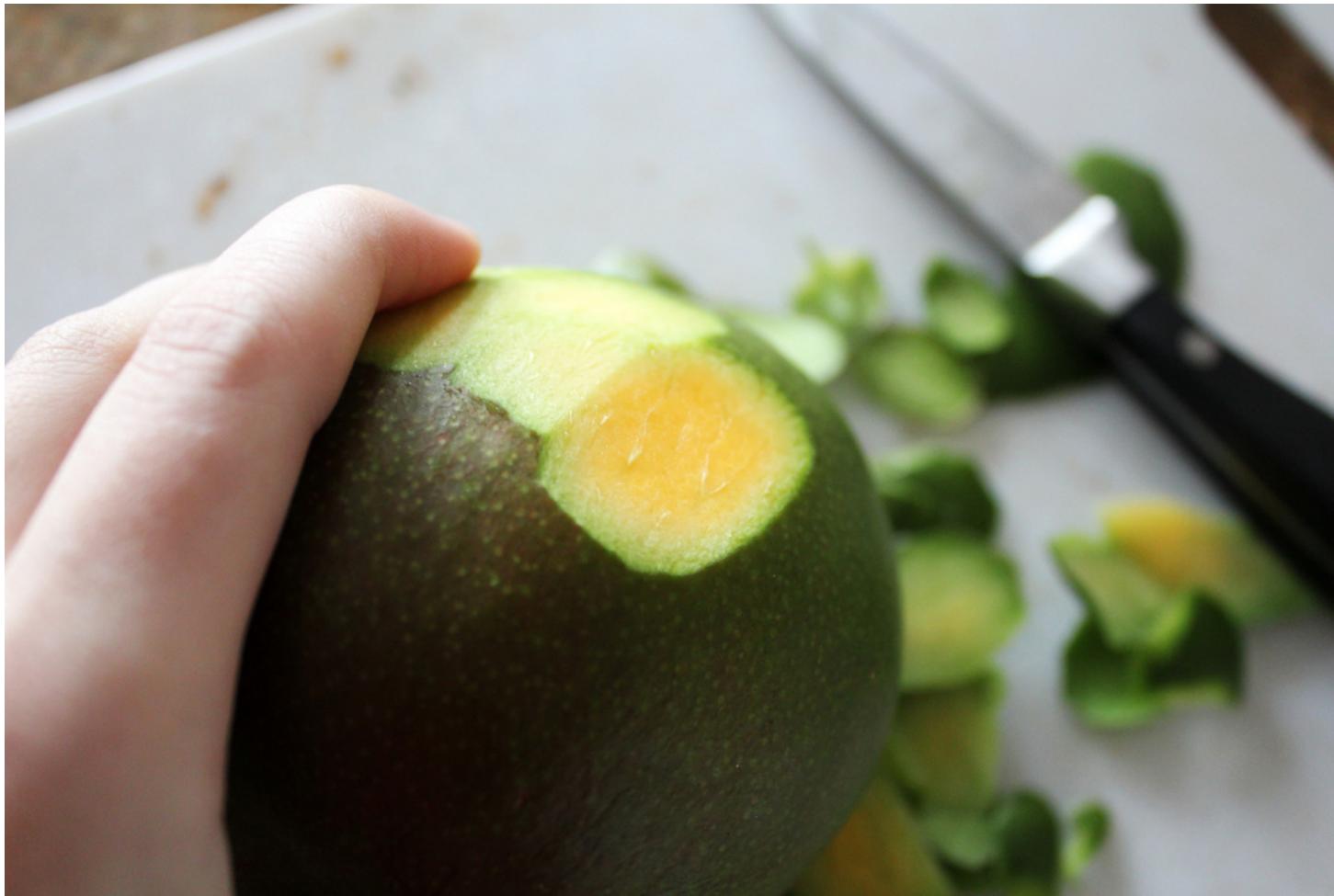


The “tool problem” in DH

cious archives are beyond the
scope of this talk.

I'm an academic — I don't
need to {get the job done.
produce working code.

OpenRefine is your knife for chopping data-fruit



Fanfic as “public literary analysis”



Magical Guide to HP



Harry Potter and
Colleagues of his
Brother Soren (?!)



Hogwarts Guide to Living

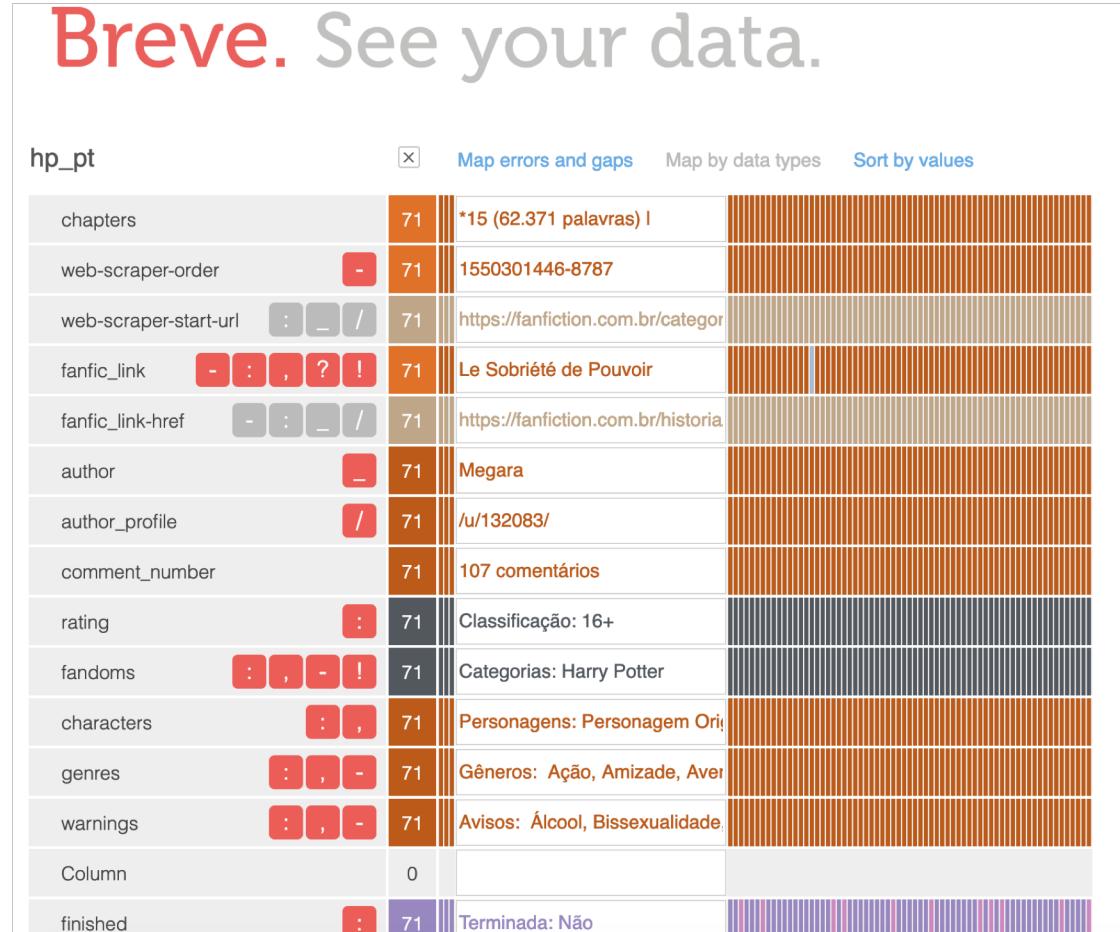


Harry Potter and the
Western Dream

Metadata varies

	Italian	Japanese	Portuguese	Spanish	Russian
Characters	X	X	X		X
Character pairs	X				X
Genres	X		X	X	X
Rating	X		X		X
Length		X		X	X
Comments			X	X	
Readers				X	
Status					X
Votes/Faves		X			

Preview your file



<http://breve.designhumanities.org/app/>