



HARMONY



Harmony

Open source NLP/AI tool for psychologists

harmonydata.ac.uk

github.com/harmonydata/

Thomas Wood

 **Fast Data Science** fastdatascience.com



Harmony

Harmony is an open source tool for psychologists to use for free around the world

Uses Natural Language Processing (NLP)

MIT License

It's not a monetised product

Python+R libraries

Wellcome funded



MEET THE TEAM



**Eoin
McElroy**

Psychologist
Ulster
University



**Bettina
Moltrecht**

Psychologist
University College
London



**Mauricio
Hoffmann**

Psychologist
Universidade
Federal de Santa
Maria



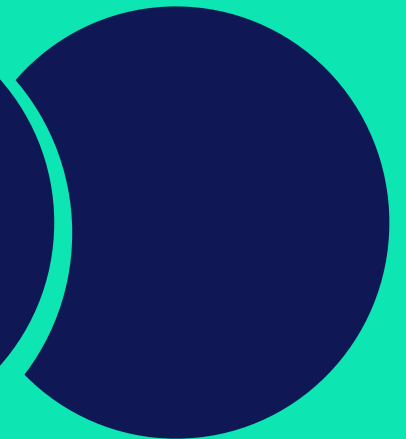
**George
Ploubidis**

Psychologist
University College
London



**Thomas
Wood**

Data science/software
Fast Data
Science



**John
Rogers**

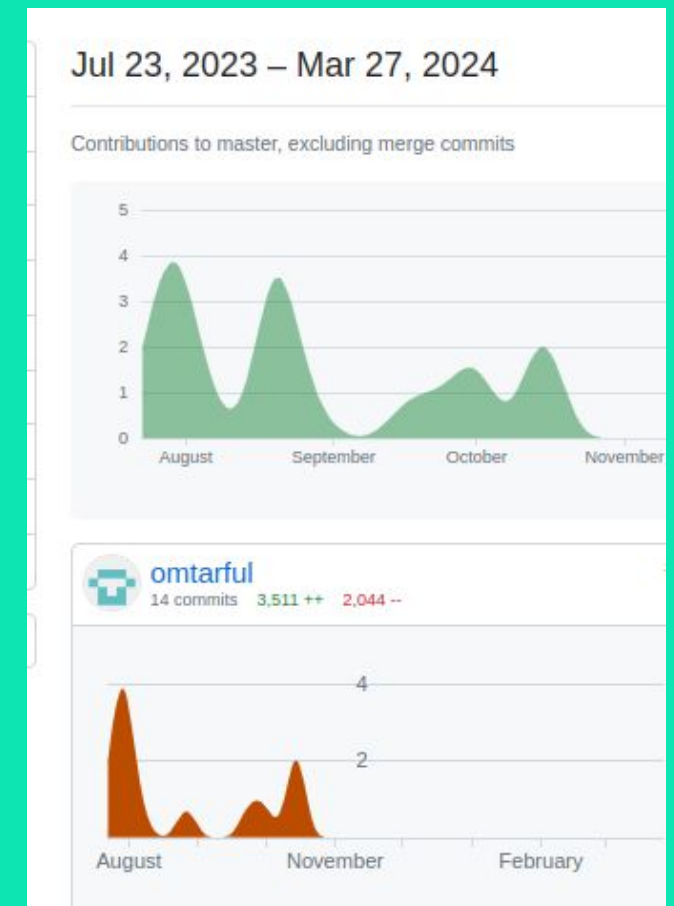
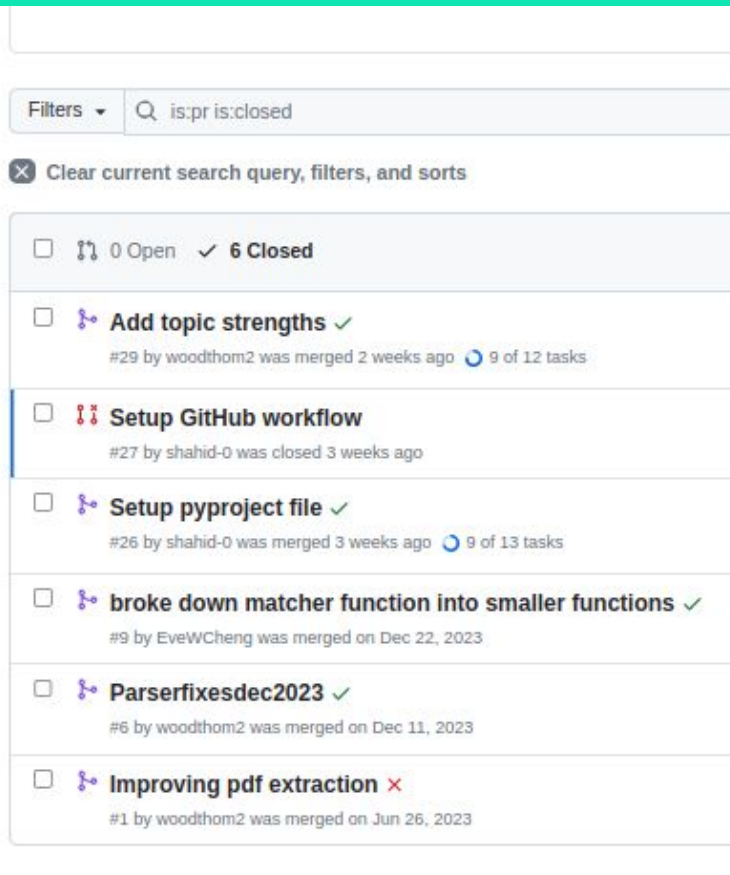
Front end
development



MEET THE TEAM

...plus people around the world who have made pull requests to Github!

shahid-0
EveWCheng
0x48piraj
nikhildevre
omtarful



Partners





Examples of open source projects



Ubuntu

Mix of licences



Firefox

Mozilla License

pandas Python library → BSD 3 License

psych, R package → GPL-3 License

 HARMONY → MIT License

GAD-7 Anxiety

Over the <u>last two weeks</u> , how often have you been bothered by the following problems?	Not at all	Several days	More than half the days	Nearly every day
1. Feeling nervous, anxious, or on edge	0	1	2	3
2. Not being able to stop or control worrying	0	1	2	3
3. Worrying too much about different things	0	1	2	3
4. Trouble relaxing	0	1	2	3
5. Being so restless that it is hard to sit still	0	1	2	3
6. Becoming easily annoyed or irritable	0	1	2	3
7. Feeling afraid, as if something awful might happen	0	1	2	3

GAD-7 Anxiety

Over the <u>last two weeks</u> , how often have you been bothered by the following problems?	Not at all	Several days	More than half the days
1. Feeling nervous, anxious, or on edge	0	1	2
2. Not being able to stop or control worrying	0	1	2
3. Worrying too much about different things	0	1	2
4. Trouble relaxing	0	1	2
5. Being so restless that it is hard to sit still	0	1	2
6. Becoming easily annoyed or irritable	0	1	2
7. Feeling afraid, as if something awful might happen	0	1	2

Column totals ____ + ____ + ____

Total score

Beck's Anxiety Inventory

Numbness or tingling

Feeling hot

Wobbliness in legs

Unable to relax

Fear of worst happening

Dizzy or lightheaded

Heart pounding / racing

Unsteady

Terrified or afraid

Nervous

Feeling of choking

Hands trembling

Shaky / unsteady

Fear of losing control

Difficulty in breathing

Fear of dying

Scared

Indigestion

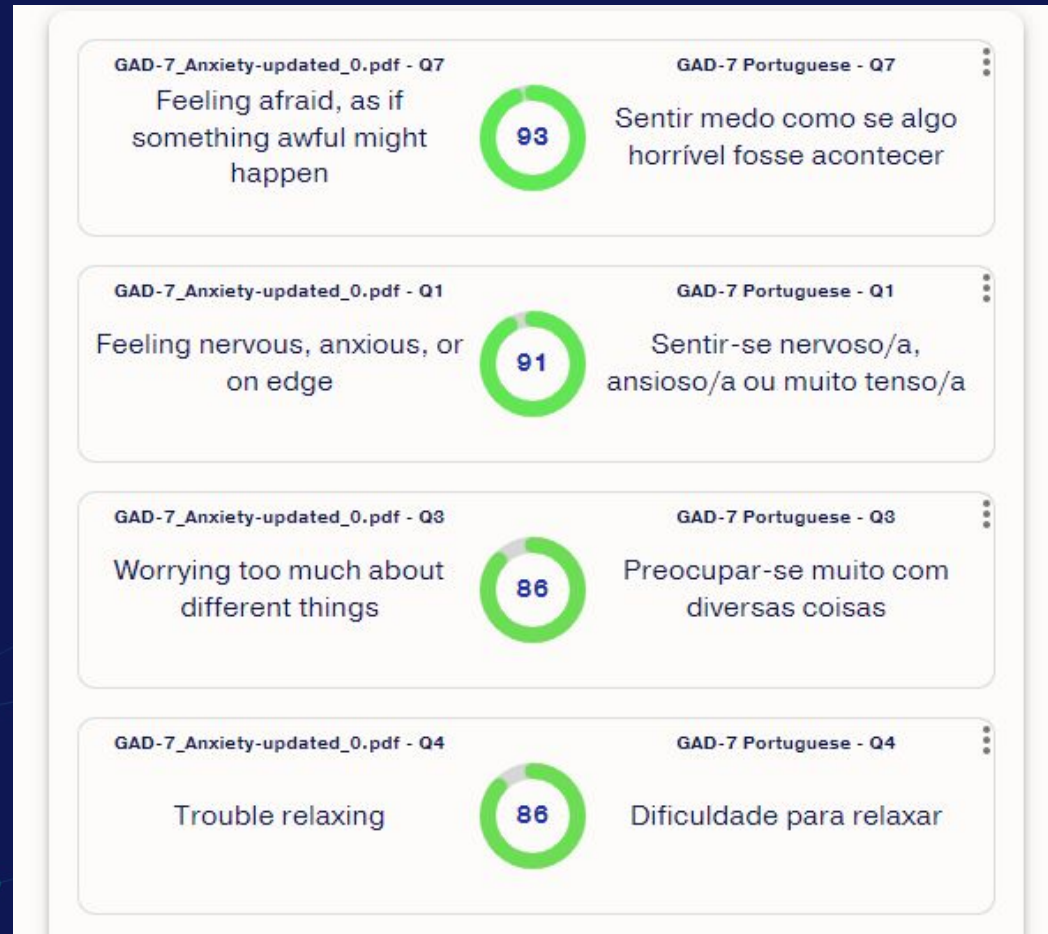
Faint / lightheaded

Face flushed

Hot / cold sweats

	A	B	C	D	
1	Measure	Cohort	Age (Range)	Age (Year)	Low Mood
2	SF-36 (10 items)	ALSPAC	20s	18	6. Have you felt downhearted and person
3	MFQ	ALSPAC	20s	18	1. I felt miserable or unhappy / 3. felt happy
4	SF-36 (10 items)	ALSPAC	20s	21	6. Have you felt downhearted and person
5	MFQ	ALSPAC	20s	21	1. I felt miserable or unhappy / 3. felt happy
6	MFQ	ALSPAC	20s	22	1. I felt miserable or unhappy / 3. felt happy
7	Malaise Inventory (24-item version)	NCDS	20s	23	3. Do you often feel depressed?
8	MFQ	ALSPAC	20s	23	1. I felt miserable or unhappy / 3. felt happy

Use Harmony for fast harmonisation across languages



GAD-7_Anxiety-updated_0.pdf - Q7 Feeling afraid, as if something awful might happen	93	GAD-7 Portuguese - Q7 Sentir medo como se algo horrível fosse acontecer
GAD-7_Anxiety-updated_0.pdf - Q1 Feeling nervous, anxious, or on edge	91	GAD-7 Portuguese - Q1 Sentir-se nervoso/a, ansioso/a ou muito tenso/a
GAD-7_Anxiety-updated_0.pdf - Q3 Worrying too much about different things	86	GAD-7 Portuguese - Q3 Preocupar-se muito com diversas coisas
GAD-7_Anxiety-updated_0.pdf - Q4 Trouble relaxing	86	GAD-7 Portuguese - Q4 Dificuldade para relaxar

A world map composed of white dots on a blue background. The map is centered and occupies most of the frame. A white rectangular box is overlaid horizontally across the middle of the map, containing a URL.

<https://harmonydata.ac.uk/app>

DEVELOPMENT OF HARMONY

System Usability Scale

INSTRUÇÕES: Para cada uma das afirmações abaixo, selecione a caixa (apenas uma por linha) que melhor descreve a sua reação ao website que acabou de ver.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. I needed to learn a lot of things before I got going with this system.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. I think I would like to use this system frequently.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. I think I would like to learn more about this system.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. I would like to use this system often.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. I would like to use this system every day.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. I would like to use this system a lot.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. I would like to use this system a great deal.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. I would like to use this system a very great deal.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. I would like to use this system a great deal.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. I would like to use this system a very great deal.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Read to question

Compare with Son

make revis

tracks any edits for orgs

"edit function" "available changes" profile

The screenshot shows the HARMONY application interface. At the top left is the HARMONY logo. Below it is a section titled "Options" with a "Match Threshold" slider and a "Show within instrument matches" button. To the right of the slider, a handwritten note asks "what does that mean?". Below the options section, there is a list of items with associated scores in circles. Handwritten notes include "show 'antonyms'" and "what is this for?". On the right side of the interface, there is a vertical list of items with scores, and a handwritten note "edit!" with a downward arrow, and another note "sorted by similarity" with a downward arrow.

Options

Match Threshold

Show within instrument matches

Many fears, easily scared 74

Often fights with other children or bullies them 77

Often unhappy, down-hearted or tearful 70

Often unhappy, down-hearted or tearful 76

Feeling nervous, anxious, or on edge 71

Not being able to sleep or control worrying 76

Feeling afraid, as if something might happen

Picked on or bullied by other children

Becoming easily annoyed or irritable

Feeling down, depressed or hopeless

Feeling afraid, as if something awful might happen

Worrying too much about different things

edit!

sorted by similarity

low

How does Harmony work?

1. Process the PDF to get the question items out
2. Convert all questions to sentence embeddings
3. Calculate the cosine similarity



How does Harmony work?

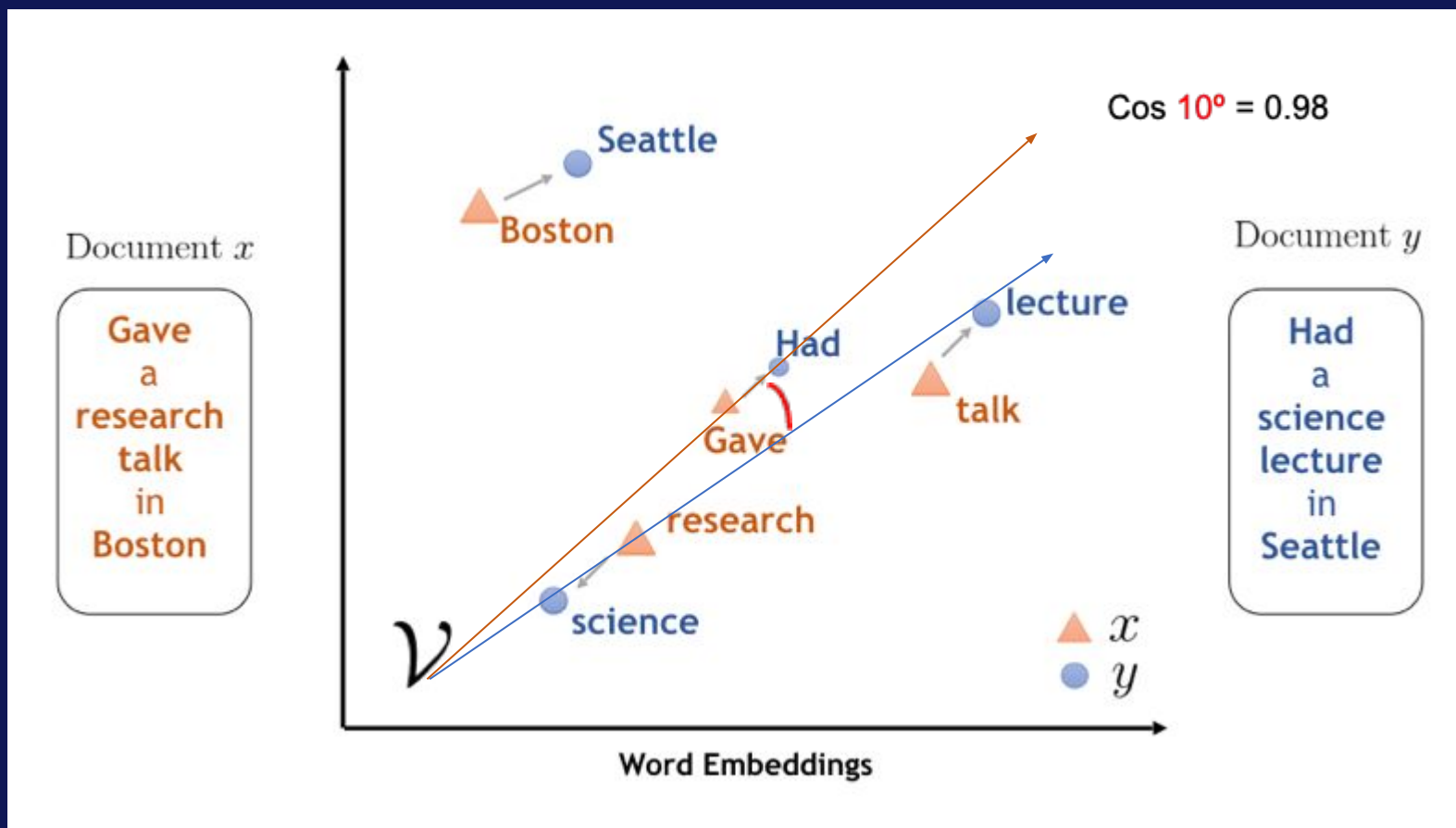
Matches items by semantic content using Natural Language Processing

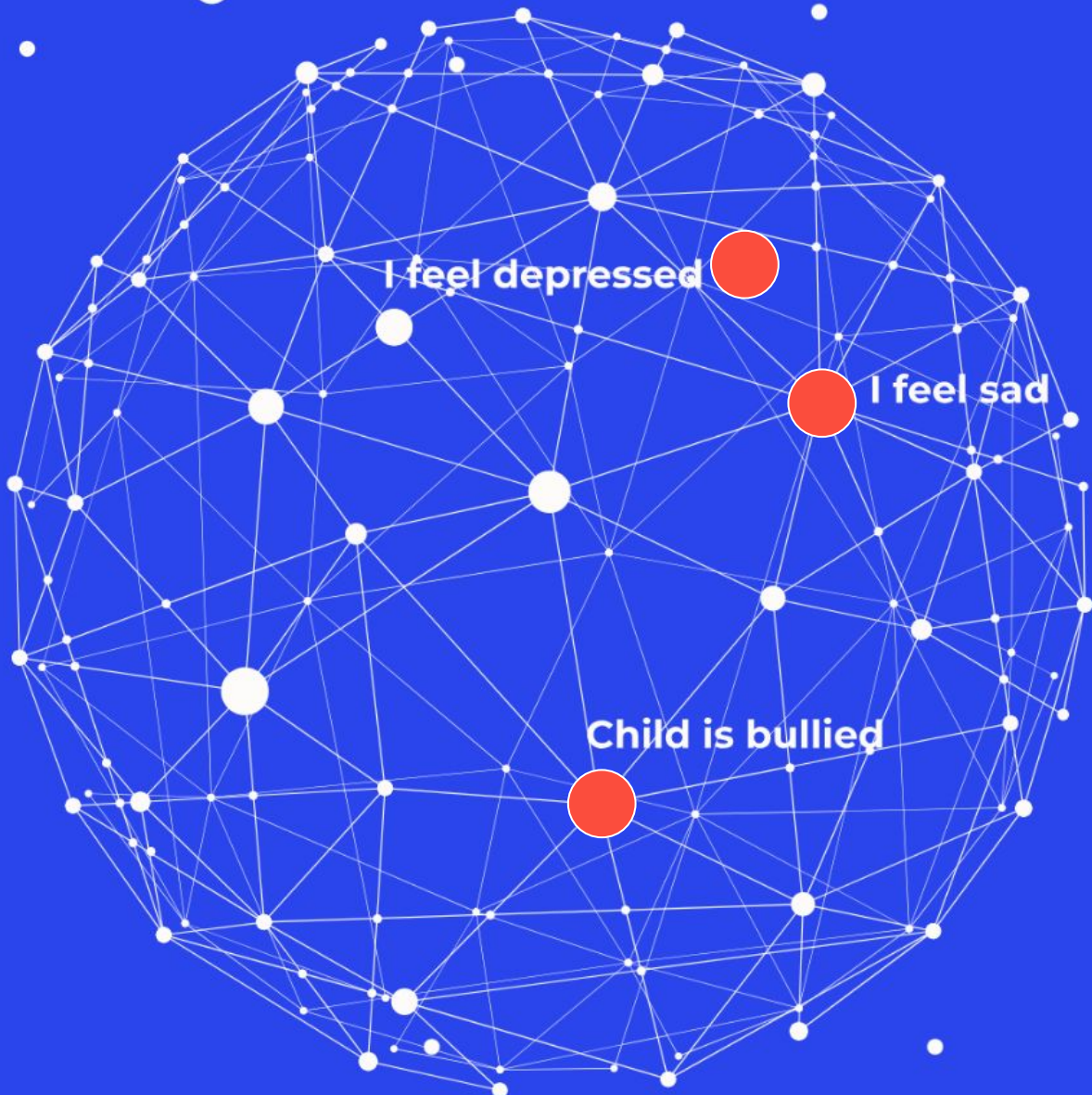
Steps:

1. Break sentences into words (tokens)
2. Attribute values to words (vectors)
3. Place them into a vector space
4. Place synonyms in close coordinates
5. Use sentence Transformer (neural network) to combine the vectors of an entire text (Attention model)

How does Harmony work?

- Matches items by semantic content using Natural Language Processing
- Transformer model converts semantic words into numeric vectors





Sentences have a related meaning

Vectors point in a similar direction

Cosine score is high – close to 1



Sentences have unrelated meaning

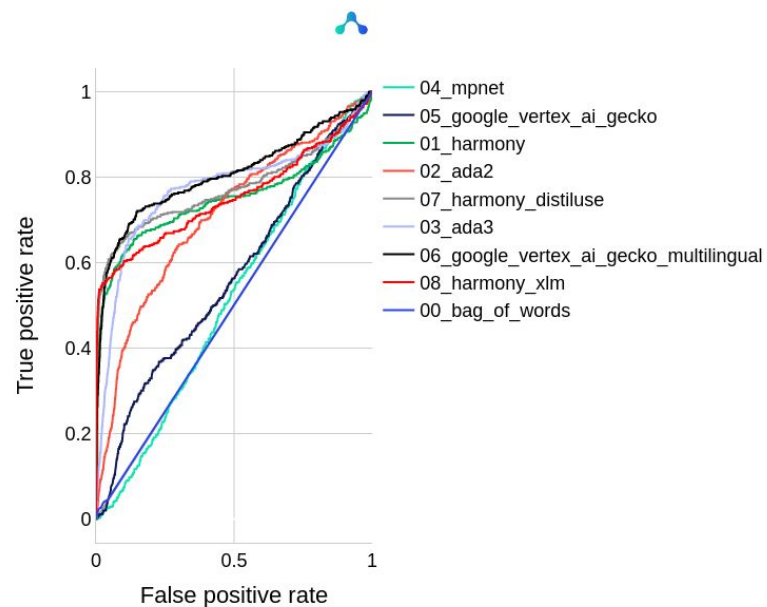
Vectors are orthogonal

Cosine score is low – closer to 0

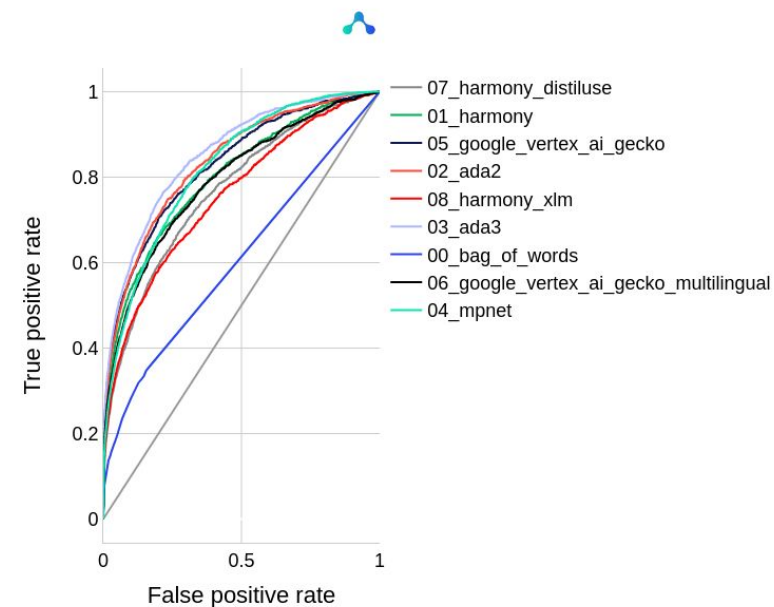
Evaluating Harmony



ROC on GAD 7 multilingual dataset



ROC on McElroy et al childhood dataset



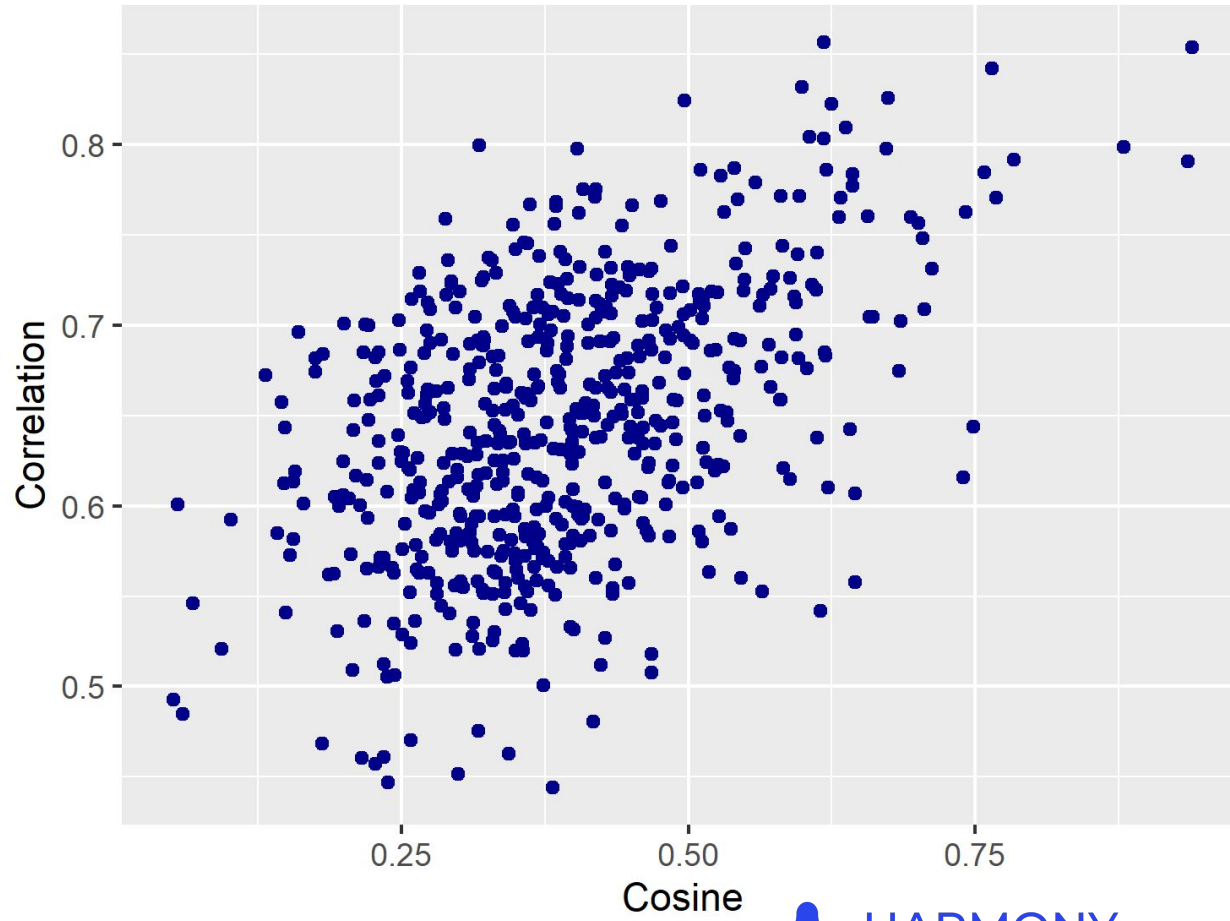


Real correlations?

	A	B	C				
1	Questionnaire	Item number	Content				
2	IDQ						
3	IDQ						
4	IDQ						
5	IDQ						
6	IDQ						
7	IDQ						
8	IDQ						
9	IDQ						
10	IDQ						
11	IAQ						
12	IAQ						
13	IAQ						
14	IAQ						
15	IAQ						
16	IAQ						
17	IAQ						
18	IAQ						
19	PHQ						

	A	B	C	D	E
	Supplementary File 2. Correlaiton and cosine coefficients for item pairs				
		from	to	spearman	cosine
1		1	10	0.719538559	0.61149627
2		2	11	0.719244021	0.445720732
3		3	12	0.731182941	0.711875081
4		4	13	0.665979411	0.571581244
5		5	14	0.703580795	0.511955619
6		6	15	0.709961188	0.297138691
7		7	16	0.691486691	0.50184983
8		8	17	0.608636306	0.259960353
9		9	18	0.738485659	0.37026453
10		10	19	0.798613012	0.879561961
11		11	2	0.824578169	0.496798843
12		12	1	0.573991369	0.572757006
13		13	21	0.628075543	0.655307412
14		14	22	0.625625193	0.354875147
15		15	23	0.718334673	0.681931853
16		16	24	0.683265023	0.334293664
17		17	25	0.587003714	0.53709048

Real correlations?



MCElroy et al,

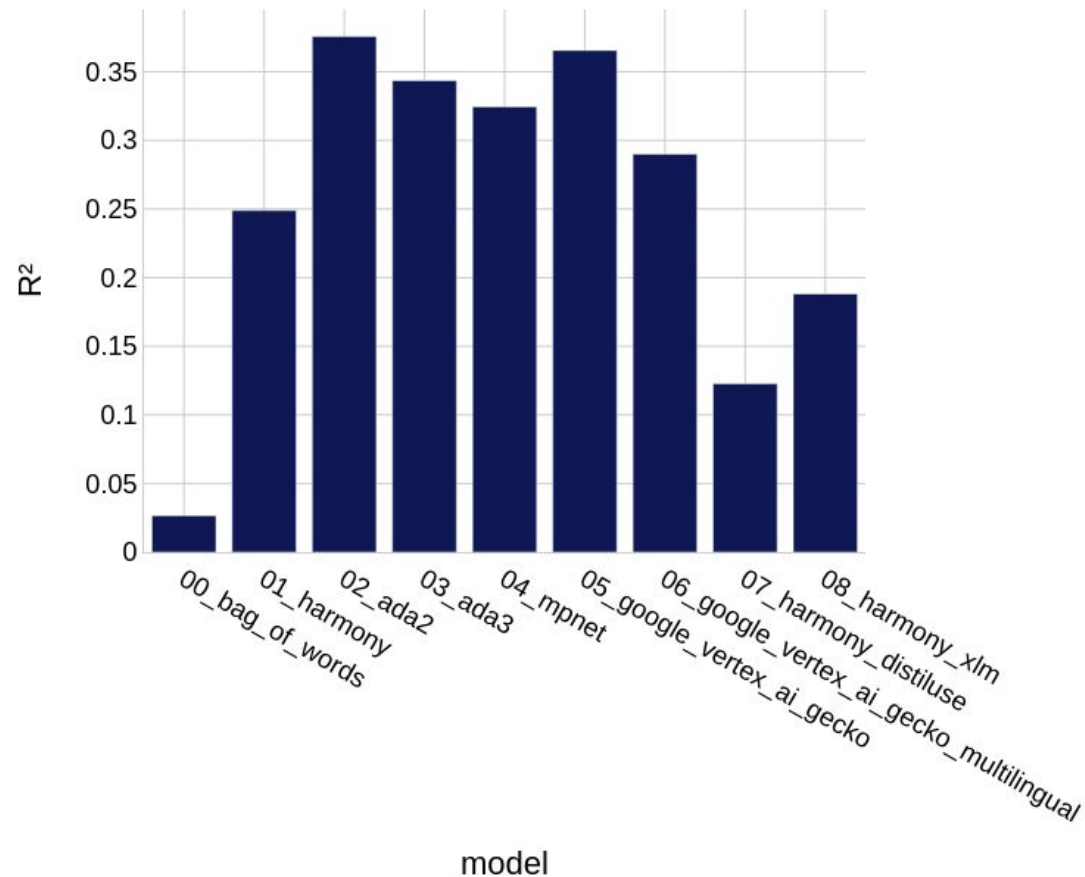
Using natural language processing to facilitate the harmonization of mental health questionnaires: a validation study using real-world data



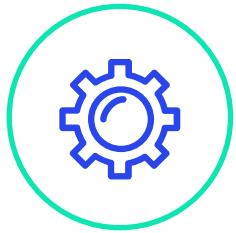


Real correlations?

R² on McElroy et al cosine correlation dataset



How to get involved



Clone/Fork the Github

- Make your changes in your fork
- Please run all the unit tests
- Check that the API and front end all run
- Make a pull request
- We merge your changes into main!



Use Harmony in research

- We'd love to see how Harmony can be used across social sciences, e.g. political sciences, law, market research, etc



Promote Harmony on social media

- Share on Instagram, LinkedIn, Twitter



Upcoming hackathon

Planned for 3 June in UCL, London
harmonydata.ac.uk/hackathon

Ideas so far:

1. Harmony needs improvement in parsing PDFs (check Kaggle)
2. Improving the matching algorithm
3. Add different LLMs
4. <https://harmonydata.ac.uk/ideas/>



Next steps

1. Extend Harmony for new LLMs
2. Improve the matching
3. Can we link Harmony to more psychology databases?
4. Are there use cases outside psychology?
 - Market research? (Surveys about new products?)
 - Pharma? (Informed Consent Forms?)
 - Finance?
 - Legal?
 - National Archives?
 - Your industry?
5. More collaborations...



HARMONY

Thank you!

Question, Suggestions and (maybe) answers

@harmony_data

@fastdatascienc1

<https://harmonydata.ac.uk/app/>



<https://discord.gg/harmonydata>

<https://github.com/harmonydata>



Fast Data Science



CENTRE FOR
LONGITUDINAL
STUDIES

