#### BEYOND SYNTAX: ON THE POWER AND POTENTIALITY OF DEEP OPEN SOURCE COMMUNITIES

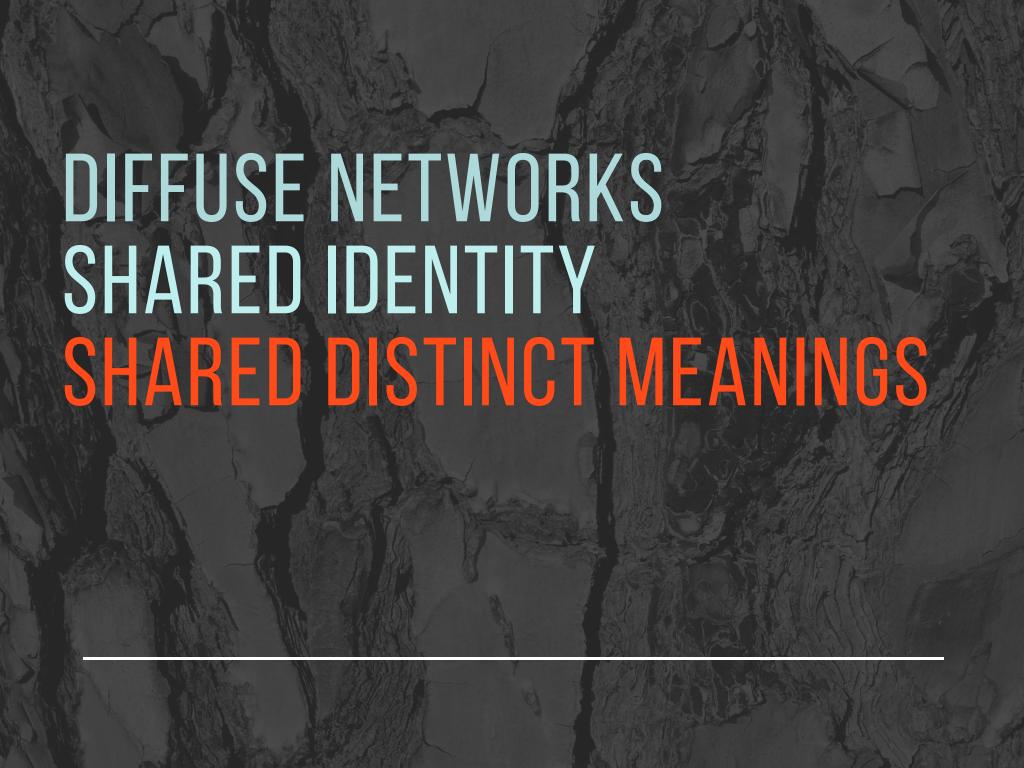
#### BEYOND SYNTAX?

#### NOT JUST A LANGUAGE?

"A RELATIVELY DIFFUSE SOCIAL NETWORK HAVING SHARED IDENTITY, DISTINCTIVE MEANINGS AROUND CERTAIN IDEAS, PRACTICE AND OBJECTS, AND A SENSE OF MARGINALIZATION FROM OR RESISTANCE TO THE PERCEIVED CONVENTIONAL SOCIETY..."

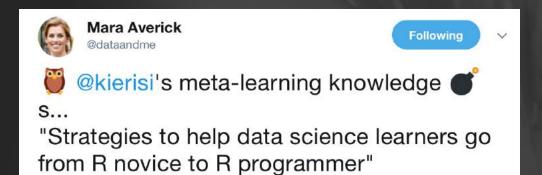


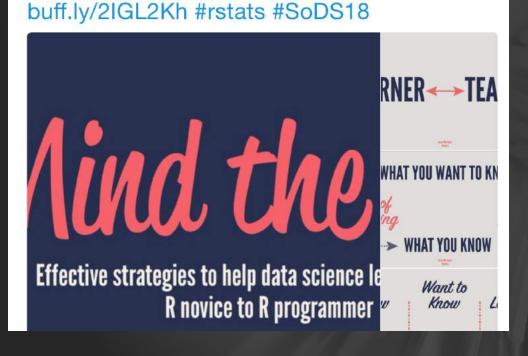




#### DIFFUSE NETWORKS SHARED IDENTITY SHARED DISTINCT MEANINGS RESISTANCE OR MARGINALISATION?

## SUBCULTURES HAVE THE POWER TO CREATE VALUE





#### OPEN SOURCE HAS VALUE

#### AN EVOLUTIONARY TOURNEY

#### SHARING CREATES VALUE

#### METHODOLOGIES

### METHODOLOGIES IDEOLOGIES

# METHODOLOGIES IDEOLOGIES STRATEGIES

# SUBCULTURES HAVE THE POTENTIAL TO CREATE VALUES

#### LOOKING FOR NUANCE

## THE POTENTIAL TO CREATE CHANGE

# ETHICS IN R

Both within and without

#### **Ethical Machine Learning**

#### A vignette designed to assist with spotting and preventing proxy bias.

Machine Learning systems often inherit biases against protected classes and historically disparaged groups via their training data [1]. Though some biases in features are straightforward to detect (ex: age, gender, race), others are not explicit and rely on subtle correlations in machine learning algorithms to understand. The incorporation of unintended bias into predictive models is called *proxy discrimination*.

Example of proxy bias in a bank's loan decision workflow.

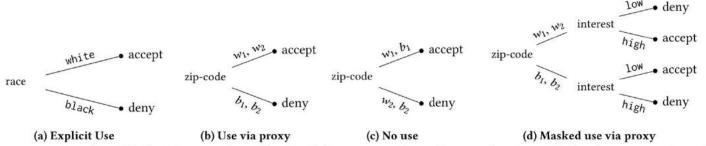
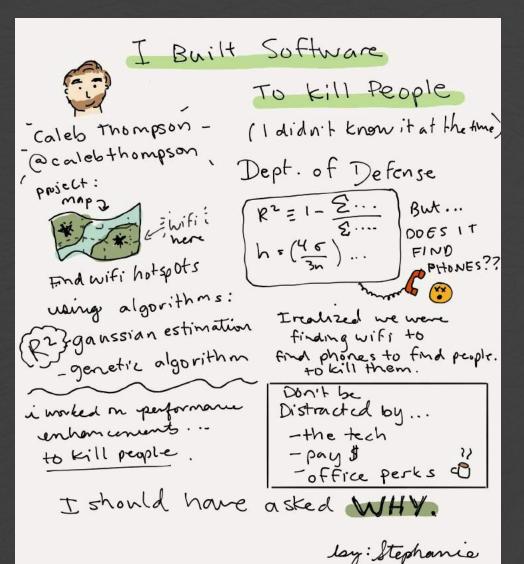


Figure 1: Examples of models (decision trees) used by a bank for accepting home loan applications. The bank uses race, zip-code, and customer's level of interest. Zip-codes  $w_1$  and  $w_2$  are predominantly white, whereas zip-codes  $b_1$  and  $b_2$  are predominantly black. Interest in loan, (high or low), is independent of race.



Thompson and Nemeth, 2018



## IF WE DON'T CHOOSE OUR VALUES, WHO WILL?

Not an absence of a presence of values, an absence of the presence of thought for them

# BEYOND SYNTAX: CREATE VALUE, EXPRESS YOUR VALUES

www.stephdesilva.com @stephdesilva

#### ACKNOWLEDGEMENTS: BETTER, TOGETHER

This piece benefited enormously from input, suggestions and challenges posed by Jesse Mostipak, Di Cook, Kim Fitter, John Ormerod and Charles Gray. Thank you all.

#### REFERENCES

Averick, M. (2018): Twitter Post,

<a href="https://twitter.com/dataandme/status/999834698997170178">https://twitter.com/dataandme/status/999834698997170178</a>, accessed 18/05/27.

Bailey, P. (2018): Ethical Machine Learning: A vignette designed to assist with spotting and preventing proxy bias, <a href="https://github.com/dynamicwebpaige/runconf18">https://github.com/dynamicwebpaige/runconf18</a>, accessed 18/05/23.

Coleman, E.G., (2012): Coding Freedom: The Ethics and Aesthetics of Hacking, Princeton University Press, Princeton.

Haenfler, R. (2014): Subcultures: The Basics, Oxon, Routledge.

Keyes, O. (2016): "Ethics, scientific consent and OKCupid", Ironholds,

<a href="https://ironholds.org/scientific-consent/">https://ironholds.org/scientific-consent/</a>, accessed 18/05/23.

Keyes, O. (2017): "Stop Mapping Names to Gender", Ironholds,

<a href="https://ironholds.org/names-gender/">https://ironholds.org/names-gender/</a>, accessed 18/05/23.

Keys, O. (2017b): "Behaviour and Community in Open Science", Ironholds,

<a href="https://ironholds.org/not-about-image/">https://ironholds.org/not-about-image/</a>, accessed 18/05/23.

#### REFERENCES

Leek, J., (2017): "Reproducibility and replicability is a glossy science now so watch out for the hype", *Simply Statistics*, <a href="https://simplystatistics.org/2017/03/02/rr-glossy/">https://simplystatistics.org/2017/03/02/rr-glossy/</a>, accessed 2018/05/23.

Leek, J., (2018): "The value of re-analysis", Simply Statistics,

<a href="https://simplystatistics.org/2012/12/18/the-value-of-re-analysis/">https://simplystatistics.org/2012/12/18/the-value-of-re-analysis/</a>, accessed 2018/05/23.

Leek, J., Peng, R.D., (2015): "Opinion: Reproducible research can still be wrong: Adopting a prevention approach", *Proceedings of the National Academy of Sciences of the United States of America*\_ Volume 112, Number 6, p. 1645-1646,

<a href="http://www.pnas.org/content/112/6/1645">http://www.pnas.org/content/112/6/1645</a>, accessed 18/05/23.

Thompson, C., Nemeth, S. (2018): I Built Software to Kill People,

<a href="https://twitter.com/stephaniecodes/status/992419444810010625">https://twitter.com/stephaniecodes/status/992419444810010625</a>, accessed 18/05/23.