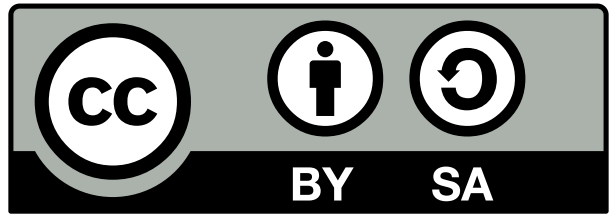
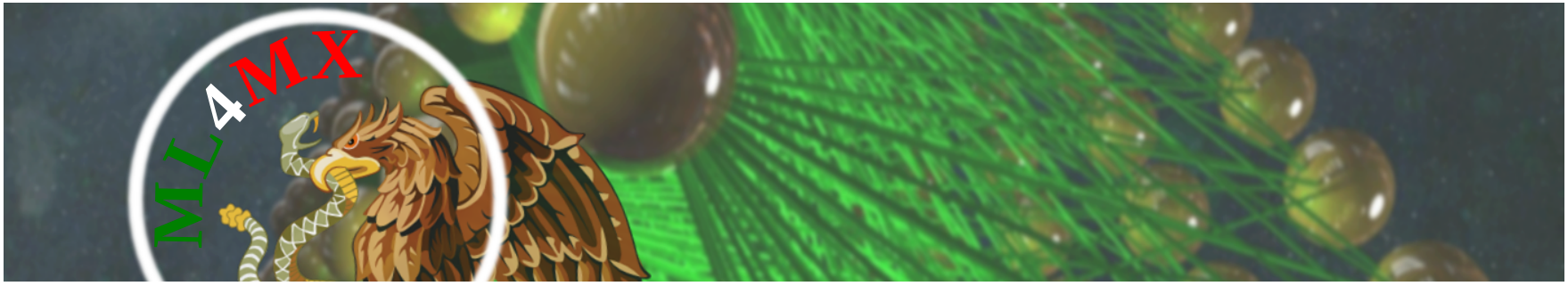


This presentation is released under the terms of the
Creative Commons Attribution-Share Alike license.

You are free to reuse it and modify it as much as you want as long as
(1) you mention Miguel P Xochicale as being the original author,
(2) you re-share your presentation under the same terms.

You can download the sources of this presentation here:
https://github.com/mxochicale/chamaleon-mx



Challenges in Reproducibility of Machine Learning, Open Source Tools and Mexican Open Datasets

@ML4MX #ML4MX

Miguel P Xochicale

@_mxochicale @mxochicale

https://github.com/mxochicale/chamaleon-mx

REPRODUCIBILITY IN ML & AI

REPRODUCIBILITY PROBLEM IN AI

"In Artificial Intelligence, only 6% of 400 included code."

[Gundersen O. E., **Missing data hinder replication of AI studies**, Feb 2018]

REPRODUCIBILITY PROBLEM IN AI

"Most Artificial Intelligence researchers don't report the source code, or the data that AI are trained on."

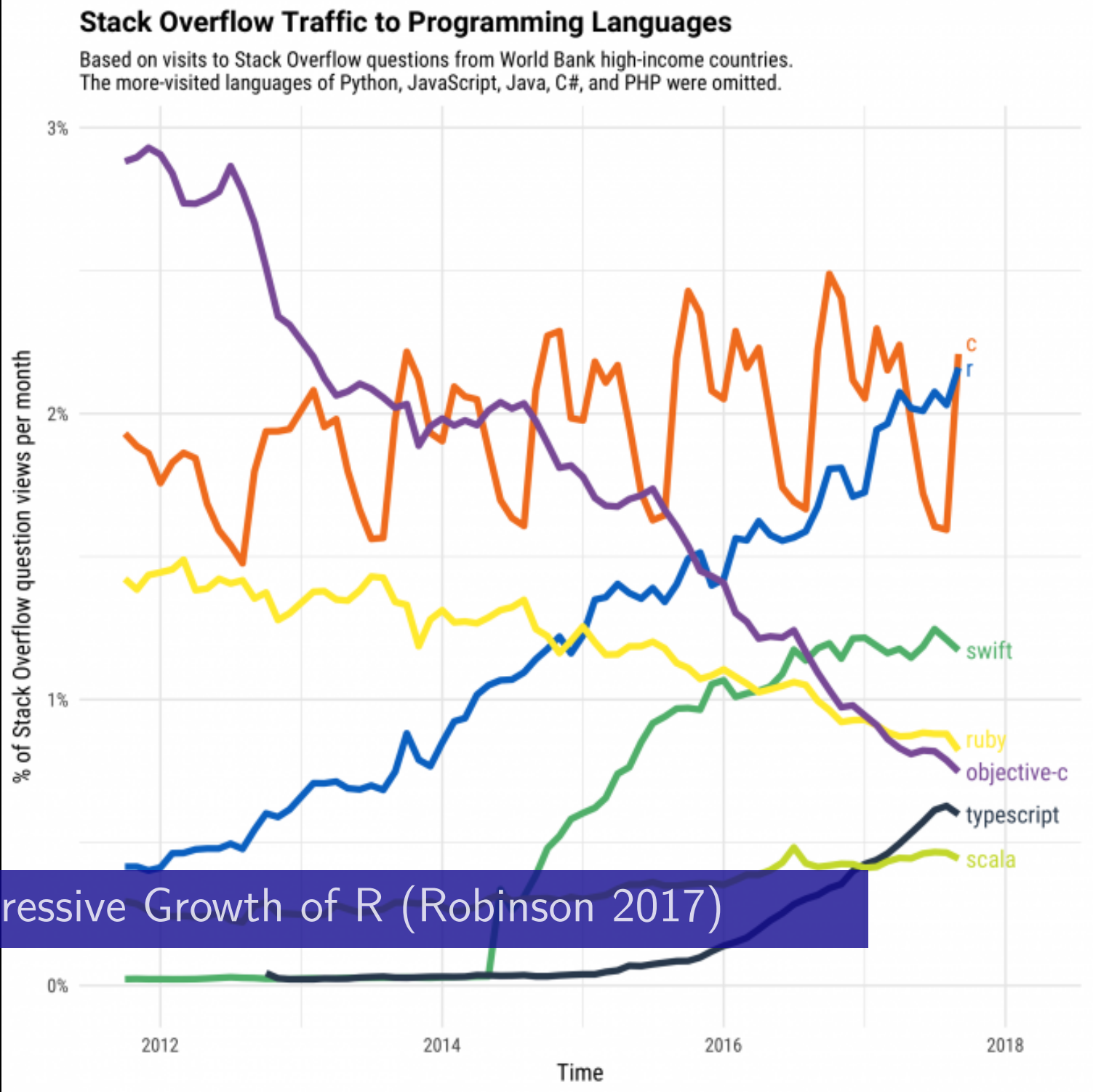
[Geib C., **Scientists Can't Replicate AI Studies. That's Bad News.**, Feb 2018]

REPRODUCIBILITY CRISIS IN ML

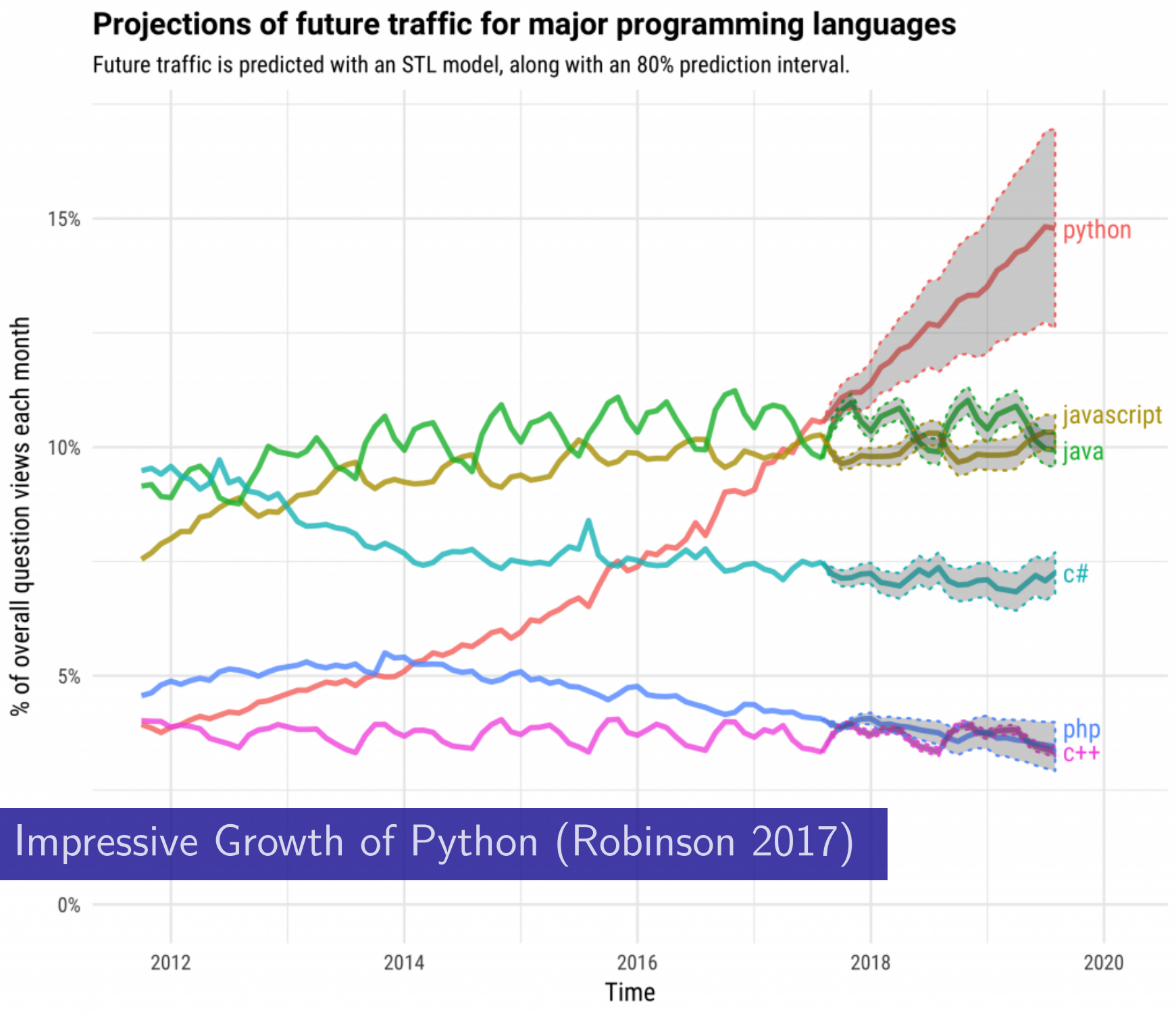
"Machine Learning is by far the worst environment I've ever found for collaborating and keeping track of changes."

[Warden P., **The Machine Learning Reproducibility Crisis.**, March 2018]

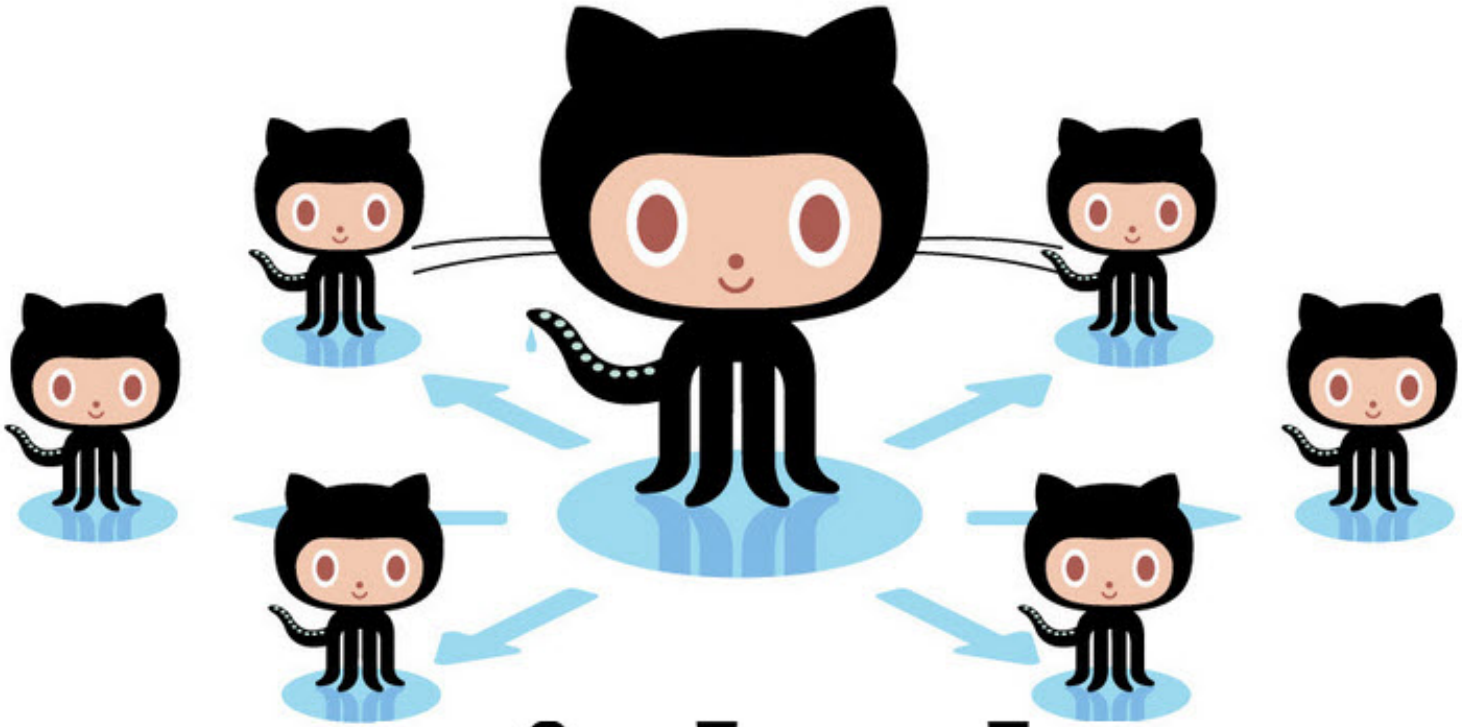
OPEN SOURCE SOFTWARE TOOLS



The Impressive Growth of R (Robinson 2017)



The Impressive Growth of Python (Robinson 2017)



Learn to use GitHub (https://github.com/)

MEXICAN OPEN DATASETS

MEXICAN OPEN DATASETS?

* Portal de Datos Abiertos del Gobierno de la República [https://datos.gob.mx/]

  - Calidad de Aire de SINAICA [Formato json]
  - Condiciones Atmosféricas [Formato json]
  - Gasolineras y Precios Comerciales de Gasolina y Diesel [Formato xml]
  - Contrataciones Abiertas [Formato json]

* Comunidad de práctica de datos abiertos en Mexico [twitter @OpenDataMX]

* Construimos herramientas digitales para fortalecer a la ciudadanía. [twitter @CivicaDigical]

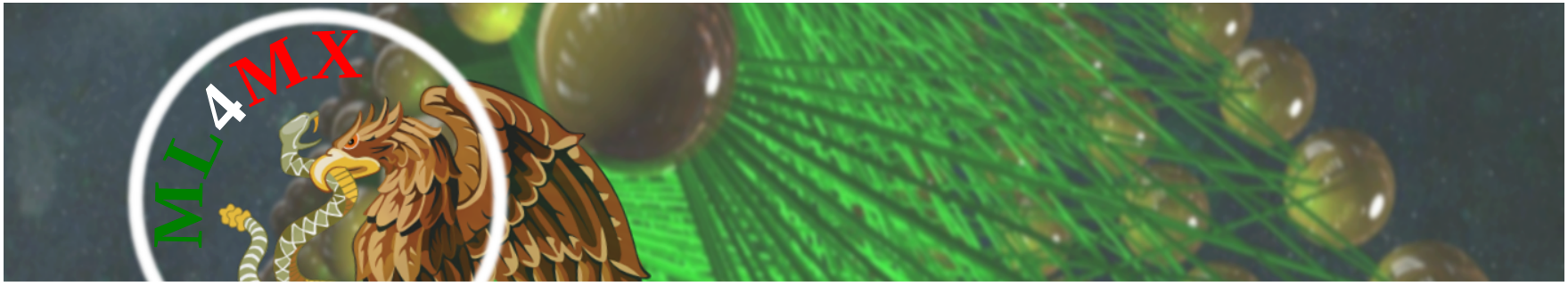
@DatosGobMX so the questions are: 1, who in MEX is doing #MachineLearning with #OpenData of @DatosGobMX ? 2. Can u share the contact of those persons?
12:11 PM - 20 Feb 2018
1 Retweet 2 Likes
1
2
Tweet your reply
STILL WAITING FOR REPLIES!

PREGUNTAS?

TAKE AWAY MESSAGES

- chose LIBRE over close, SIMPLE over easy
- learn / teach / build / share
- give your stuff away
- LEARN FROM:
  https://github.com/ml4mx
  https://rescience.github.io

[Brett M, **https://github.com/matthew-brett/thw-intro.**, March 2018]



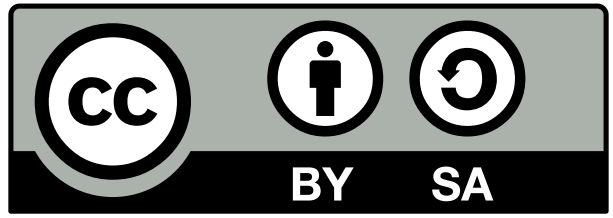
Challenges in Reproducibility of Machine Learning, Open Source Tools and Mexican Open Datasets

@ML4MX #ML4MX

Miguel P Xochicale

@_mxochicale @mxochicale

https://github.com/mxochicale/chamaleon-mx



This presentation is released under the terms of the
Creative Commons Attribution-Share Alike license.

You are free to reuse it and modify it as much as you want as long as
(1) you mention Miguel P Xochicale as being the original author,
(2) you re-share your presentation under the same terms.

You can download the sources of this presentation here:
https://github.com/mxochicale/chamaleon-mx