

# Aula 9 – O Estado-da-arte e o Futuro

**Luíz Fernando Esser**

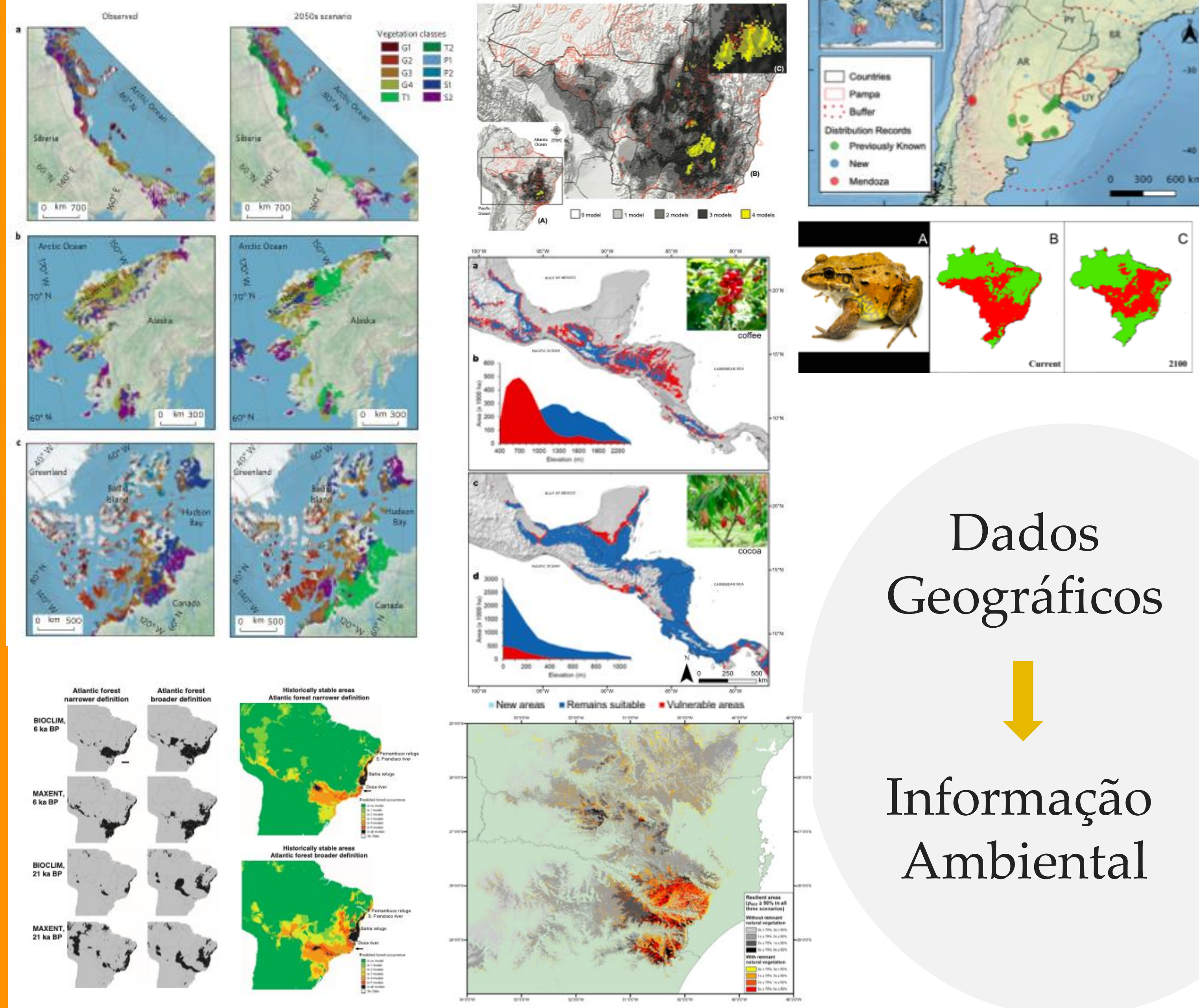
**Fundamentos de Modelagem de Distribuição de Espécies no R**



# Na última aula...

- ✓ Atributos Funcionais
- ✓ Padrões Biogeográficos
- ✓ Espécies invasoras
- ✓ Distinguir Espécies
- ✓ Populações
- ✓ Microrefúgios
- ✓ Espécies cultivadas
- ✓ Sistemas de cultivo

Garbage in, garbage out.



Dados Geográficos



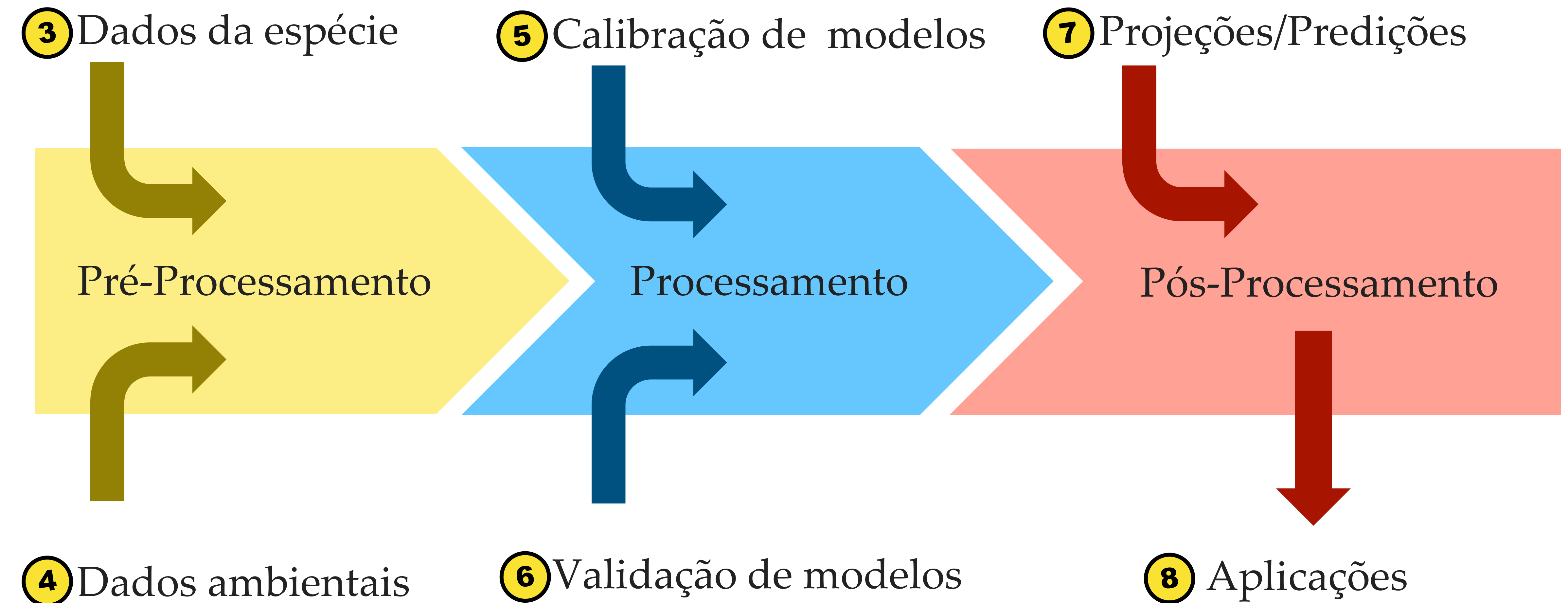
Informação Ambiental



Coisas que fazem um bom trabalho

(tudo aqui é justificável)

# Framework - SDM



# O que é preciso cuidar...

## ③ Dados da espécie

- Presenças foram coletadas de mais de uma base de dados?
- Estão sendo usados dados de toda a região de distribuição?  
(i.e. existe suficiência amostral?)
- Data cleaning excluindo dados dentro da mesma célula do raster?
- Quantas presenças tem? Aonde elas estão (mapa)?
- Diferença entre background e pseudoausência.
- Quantidade adequada de dados de BG ou PA.
- Como dados de BG ou PA foram obtidos? Faz sentido com o conceito?
- Quantos datasets de BG/PA foram gerados (evitando incerteza)?



# O que é preciso cuidar...

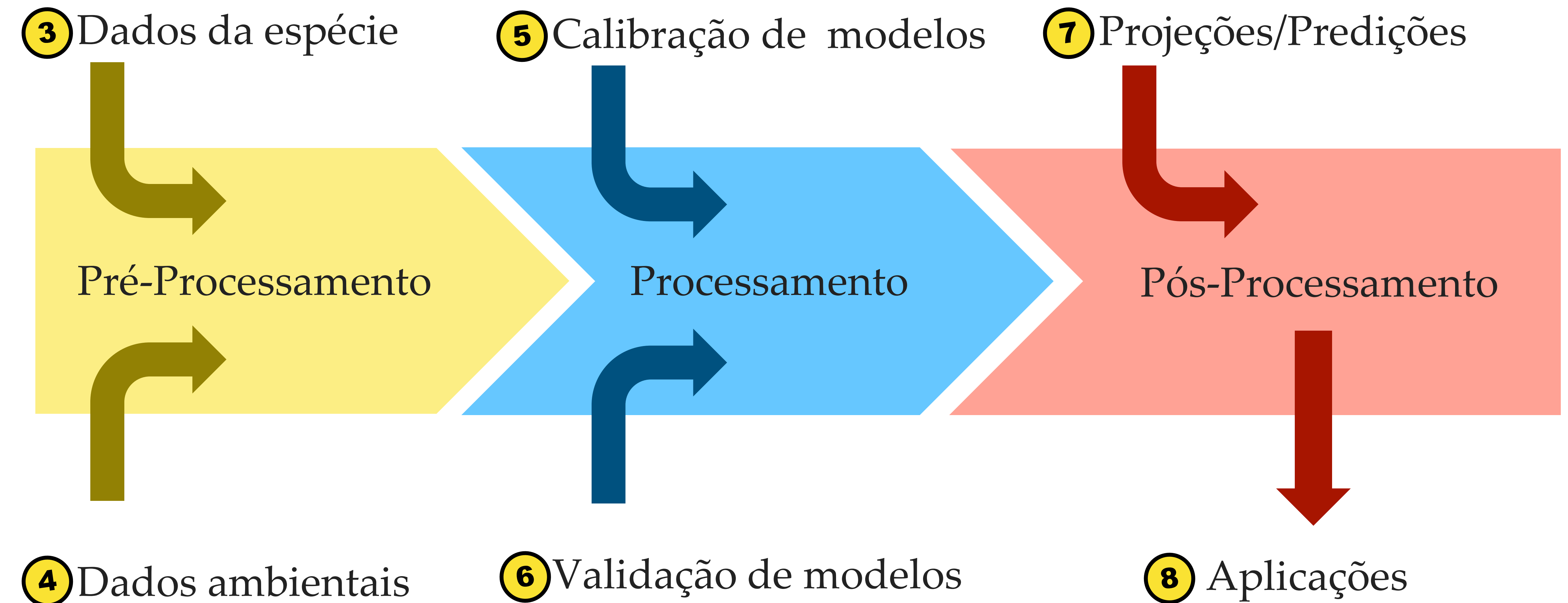
## ③ Dados da espécie



#...i think this meme demonstrated the importance of sample size better and more efficiently than any science class i've ever taken



# Framework - SDM





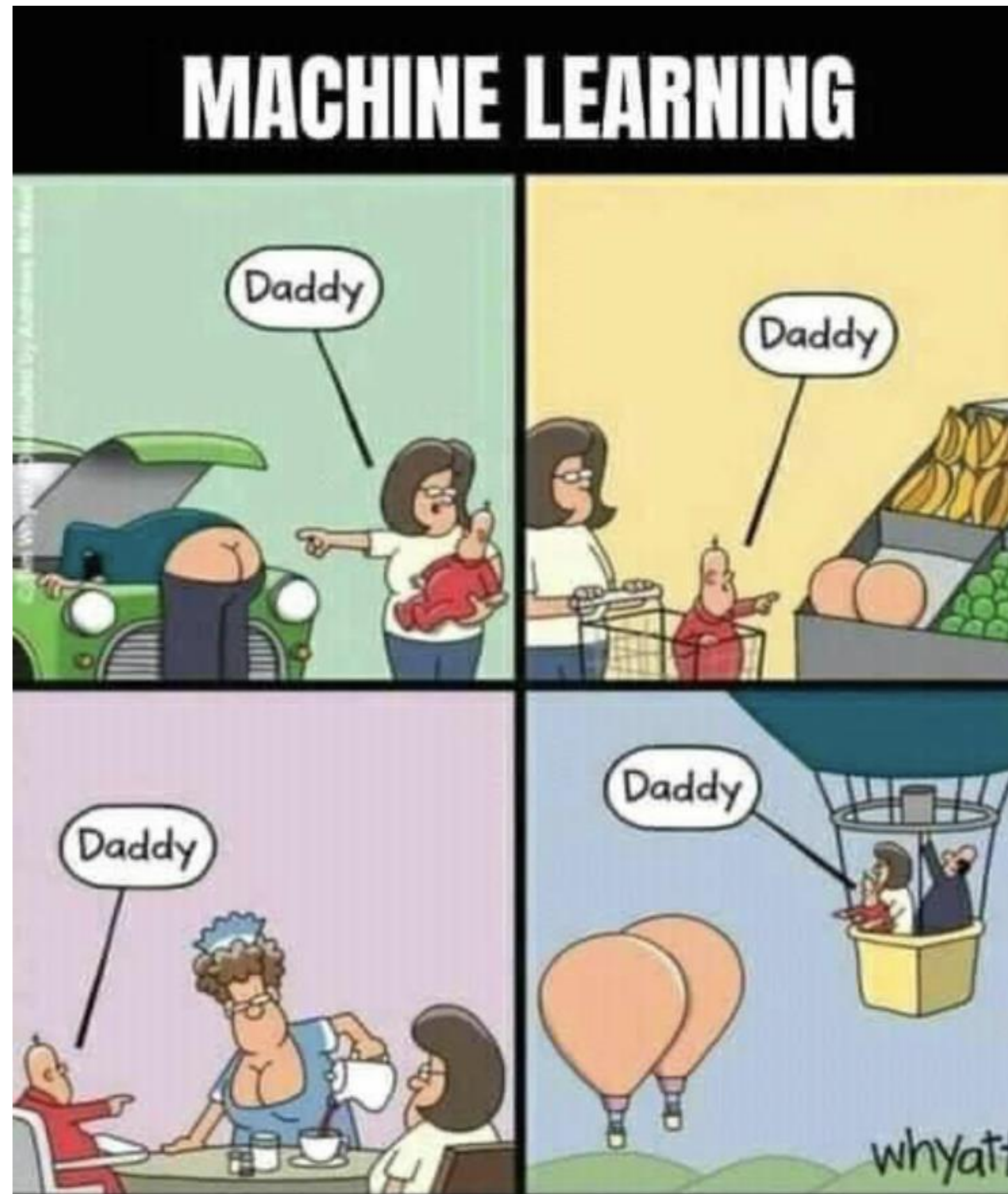
# O que é preciso cuidar...

## ④ Dados ambientais

- Foram consideradas apenas variáveis com relação causal?
- Foi removida a multicolinearidade?
- Qual a área de distribuição das variáveis?
- Proporção de variáveis por número de presenças.

# O que é preciso cuidar...

## ④ Dados ambientais



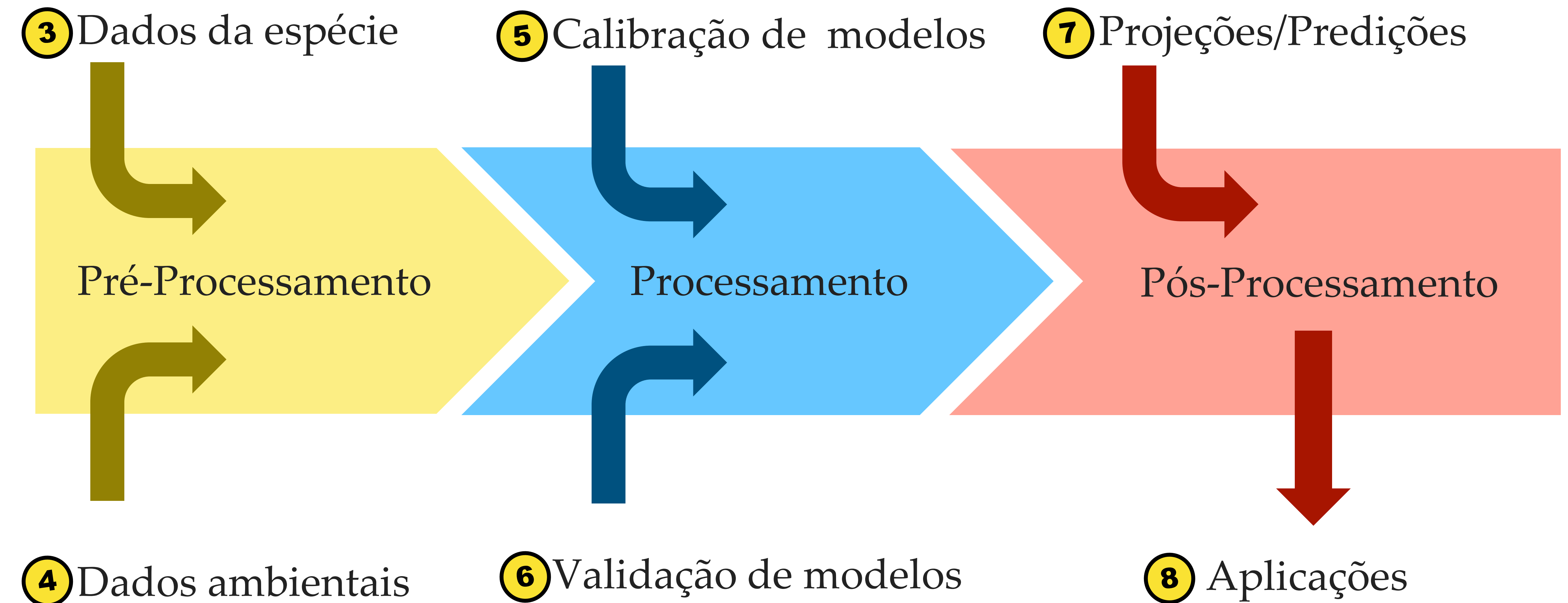
Dados  
Geográficos



Informação  
Ambiental



# Framework - SDM



# O que é preciso cuidar...

## ⑤ Calibração de modelos

- Os algoritmos passaram por "tuning" ou foram usados no padrão?
- Os dados da espécie suportam o algoritmo usado (i.e. existe confusão entre pseudoausência e background)?
- Curvas de resposta condizem com uma resposta biológica?
- Quantos modelos foram gerados no total?



# O que é preciso cuidar...

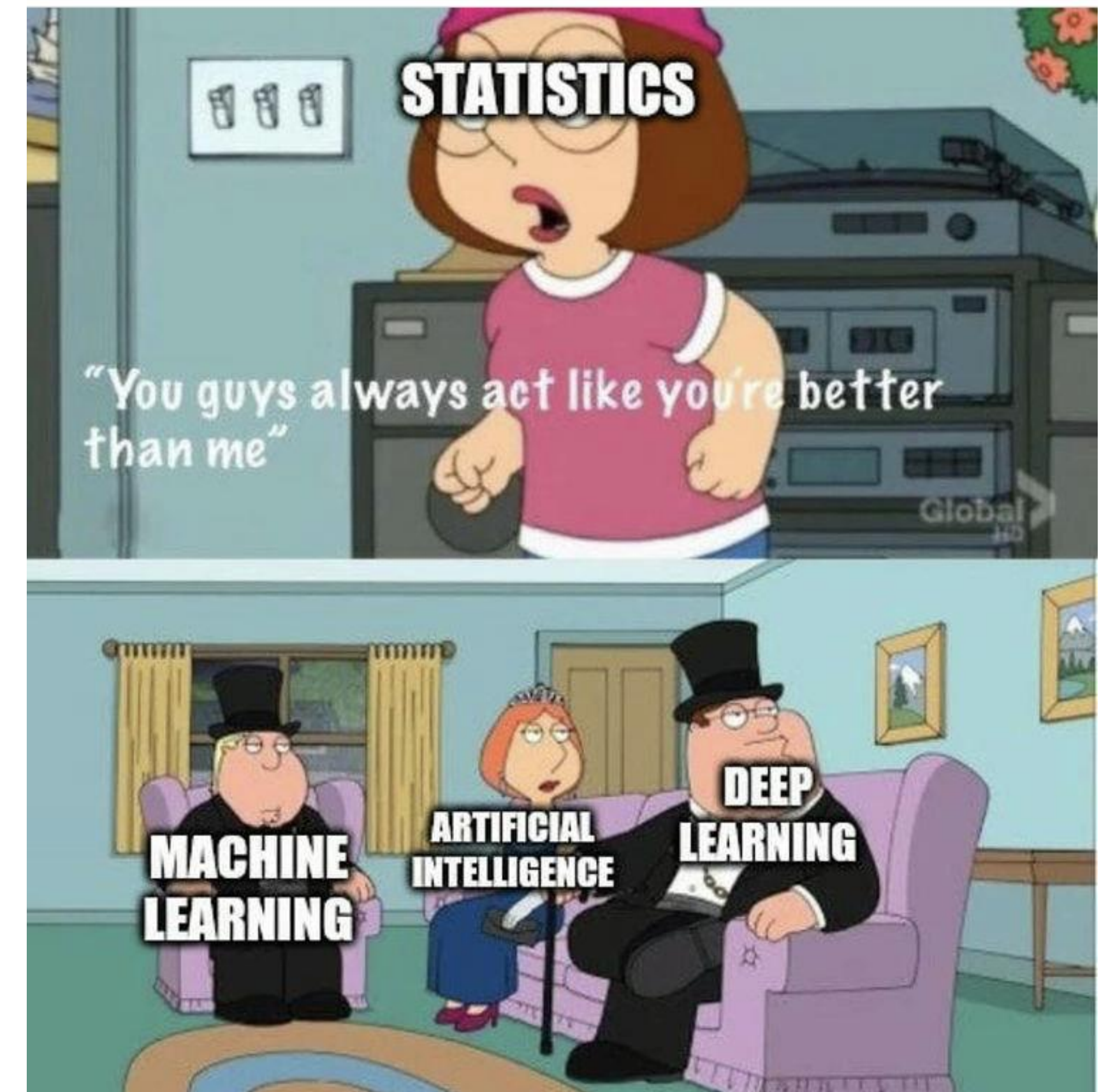
## ⑤ Calibração de modelos

	A parrot	Machine learning algorithm
Learns random phrases	✓	✓
Doesn't understand shit about what it learns	✓	✓
Occasionally speaks nonsense	✓	✓
Is a cute birdie parrot	✓	✗

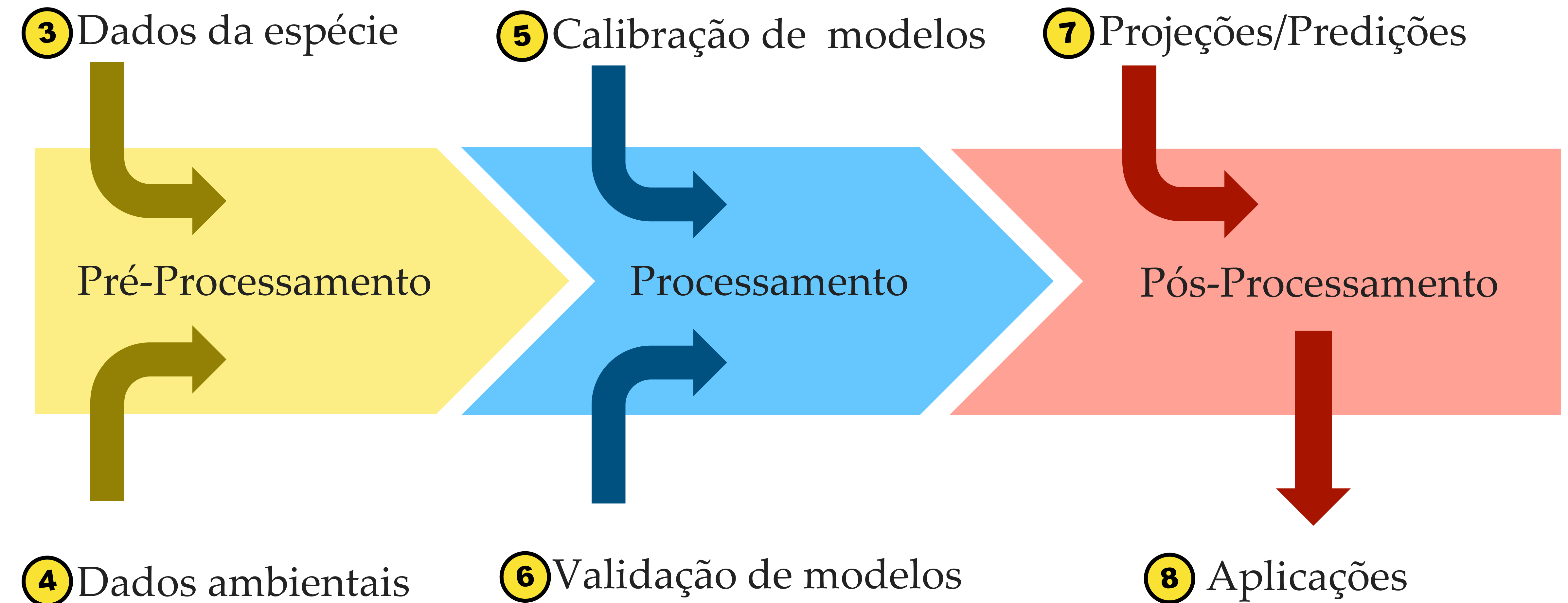


thenullhypotenuse  
Mars

...



# Framework - SDM





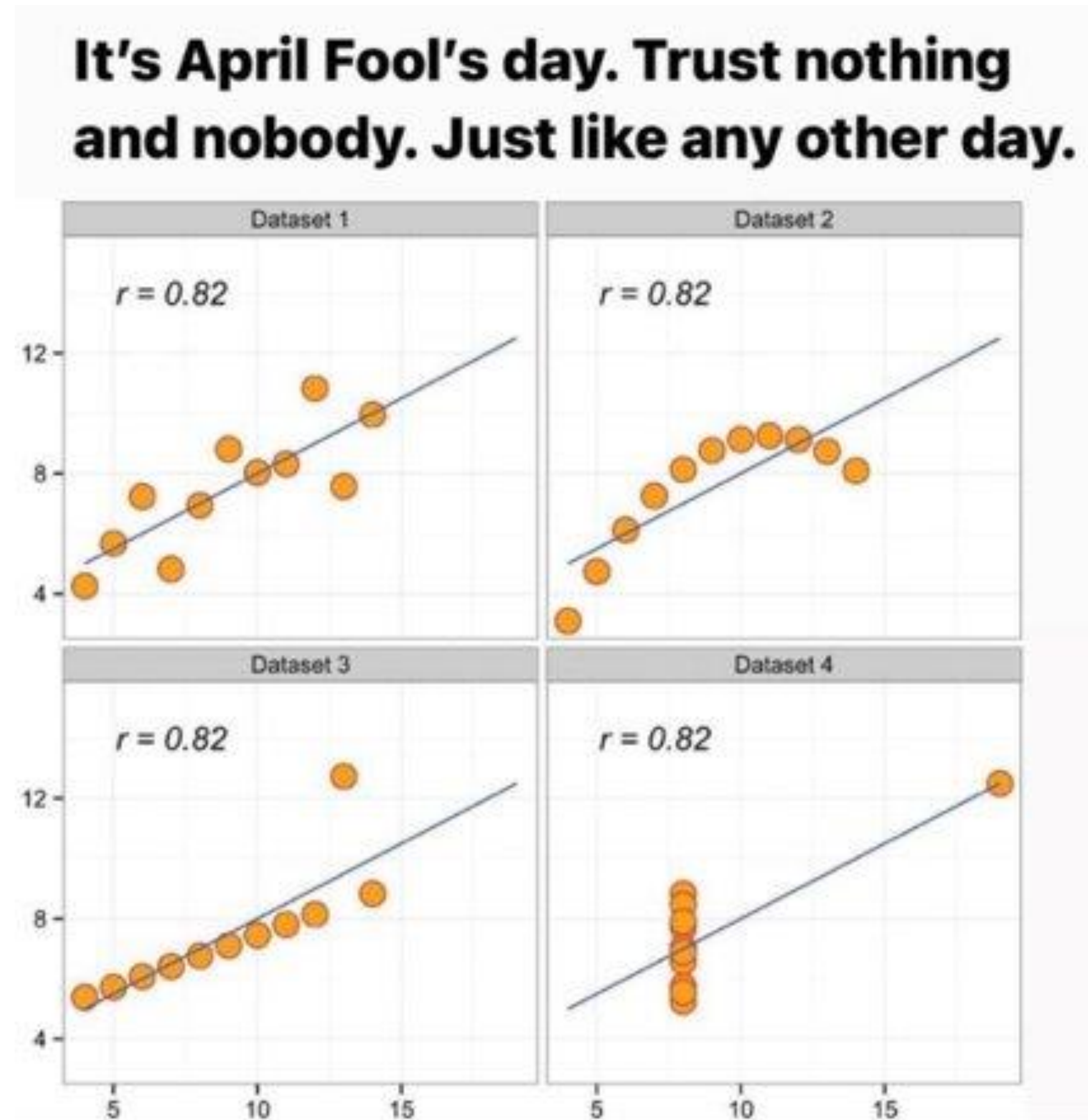
# O que é preciso cuidar...

## ⑥ Validação de modelos

- Os dados de teste estão separados dos dados de treino?
- Qual a proporção de dados de treino e de teste?
- A métrica utilizada responde à pergunta?
- Os limiares utilizados são subjetivos?
- Está ocorrendo overfitting?
- Quais os valores das métricas utilizadas?

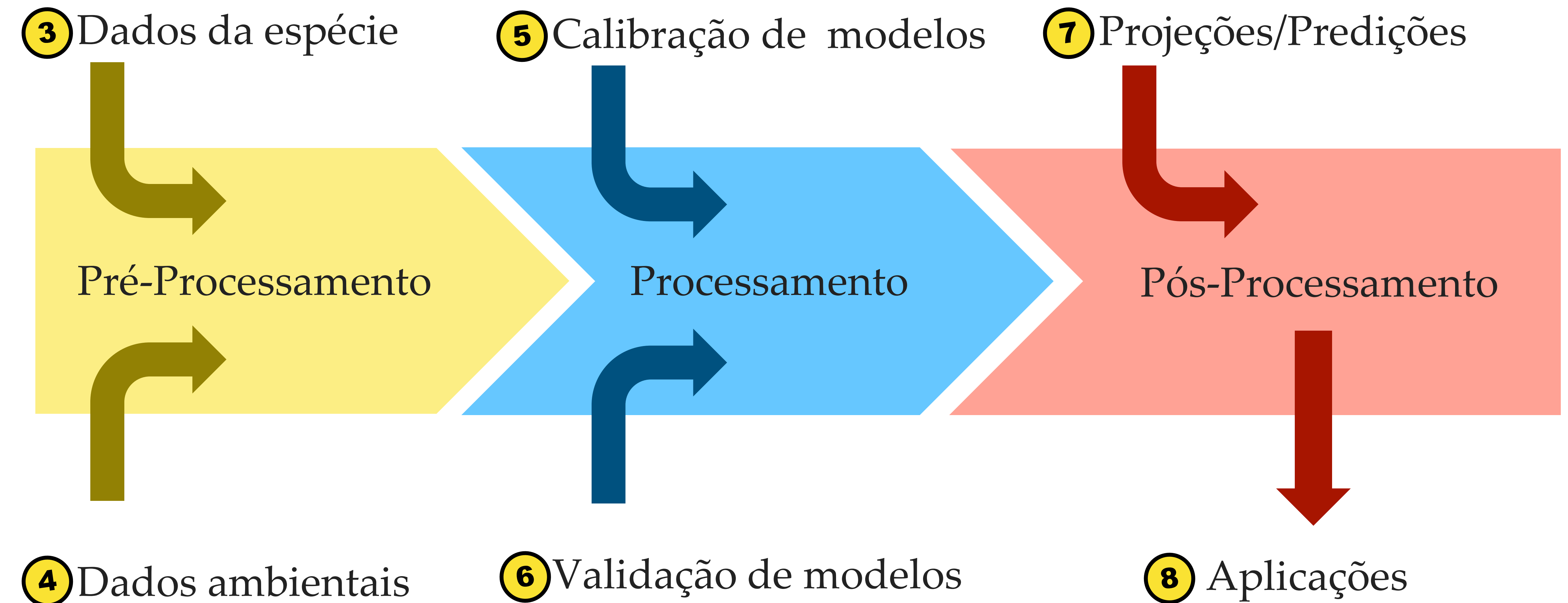
# O que é preciso cuidar...

## ⑥ Validação de modelos





# Framework - SDM



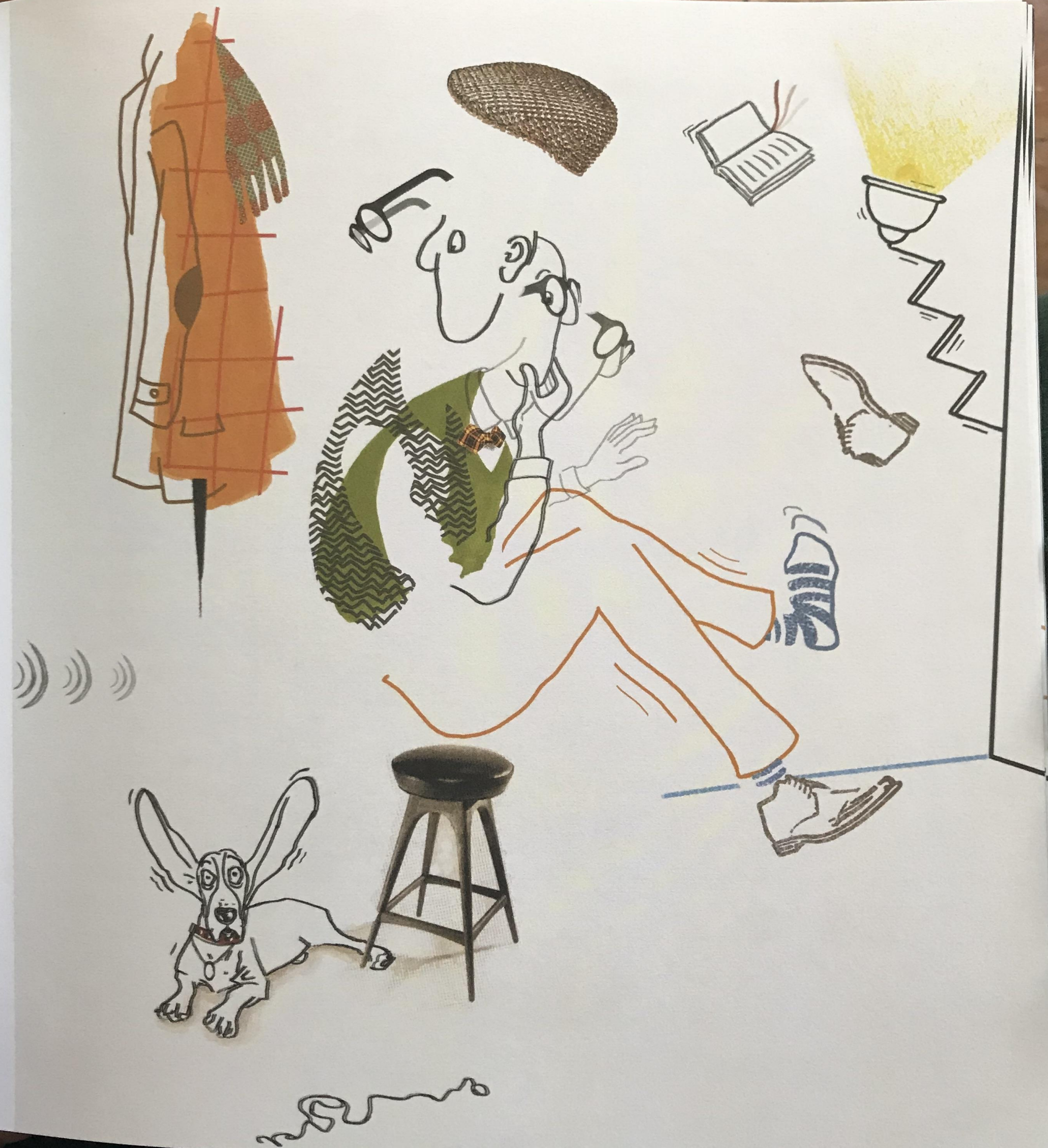
# O que é preciso cuidar...

## 7 Projeções/Predições

- Os cenários projetados respondem às perguntas?
- Qual a quantidade de GCMs?
- Como foi feito o ensemble?
- O que foi retornado do ensemble (adequabilidade ambiental ou probabilidade de presença)?
- Está sendo apresentado um mapa de incerteza?



# O que é preciso cuidar...





Ferramentas já utilizáveis



# Possibilidades

- Ensemble of Small Models (ESM)
- Equivalência e Similaridade de Nicho
- Presence-Absence Matrices (PAMs)
- Joint Species Distribution Models (JSDMs)

## Methods in Ecology and Evolution



Methods in Ecology and Evolution 2015, 6, 1210–1218

doi: 10.1111/2041-210X.12403

### Overcoming limitations of modelling rare species by using ensembles of small models

Frank T. Breiner<sup>1,2\*</sup>, Antoine Guisan<sup>2,3</sup>, Ariel Bergamini<sup>1†</sup> and Michael P. Nobis<sup>1†</sup>

<sup>1</sup>Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland; <sup>2</sup>Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; and <sup>3</sup>Institute of Earth Surface Dynamics, University of Lausanne, 1015 Lausanne, Switzerland

#### ORIGINAL ARTICLE

doi:10.1111/j.1558-5646.2008.00482.x

### ENVIRONMENTAL NICHE EQUIVALENCY VERSUS CONSERVATISM: QUANTITATIVE APPROACHES TO NICHE EVOLUTION

Dan L. Warren<sup>1,2</sup>, Richard E. Glor<sup>3,4</sup> and Michael Turelli<sup>1,5</sup>

<sup>1</sup>Department of Evolution and Ecology, University of California, Davis, California 95616

<sup>2</sup>E-mail: danwarren@ucdavis.edu

<sup>3</sup>Department of Biology, University of Rochester, Rochester, New York, 14627

<sup>4</sup>E-mail: rglor@mail.rochester.edu

<sup>5</sup>E-mail: mturelli@ucdavis.edu

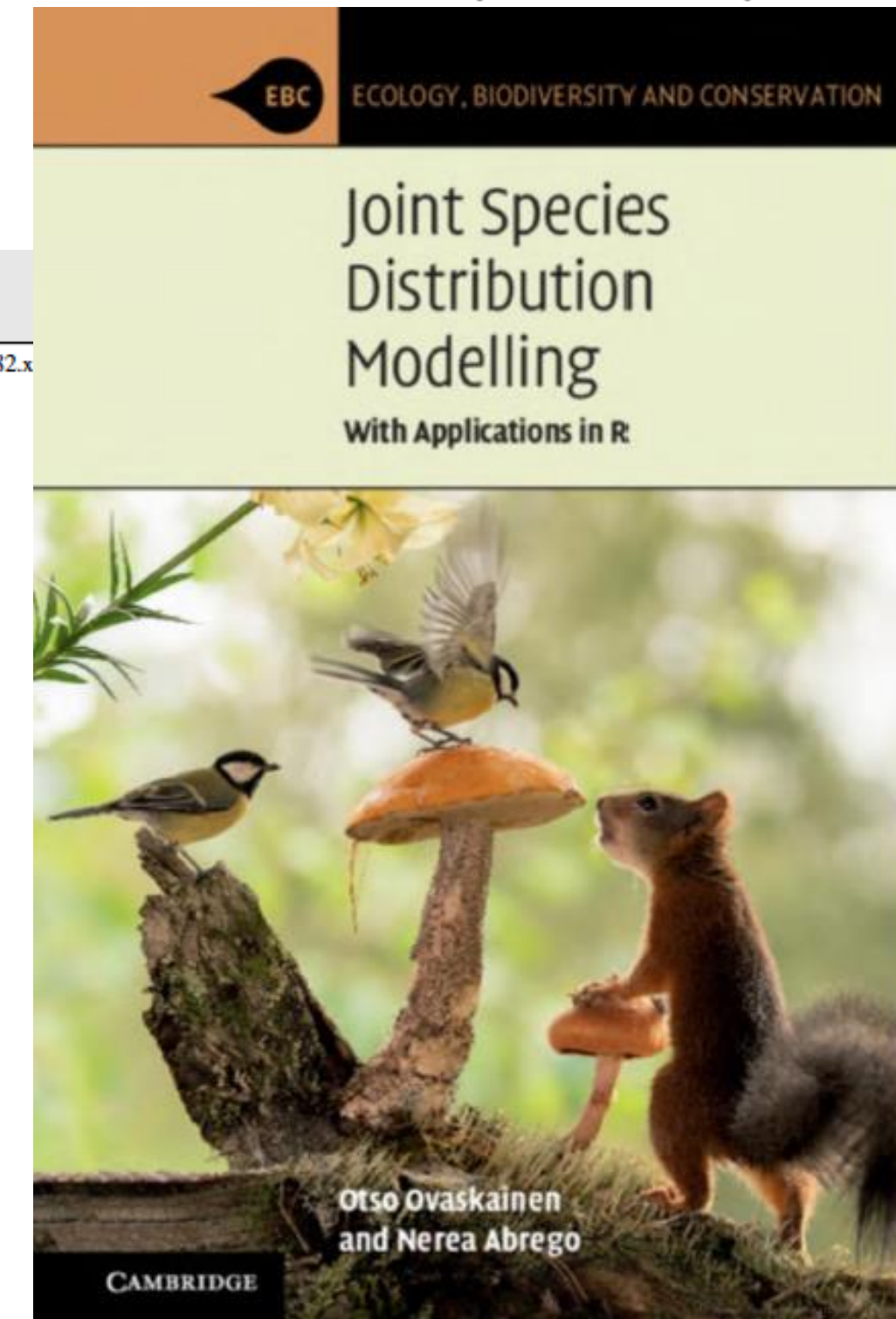
Biodiversity Informatics, 10, 2015, pp. 22–34

### INDICES OF BIODIVERSITY PATTERN BASED ON PRESENCE-ABSENCE MATRICES: A GIS IMPLEMENTATION

JORGE SOBERÓN<sup>1,2</sup> AND JEFF CAVNER<sup>1</sup>

<sup>1</sup>Biodiversity Institute, University of Kansas, Lawrence, KS 66045, USA

<sup>2</sup>Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA





# Possibilidades



**Adam Charles**

@BaselessPursuit



I was surprised when my daughter said she learned R at school yesterday, and then I remembered that she's 4 and she meant the letter.

My priors are all too skewed

8:49 PM · 17 Mar 21 · [Twitter Web App](#)

**2,261** Retweets **296** Quote Tweets





# Possibilidades





Coisas que podem mudar...



# Coisas que podem mudar...

## Positive-Unlabeled Learning

- Um único algoritmo implementado em R.

## Teoria de redes



## Package ‘AdaSampling’

May 21, 2019

**Type** Package

**Title** Adaptive Sampling for Positive Unlabeled and Label Noise Learning

**Version** 1.3

**Author** Pengyi Yang

**Maintainer** Pengyi Yang <yangpy7@gmail.com>

**Description** Implements the adaptive sampling procedure, a framework for both positive unlabeled learning and learning with class label noise. Yang, P., Ormerod, J., Liu, W., Ma, C., Zomaya, A., Yang, J. (2018) <doi:10.1109/TCYB.2018.2816984>.

**License** GPL-3

**Encoding** UTF-8

**Depends** R (>= 3.4.0)

**LazyData** true

**Imports** caret (>= 6.0-78) , class (>= 7.3-14), e1071 (>= 1.6-8), stats, MASS

**BugReports** <https://github.com/PengyiYang/AdaSampling/issues>

**Suggests** knitr, rmarkdown

**VignetteBuilder** knitr

**URL** <https://github.com/PengyiYang/AdaSampling/>

**RoxygenNote** 6.1.1



# Four stages of learning Data Science



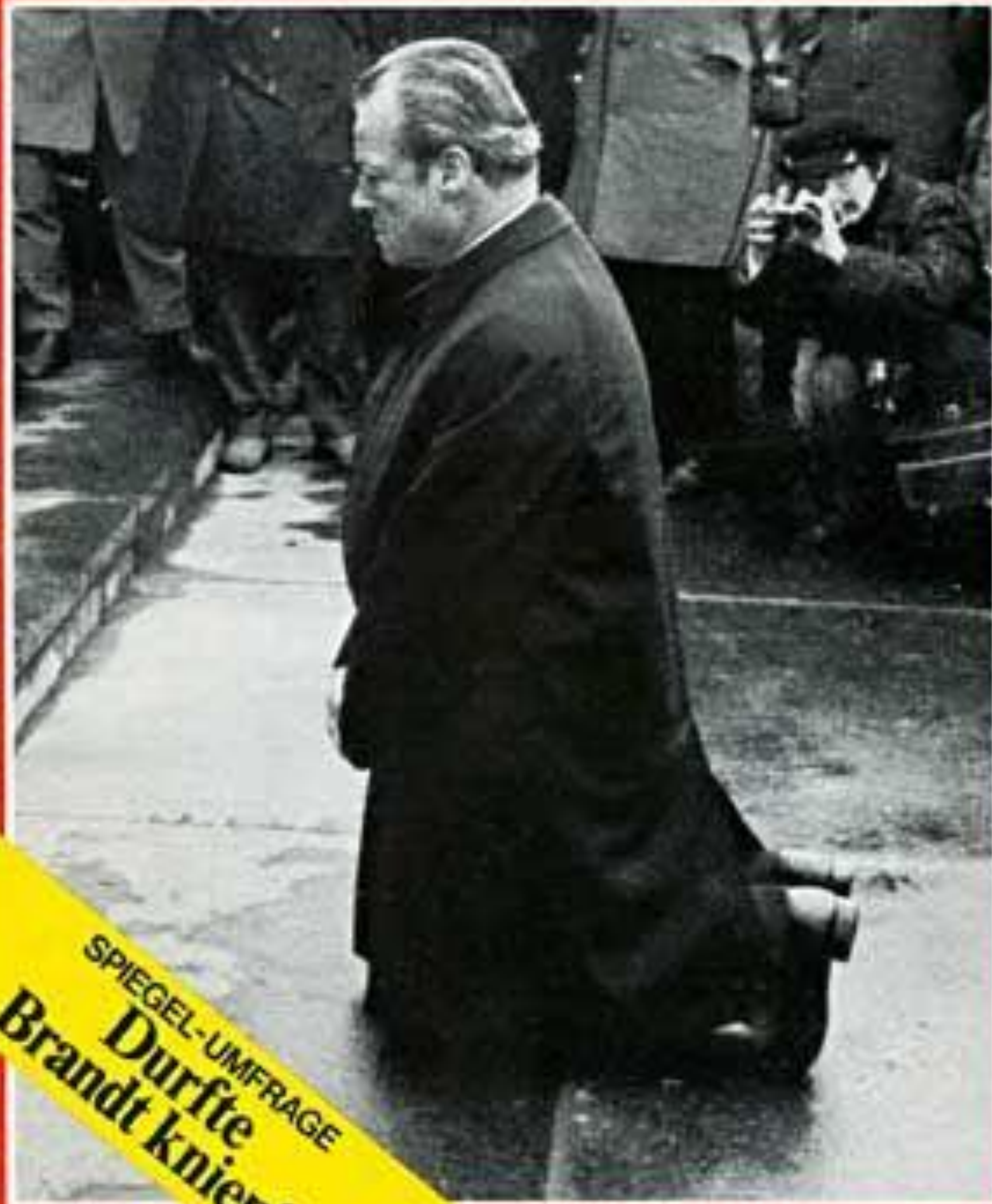


Mas tudo pode mudar...  
Tudo deve mudar...



# DER SPIEGEL

14. DEZEMBER 1970, NR. 50  
IN: JAHRSBAND - 300 L. 200  
HNO. HANDELSSTRECKEN  
C 4390 C



SPIEGEL-UMFRAGE  
**Durfte  
Brandt knien?**

Willy Brandt ajoelhado no gueto de Varsóvia