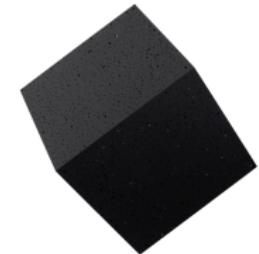


# **Edward: A library for probabilistic modeling, inference, and criticism**

Dustin Tran, David M. Blei  
Columbia University

Matt Hoffman, Rif A. Saurous, Eugene Brevdo,  
Kevin Murphy  
Google Brain

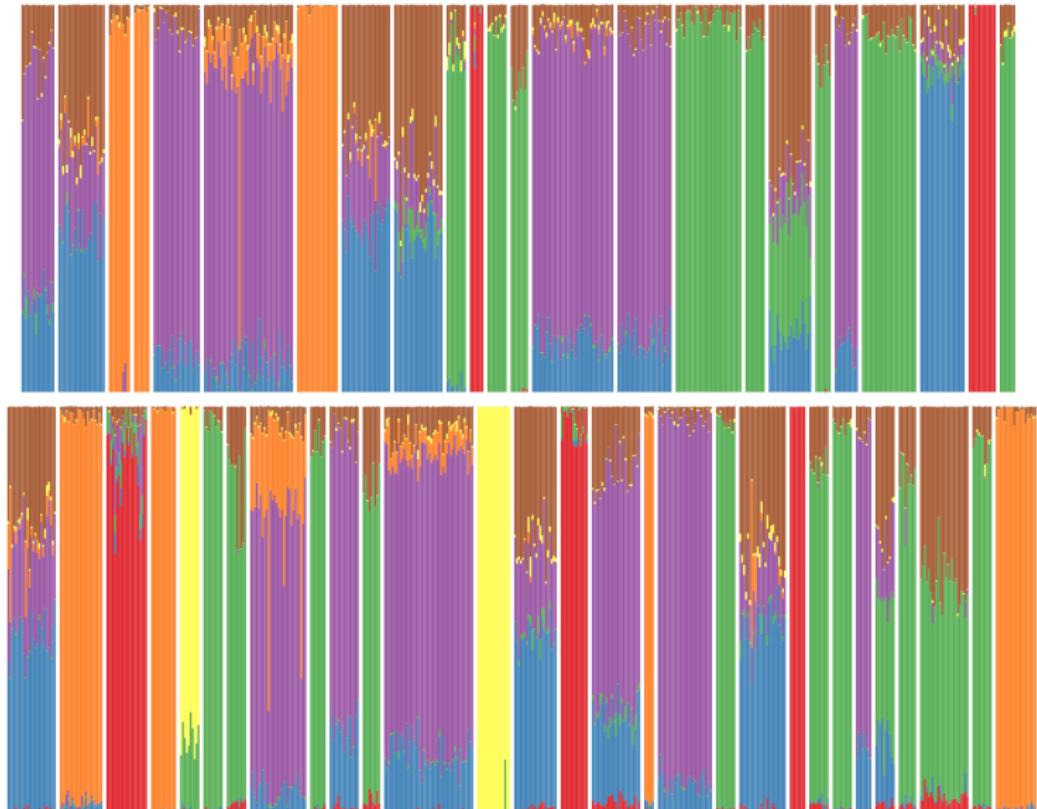


[edwardlib.org](http://edwardlib.org)

<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Game Season Team Coach Play Points Games Giants Second Players	Life Know School Street Man Family Says House Children Night	Film Movie Show Life Television Films Director Man Story Says	Book Life Books Novel Story Man Author House War Children	Wine Street Hotel House Room Night Place Restaurant Park Garden
<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
Bush Campaign Clinton Republican House Party Democratic Political Democrats Senator	Building Street Square Housing House Buildings Development Space Percent Real	Won Team Second Race Round Cup Open Game Play Win	Yankees Game Mets Season Run League Baseball Team Games Hit	Government War Military Officials Iraq Forces Iraqi Army Troops Soldiers
<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>
Children School Women Family Parents Child Life Says Help Mother	Stock Percent Companies Fund Market Bank Investors Funds Financial Business	Church War Women Life Black Political Catholic Government Jewish Pope	Art Museum Show Gallery Works Artists Street Artist Paintings Exhibition	Police Yesterday Man Officer Officers Case Found Charged Street Shot

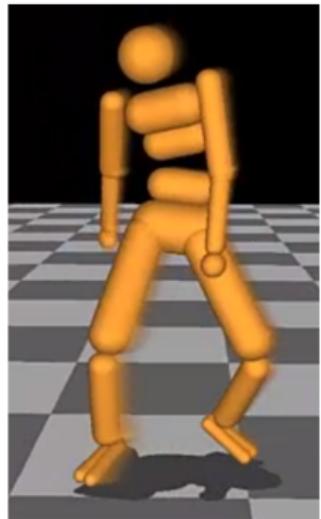
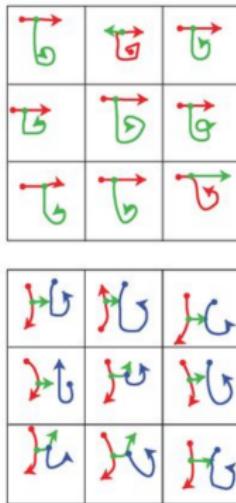
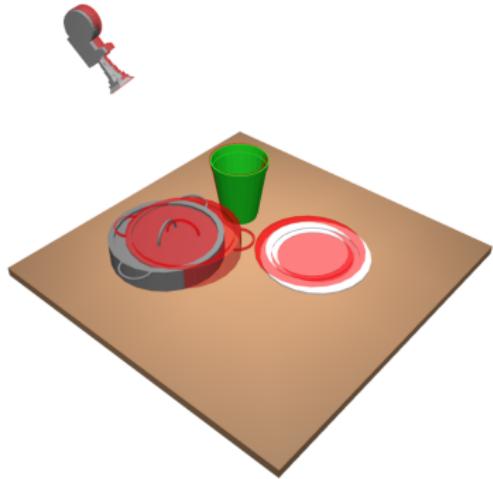
Topics found in 1.8M articles from the New York Times

[Hoffman, Blei, Wang, Paisley 2013]



Population analysis of 2 billion genetic measurements

[Gopalan+ 2017]



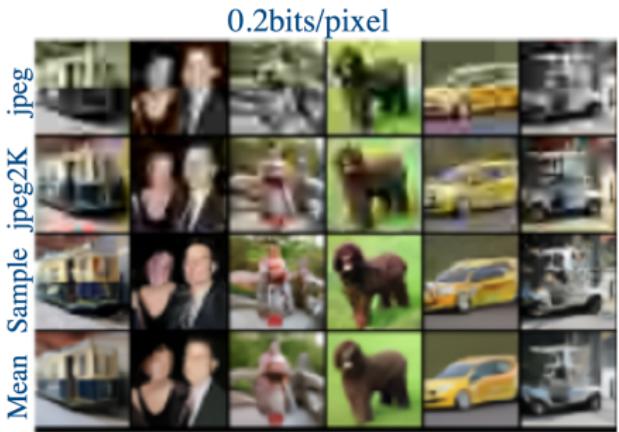
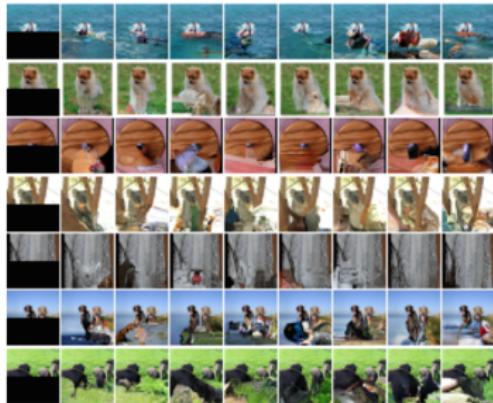
Understanding scenes, concepts, and control

[Eslami+ 2016; Lake+ 2015]



Exploratory analysis of 1.7M taxi trajectories, in Stan

[Kucukelbir+ 2017]



Compression and content generation

[Van der Oord+ 2016; Gregor+ 2016]

# George E.P. Box (1919 - 2013)

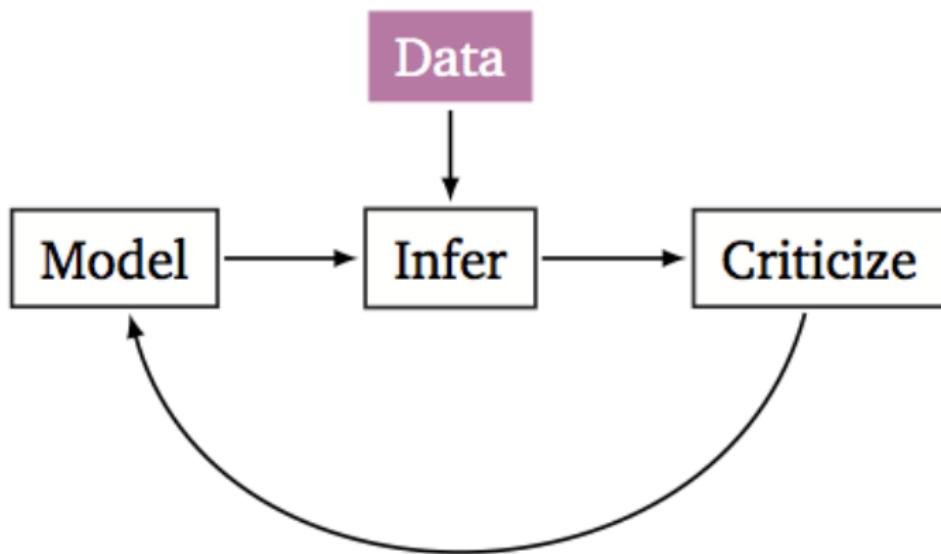


An iterative process for science:

1. Build a model of the science
2. Infer the model given data
3. Criticize the model given data

[Box & Hunter 1962, 1965; Box & Hill 1967; Box 1976, 1980]

## Box's Loop



Edward is a library designed around this loop.

[Box 1976, 1980; Blei 2014]

**Edward** is a probabilistic programming language built on TensorFlow.

### *Modeling*

- Composable Turing-complete language of random variables.
- Many data types, tensor vectorization, broadcasting, 3rd party support.
- Examples: Graphical models, neural networks, probabilistic programs.

### *Inference*

- Composable language for hybrids, message passing, data subsampling.
- Infrastructure to develop your own algorithms.
- Examples: Black box VI, Hamiltonian MC, stochastic gradient MCMC.

### *Criticism*

- Examples: Scoring rules, hypothesis tests, predictive checks.

Features include autodiff, multi-GPUs, distributed, XLA, quantization.

[Code](#)[Issues 99](#)[Pull requests 14](#)[Projects 1](#)[Pulse](#)[Graphs](#)

A library for probabilistic modeling, inference, and criticism. Deep generative models, variational inference. Runs on TensorFlow. <http://edwardlib.org>

[bayesian-methods](#)[deep-learning](#)[machine-learning](#)[data-science](#)[tensorflow](#)[neural-networks](#)[statistics](#)[probabilistic-programming](#)[1,680 commits](#)[17 branches](#)[24 releases](#)[45 contributors](#)Branch: [master](#) ▾[New pull request](#)[Find file](#)[Clone or download](#) ▾ [dustinvtran committed on GitHub](#) Improved training of Wasserstein GANs (#625) ...

Latest commit c92c141 a day ago

 [docker](#)

Make miscellaneous revisions to documentation (#562)

a month ago

[Sign Up](#)[Log In](#)

D

D

D

[all categories](#) ▾[Latest](#)[Top](#)

Topic

Category

Users

Replies

Views

Activity

Iterative estimators ("bayes filters") in Edward?



5

21

7h

Tutorial for multiple variational methods using Poisson regression?



2

20

1d



blei-lab/edward

A library for probabilistic modeling, inference, and criticism. <http://edwardlib.org>

Faez Shakil @faezs

Hi @dustinvtran, thanks for edward, the library and surrounding literature have been immense fun to get into. Would you be able to tell me whether it'd be relatively painless to get the inference compute graphs from Ed as native tensorflow graphdefs and use them on mobile platforms? Or would I have to port a bunch of custom ops

Jan 23 02:47

PEOPLE REPO INFO



We have an active community of several hundred users & many contributors.

# Who is Using Edward?

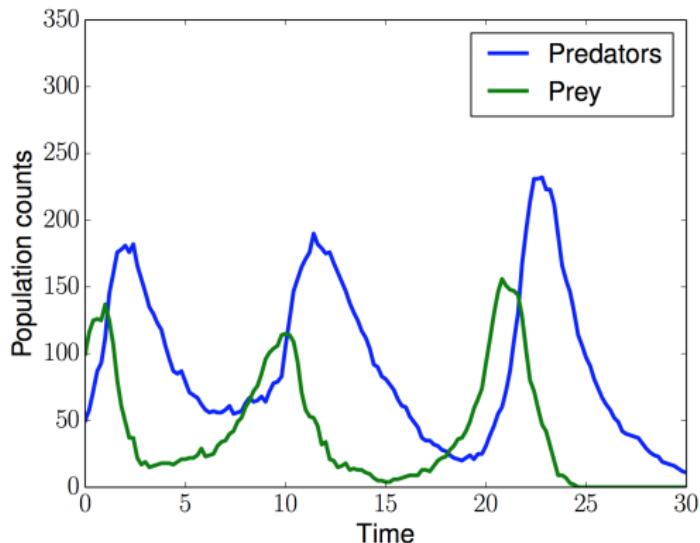
## Users

1. Machine learning enthusiasts, data scientists, business analysts  
*(ex. hierarchical GLMs, mixture models, MAP, MCMC, ...)*
2. Probabilistic graphical modeling community  
*(ex. latent Dirichlet allocation, variational inference, Gibbs)*
3. Bayesian deep learning community  
*(ex. deep generative models, Bayesian NNs, black box inference)*

## Developers

1. David Blei's group
2. Google Brain (*in conception/design*)
3. Matt Hoffman (*conjugacy*), Emily Fox's group (*time series + SGMCMC*),  
Justin Bayer (*stochastic RNNs*), John Pearson (*neuroscience*), a few  
Master's/Ph.D. students.
4. Collaboration continues to evolve. Contact us! (+visit the Forum)

# Simulator-based Model in Population Ecology



```
1 beta = LogNormal(tf.zeros(3), tf.ones(3))
2 x1 = [] # prey population
3 x2 = [] # predator population
4 for t in range(1000):
5     x1[t+1] = beta[0] * x1[t] - beta[1] * x1[t] * x2[t] + Normal(0.0, 10.0)
6     x2[t+1] = -beta[1] * x2[t] + beta[1] * x1[t] * x2[t] + Normal(0.0, 10.0)
```