

BOOMS & BUSTS

THE STORY OF US

Matthew J. Orr



HOW DID WE GET HERE?



Business, 1996 – 2017

WIRED

THE DATA

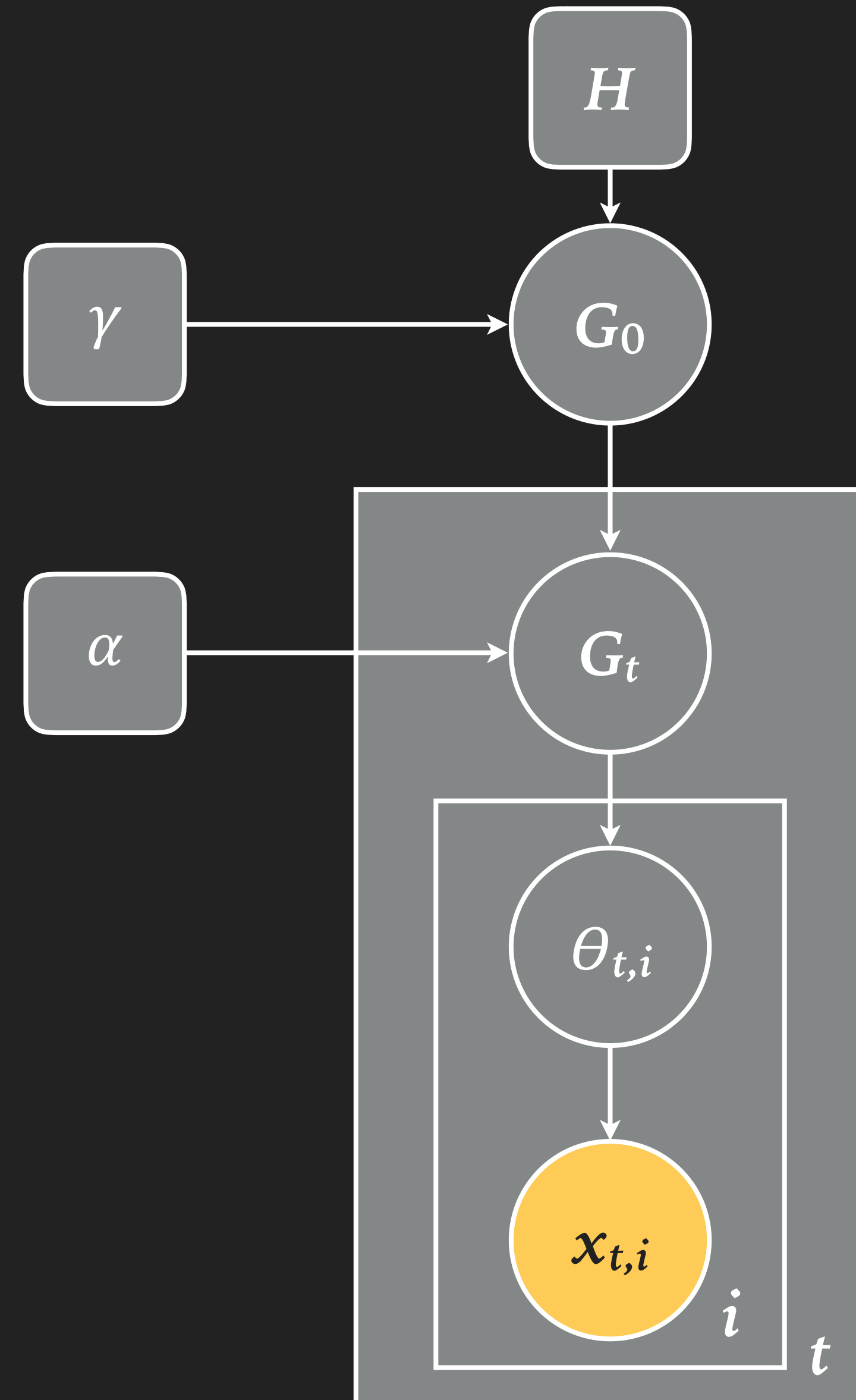
- ▶ **23,918** articles written at a college-grad reading level
- ▶ **17,918,810** words averaging 6 characters each
- ▶ 89,594-minute reading time at 200 wpm (over **2 months**)
- ▶ Stored as JSON

TOOLS & TECHNIQUES

- ▶ NLTK, Gensim
- ▶ Tokenization
- ▶ Part-of-Speech Tagging
- ▶ Lemmatization
- ▶ HDP-LDA Model

HIERARCHICAL DIRICHLET PROCESS (HDP)

- ▶ The HDP-LDA (or HDP mixture) model is a natural nonparametric generalization of Latent Dirichlet Allocation (LDA), where the **number of topics can be unbounded** and **learnt from data**.
- ▶ Each group is a document consisting of a bag of words, each cluster is a topic, and each document is a mixture of topics.
- ▶ The HDP-LDA model identified topics in each year **independently**. The **top 10** most meaningful topics were chosen from each year based on their **coherence value**.



1996 — How It All Started

1. Human Rights: Especially Tibetan Independence
2. Speculation about Tech Giants & Newcomers
3. Apple (General)
4. Market Expansion: Globalization & the Younger Generation
5. TV and Music Industry Explore Using Websites
6. "Millenium Bug" (Y2K)
7. Steve Jobs Rejoins Apple
8. The Internet is a Great Platform
9. Finance 1
10. Finance 2

2000 — The Year of Music

1. Finance 1 (Internet Stocks)
2. Online Music Streaming & Advertising
3. Music-Related Lawsuits
4. Finance 2 (Reverse Stock Split)
5. General Tech 1
6. Music Software
7. Music Tech/Gear
8. Finance 3
9. Women Tech, Women Yes
10. Finance 4

2001 — A World in Turmoil

1. Finance 1
2. Privacy Concerns from Internet Data
3. Finance 2
4. Finance 3
5. Finance 4
6. 9/11
7. War on Terror
8. Second Intifada
9. Mobile Phones
10. Apple/Apple Stores

2003 —



1. SPAM is Everywhere
2. Security: Homeland, Software & Systems
3. Mobile Phone Technology Advancements
4. SCO-Linux Controversies
5. Junk Fax (SPAM Fax!)
6. New Machine Technology
7. Rise of WiFi in the Mainstream
8. Information Security Concerns
9. Increased Interest in Linux & Other Open-Source Tech
10. Mozilla Tov!

2004 — Are You Not Entertained?

1. Risky Behavior: salmon, Lamo and Fax.com
2. File-Sharing
3. General Tech 1
4. Linux vs. Microsoft
5. Music & Video Games
6. Spam Filters Finally Work
7. Mobile Games, Music & Ringtones
8. Mydoom & Sassy Worms
9. The Patriot Act
10. Massive Job Cuts

2007 — The Facebook

1. Social Networks
2. Apple
3. Facebook is a Thing
4. Intel
5. Google
6. iPhone
7. Apple TV
8. Real Life vs. Second Life
9. Ribbit & Other Startups
10. FCC Spectrum Auction

2015 — Silicon Celebrities

1. General Tech 1
2. Presidential Primaries:
Trump, Clinton & Sanders
3. Jack Dorsey
4. Uber
5. Women & Diversity at the
Workplace
6. Apple
7. Drones
8. Pinterest & Facebook
9. General Tech 3
10. Mark Zuckerberg

2016 — Off to the Races

1. Presidential Campaign:
Trump, Clinton & Sanders
2. Uber
3. Jack Dorsey
4. Republican Convention:
Trump and Cruz Face Off
5. Women & Diversity at the
Workplace
6. General Tech 1
7. Google Fiber
8. General Tech 2
9. Net Neutrality
10. Tim Cook

2017 — The Year of Trump

1. General Tech 1
2. Net Neutrality
3. Trump's Presidency
4. Trump's Twitter
5. Trump vs. the Courts
6. Trump around the World
7. Trump and Immigration
8. "Platform" is the Hot Buzzword,
as is "Trump"
9. Amazon
10. Autonomous Cars



**“CELEBRITY IS THE
CHASTISEMENT OF
MERIT AND THE
PUNISHMENT OF
TALENT.”**

Emily Dickinson

THANK YOU

Matthew J. Orr | matt@matthewjorr.com | Github: [mattheworr](#)