

Reading the Tea Leaves: How Humans Interpret Topic Models

Chang, Graber, Gerrish, Wang, & Blei

NIPS 2009

Presenter: Robert Thorstad

Topic Modeling

People's documents usually organized around topics

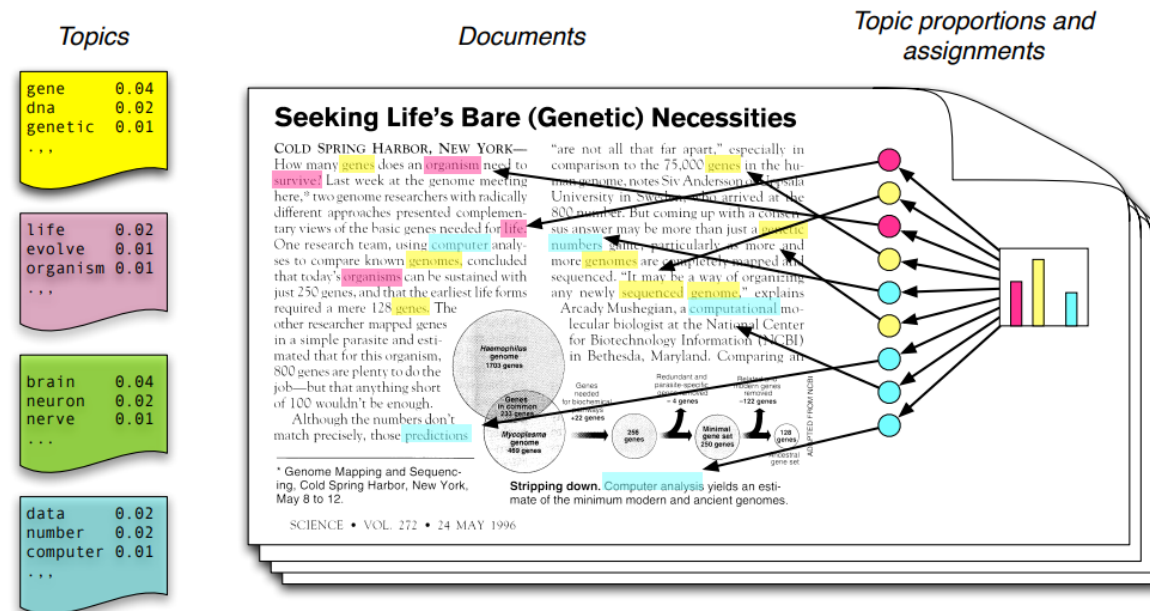


Figure: personal website of David Blei

Topic models reverse this inference: observe documents → infer topics

How know a topic is good?

Measure how well they fit new data:

Predictive Likelihood: How well do topics predict words unobserved posts?

Coherence: How much do top words in a topic “go together” in unobserved posts? (e.g. Roder, Both, & Hinneburg, 2015)

Good topics? Humans should think:

Words in a topic go together

Topics well describe document

Word Intrusion

1 / 10	floppy	alphabet	computer	processor	memory	disk
2 / 10	molecule	education	study	university	school	student
3 / 10	linguistics	actor	film	comedy	director	movie
4 / 10	islands	island	bird	coast	portuguese	mainland

Topic Intrusion

6 / 10

DOUGLAS_HOFSTADTER

Douglas Richard Hofstadter (born February 15, 1945 in New York, New York) is an American academic whose research focuses on consciousness, thinking and creativity. He is best known for ", first published in

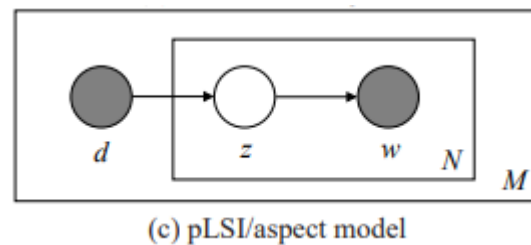
Show entire excerpt

student	school	study	education	research	university	science	learn
human	life	scientific	science	scientist	experiment	work	idea
play	role	good	actor	star	career	show	performance
write	work	book	publish	life	friend	influence	father

Studied 3 topic models

Probabilistic Latent Semantic Indexing (pLSI):

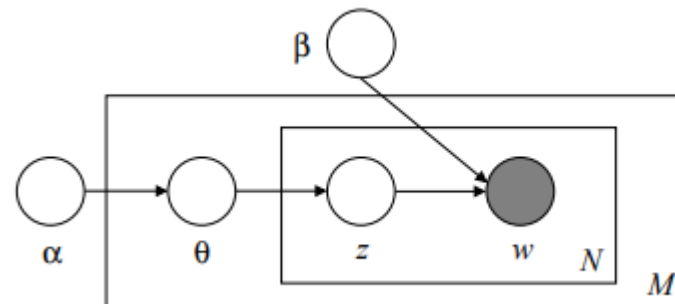
1 topic / document



(Hoffman, 1999; Figure from Blei, 2003)

Latent Dirichlet Allocation (LDA):

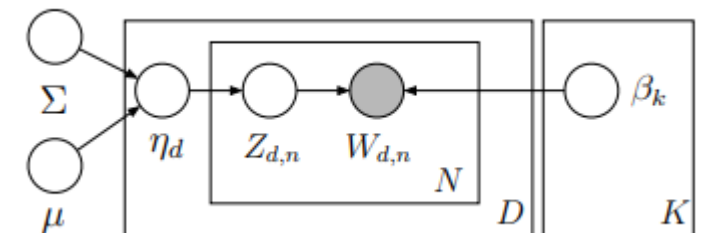
(Potentially) many topics/document



(Blei, 2003)

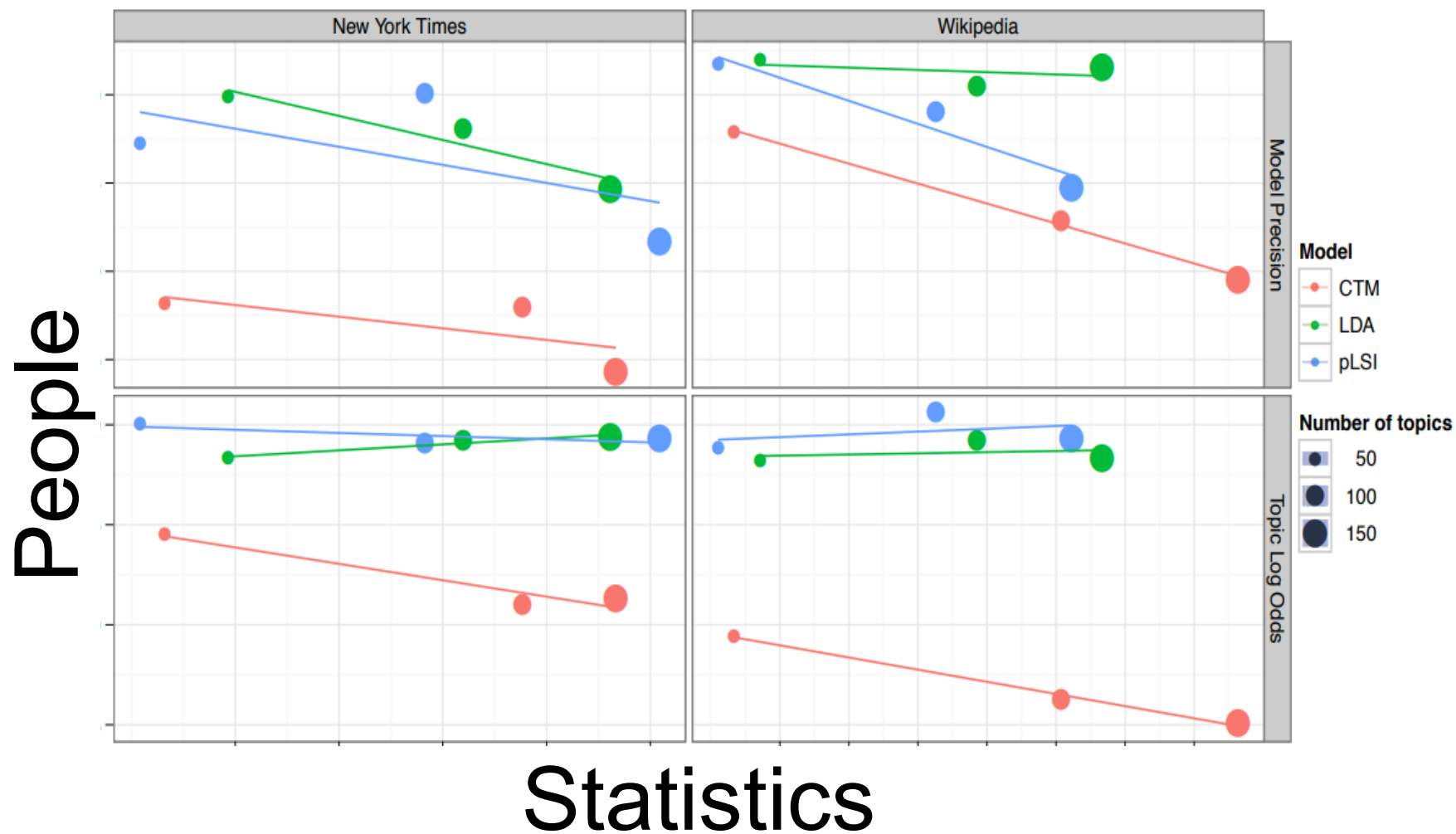
Correlated Topic Model (CTM)

Topics can be correlated



(Blei & Lafferty, 2006)

Statistics negatively correlate with human judgments!



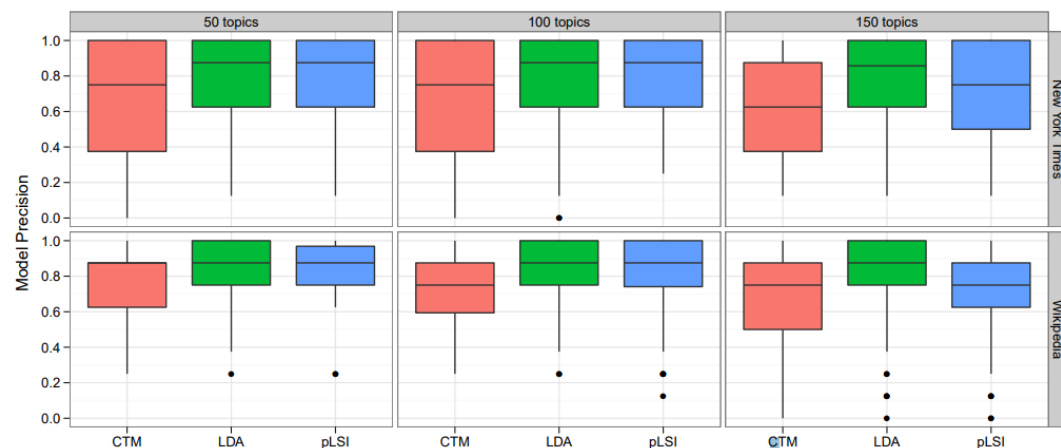
Statistically, CTM was best

Log odds:

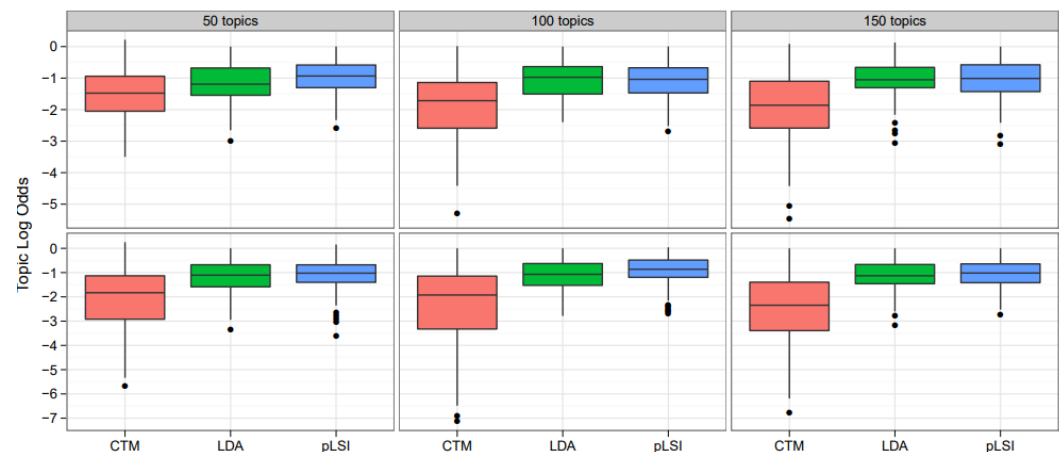
CORPUS	TOPICS	LDA	CTM	PLSI
NEW YORK TIMES	50	-7.3214 / 784.38	-7.3335 / 788.58	-7.3384 / 796.43
	100	-7.2761 / 778.24	-7.2647 / 762.16	-7.2834 / 785.05
	150	-7.2477 / 777.32	-7.2467 / 755.55	-7.2382 / 770.36
WIKIPEDIA	50	-7.5257 / 961.86	-7.5332 / 936.58	-7.5378 / 975.88
	100	-7.4629 / 935.53	-7.4385 / 880.30	-7.4748 / 951.78
	150	-7.4266 / 929.76	-7.3872 / 852.46	-7.4355 / 945.29

But, people liked LDA best

Word intrusion



Topic intrusion



Which measures agree with people?

Topic coherence

coherences	Name	C_V	C_P	C_{UMass}	$C_{one-any}$	C_{UCI}	C_{NPMI}	C_A
	\mathcal{S}	S_{set}^{one}	S_{pre}^{one}	S_{pre}^{one}	S_{any}^{one}	S_{one}^{one}	S_{one}^{one}	S_{one}^{one}
	\mathcal{P}	$\mathcal{P}_{sw(110)}$	$\mathcal{P}_{sw(70)}$	\mathcal{P}_{bd}	\mathcal{P}_{bd}	$\mathcal{P}_{sw(10)}$	$\mathcal{P}_{sw(10)}$	$\mathcal{P}_{cw(5)}$
	\mathcal{M}	$\bar{m}_{cos(nlr,1)}$	m_f	m_{lc}	m_d	m_{lr}	m_{nlr}	$\bar{m}_{cos(nlr,1)}$
	Σ	σ_a	σ_a	σ_a	σ_a	σ_a	σ_a	σ_a
using corpus	20NG	0.665	0.756	0.395	0.563	0.312	0.486	0.563
	Genomics	0.671	0.652	0.514	0.549	0.624	0.630	0.632
	RTL-Wiki	0.627	0.615	0.272	0.545	0.527	0.573	0.542
	Movie	0.548	0.549	0.093	0.453	0.473	0.438	0.431
	average	0.628	0.643	0.319	0.528	0.484	0.532	0.542
using the Wikipedia	$N = 10$	20NG	0.859	0.825	0.562	0.822	0.696	0.739
		Genomics	0.773	0.721	0.442	0.452	0.478	0.530
		NYT	0.803	0.757	0.543	0.612	0.783	0.747
		average	0.812	0.768	0.516	0.629	0.652	0.672
	$N = 5$	RTL-NYT	0.728	0.720	0.106	0.438	0.631	0.687
		RTL-Wiki	0.679	0.645	0.350	0.499	0.558	0.602
		Movie	0.544	0.533	0.143	0.454	0.447	0.465
		average	0.650	0.633	0.200	0.464	0.545	0.585
	average		0.731	0.700	0.358	0.546	0.599	0.628

Table 2: Coherence measures with strongest correlations with human ratings.

(Roder, Both, & Hinneburg, 2015)