

## **Future Innovation Predictions through Signatures of Past Novelty**

Kristoffer L Nielbo, Peter B. Vahlstrup & Anja Bechmann

`kln@cas.dk`

`knielbo.github.io`

`chcaa.io`

Center for Humanities Computing  
Aarhus University, Denmark

# Outline

- 1 Introduction
  - NGI & innovation
  - Trend detection
- 2 Methods
  - Semantics
  - Resonance
  - Persistence
- 3 Results
  - Experiment 1
  - Experiment 2
- 4 Discussion
  - Concluding remarks

## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### Introduction

NGI & innovation  
Trend detection

### Methods

Semantics  
Resonance  
Persistence

### Results

Experiment 1  
Experiment 2

### Discussion

Concluding remarks



Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### H2020-ICT, Next Generation Internet - An Open Internet Initiative

*Iteration 1* Modeled the evolution of \*innovation\* in NGI concepts → 'trend detection for (sets of) keyword frequencies'

Concept	keywords
<i>AI</i>	{ <i>ai, facial recognition, ..., deepfake</i> }
<i>IoT</i>	{ <i>iot, iiot, ..., bgp</i> }
...	...
<i>Quantum computing</i>	{ <i>quantum computing, qubit, ..., dwave</i> }

Concepts and keywords for popular tech media that characterize NGI

*Iteration 2*: mapping trends and dynamics of public debate (constrained by iter1)

Introduction

NGI & innovation

Trend detection

Methods

Semantics

Resonance

Persistence

Results

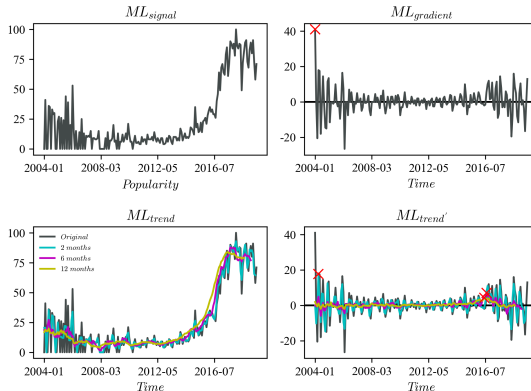
Experiment 1

Experiment 2

Discussion

Concluding remarks

# Trend detection



## Problem

reverse time order  
trend spikiness  
trend slope  
atomistic semantics  
convenience sampling

## Solution

trend reservoir  
trend resonance  
trend persistence  
distributional semantics  
add random condition

## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L. Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### Introduction

NGI & innovation

### Trend detection

### Methods

Semantics

Resonance

Persistence

### Results

Experiment 1

Experiment 2

### Discussion

Concluding remarks





- \* Reddit.com ( $\sim 7TB$ )
- \* Representation based on titles
- \* Samples are generated from keyword match in subreddit description
- \* Samples:
  1. *target trend*
  2. *random*
  3.  $\neg$ *target trend*\*

## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### Introduction

NGI & innovation  
Trend detection

### Methods

Semantics  
Resonance  
Persistence

### Results

Experiment 1  
Experiment 2

### Discussion

Concluding remarks

Model each document as a distribution on lexical topics, e.g.,  
 $s = [0.09 \ 0.78 \ 0.11 \ 0.2]$ , where each ‘topic’ is a distribution on words, and compare document similarity as the distance between any two documents with *chapter – index*  $j$  and  $k$ :

$$D_{KL}(s^{(j)} \mid s^{(k)}) = \sum_{i=1}^K s_i^{(j)} \times \log_2 \frac{s_i^{(j)}}{s_i^{(k)}}$$

Reduce interpretive load and compare relative entropy between documents on topics (“topical variation between documents”)

Works for any model that ‘embeds’ a document in a vector space, but equations have to be modified for geometrical models (e.g., *NMF*).

Introduction

NGI & innovation

Trend detection

Methods

Semantics

Resonance

Persistence

Results

Experiment 1

Experiment 2

Discussion

Concluding remarks



Compute resonance as the difference between of novelty and transience:

Novelty over window  $w$ :

$$\mathbb{N}_w(j) = \frac{1}{w} \sum_{d=1}^w D_{KL}(s^{(j)} \mid s^{(j-d)})$$

Transience:

$$\mathbb{T}_w(j) = \frac{1}{w} \sum_{d=1}^w D_{KL}(s^{(j)} \mid s^{(j+d)})$$

Resonance

$$\mathbb{R}_w(j) = \mathbb{N}_w(j) - \mathbb{T}_w(j)$$

Introduction

NGI & innovation

Trend detection

Methods

Semantics

**Resonance**

Persistence

Results

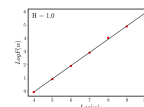
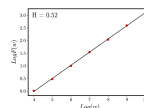
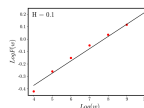
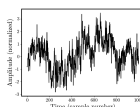
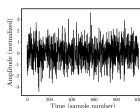
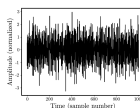
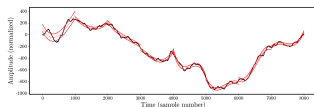
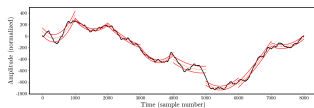
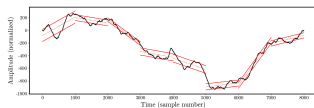
Experiment 1

Experiment 2

Discussion

Concluding remarks

# Persistence: Fractal scaling



For  $1/f^{2H+1}$  processes: anti-persistent process:  $0 < H < 0.5$ , short-range correlations only  $H = 0.5$ , and  $0.5 < H < 1$  persistent process

K. L. Nielbo, K. F. Baunvig, B. Liu, and J. Gao, "A curious case of entropic decay: Persistent complexity in textual cultural heritage," Digital Scholarship in the Humanities, 2018

M. Wevers, J. Gao, and K. L. Nielbo, "Tracking the Consumption Junction: Temporal Dependencies in Dutch Newspaper Articles and Advertisements, arxiv:1903.11461

## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

Introduction

NGI & innovation

Trend detection

Methods

Semantics

Resonance

Persistence

Results

Experiment 1

Experiment 2

Discussion

Concluding remarks



# Target: $r/\text{MachineLearning}$

## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### Introduction

NGI & innovation

Trend detection

### Methods

Semantics

Resonance

Persistence

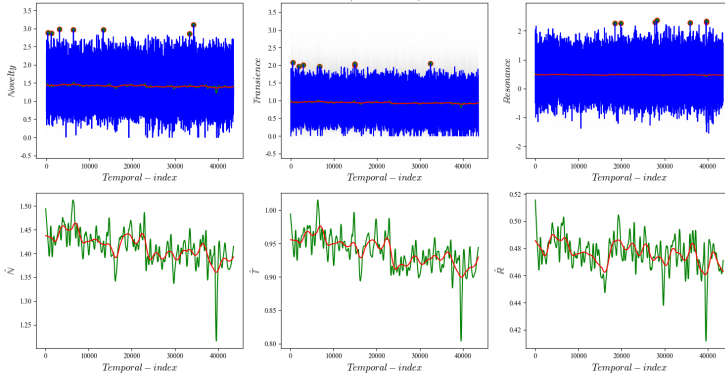
### Results

Experiment 1

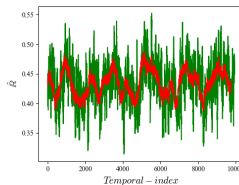
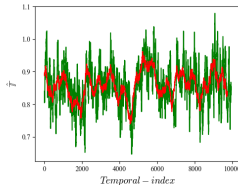
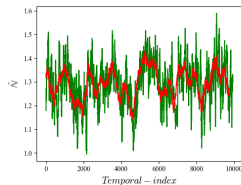
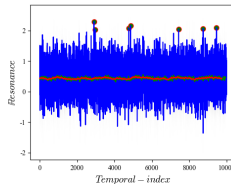
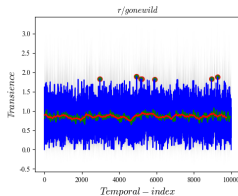
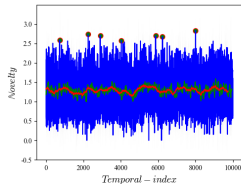
Experiment 2

### Discussion

Concluding remarks



# Random: r/gonewild



## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L. Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### Introduction

NGI & innovation  
Trend detection

### Methods

Semantics  
Resonance  
Persistence

### Results

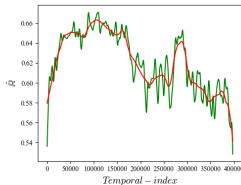
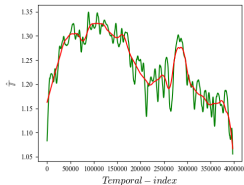
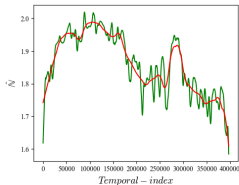
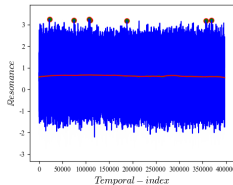
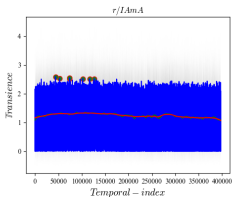
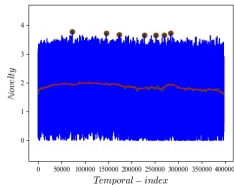
Experiment 1  
Experiment 2

### Discussion

Concluding remarks



# Non-tech: r/IAmA



## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

### Introduction

NGI & innovation  
Trend detection

### Methods

Semantics  
Resonance  
Persistence

### Results

Experiment 1  
Experiment 2

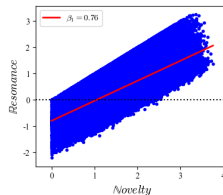
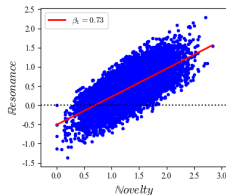
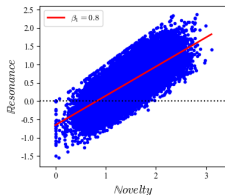
### Discussion

Concluding remarks

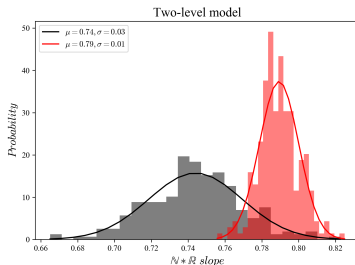


# Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io



*r/ MachineLearning, r/ gonewild, r/ IAmA*



$t_{100} = 26.03, p < .00001$

## Introduction

NGI & innovation  
Trend detection

## Methods

Semantics  
Resonance  
Persistence

## Results

Experiment 1  
Experiment 2

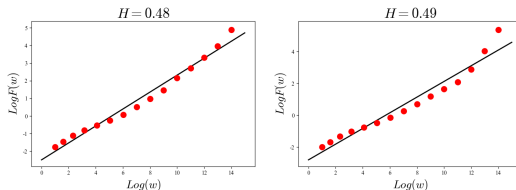
## Discussion

Concluding remarks

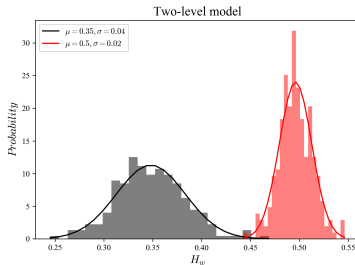


# Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io



*r/ MachineLearning, r/Python*



Introduction

NGI & innovation

Trend detection

Methods

Semantics

Resonance

Persistence

Results

Experiment 1

Experiment 2

Discussion

Concluding remarks



## Concluding remarks

## Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

For trend reservoirs we suggest that:

- \*  $N * \mathbb{R}$  slope is steeper in comparison to a random\* baseline
- \* short-range correlations with a tendency to long-term correlations on longer time scales
- \* indications of multi-fractal behavior, probably driven by subreddit “rate of change”
- \* saturated subreddits show mean-reverting behavior for resonance

So to predict future innovations in streams of data, group data on

$N * \mathbb{R}$  slope  $\geq .8$  and  $0.5 \geq H \leq 1.0$

Introduction

NGI & innovation

Trend detection

Methods

Semantics

Resonance

Persistence

Results

Experiment 1

Experiment 2

Discussion

Concluding remarks



# Future Innovation Predictions through Signatures of Past Novelty

Kristoffer L Nielbo,  
Peter B. Vahlstrup &  
Anja Bechmann  
kln@cas.dk  
knielbo.github.io  
chcaa.io

Thank you for your attention

kln@au.dk  
knielbo.github.io  
chcaa.io

slides: [http://knielbo.github.io/files/kln\\_reddit.pdf](http://knielbo.github.io/files/kln_reddit.pdf)

Introduction

NGI & innovation

Trend detection

Methods

Semantics

Resonance

Persistence

Results

Experiment 1

Experiment 2

Discussion

Concluding remarks

