



Multi-modal topic inference from video

Oron Nir, Senior Data Scientist, Media AI

Daniel Nurieli, Data Scientist, Media AI



Insights Timeline

Download Translate View Edit

Search

161 People

Satya Nadella
Microsoft CEO
Appears in 12.07% of video Show biography Find on Bing

16 Topics

Technology / Information systems / Cloud computing Web development

Technology / Big data

Science / Mathematics / Machine learning

Education / Technology

Business / Startups / Startup advice and strategy

Companies Software Smartphones Word embeddings Products Microsoft Management Integrated development environments Serverless computing

Visual studio code

124 Labels

person indoor screenshot standing man outdoor floor monitor computer laptop screen wall electronics sky +110

69 Brands

Microsoft Microsoft Visual Studio Starbucks Microsoft Cortana Microsoft Windows Microsoft Edge Java Twitter +61

Microsoft Show description

01:07:41

Microsoft Build 2018 - Opening keynote FULL

Public Created 4 months ago by Oron Nir

Embed Download Report Video editor

More videos with similar people and keywords

- SRZ33774B0101F0A_TC_EP-13_Medical_Detectives_Geheimnisse_der_Ge...
Created 10 months ago by Oron Nir
- Oprah%20Winfrey%20Receives%20Cecil%20B...%20de%20Mille%20Award%20at%20the%2...
Created 9 months ago by Oron Nir
- Why David Blatt Was Fired By The Cavaliers
Created 10 months ago by Oron Nir
- Hillary
Created 9 months ago by Oron Nir
- TonightNewsBroadcast
Created one year ago by Oron Nir

16 Topics



Technology / Information systems / **Cloud computing** **Web development**

Technology / **Big data**

Science / Mathematics / **Machine learning**

Education / **Technology**

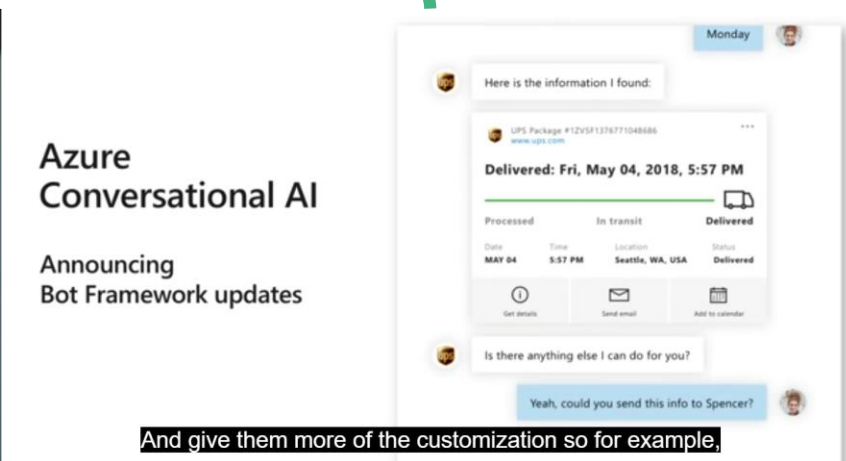
Business / Startups / **Startup advice and strategy**

- Companies
- Software
- Smartphones
- Word embeddings
- Products
- Microsoft
- Management
- Integrated development environments
- Serverless computing
- Visual studio code

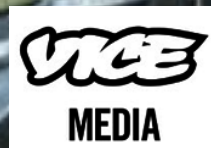


Play previous

Play next



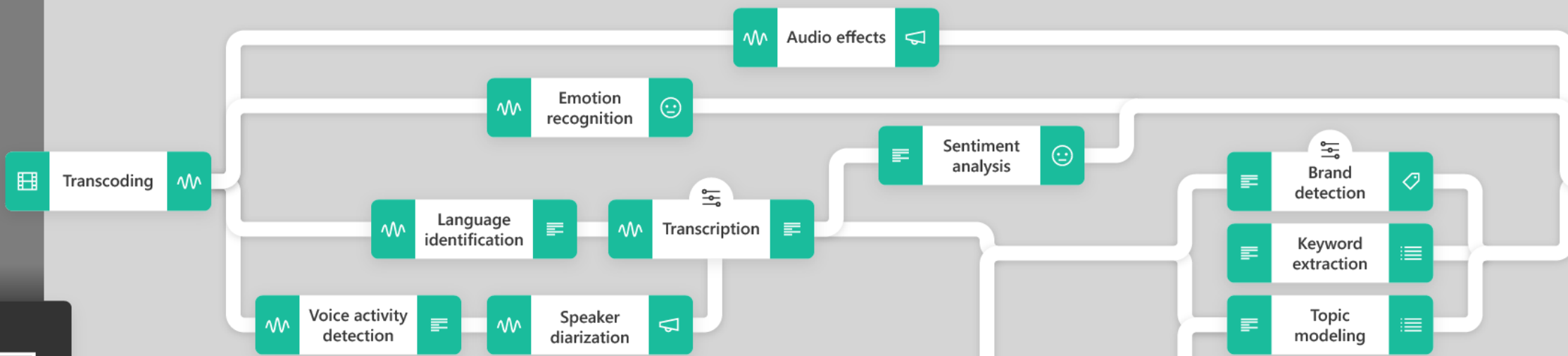
Topic inference in the media business



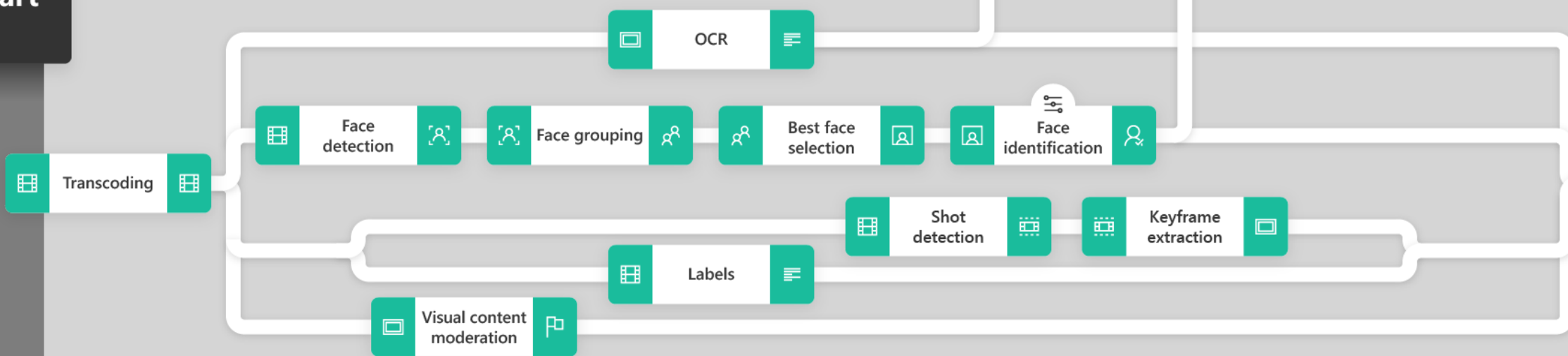
Media Topics



Audio analysis



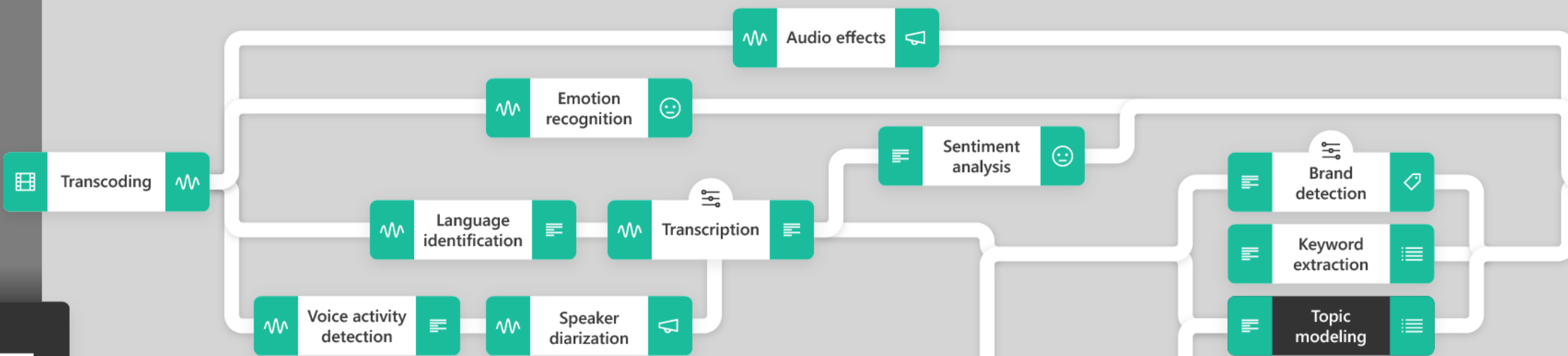
Vision analysis



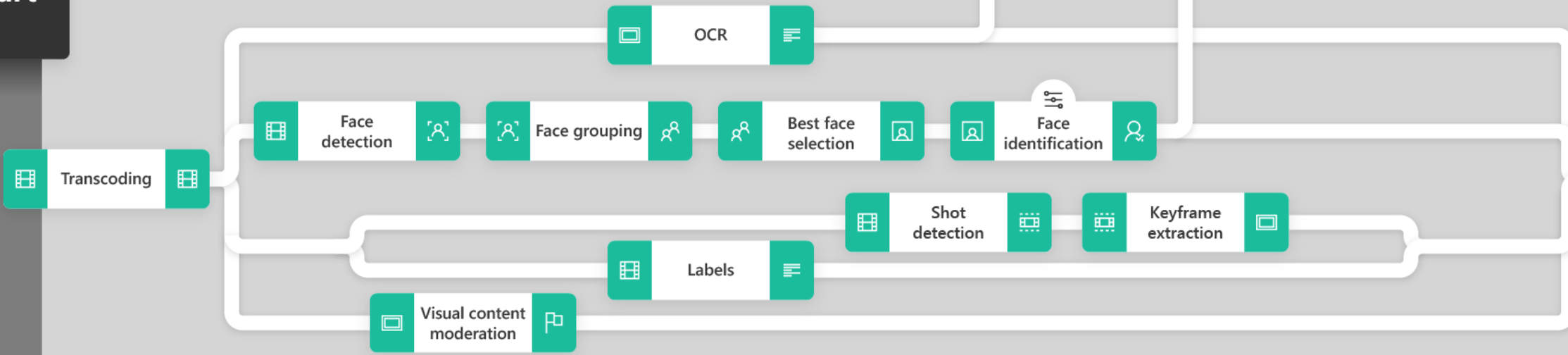
Start

Video index

Audio analysis



Vision analysis

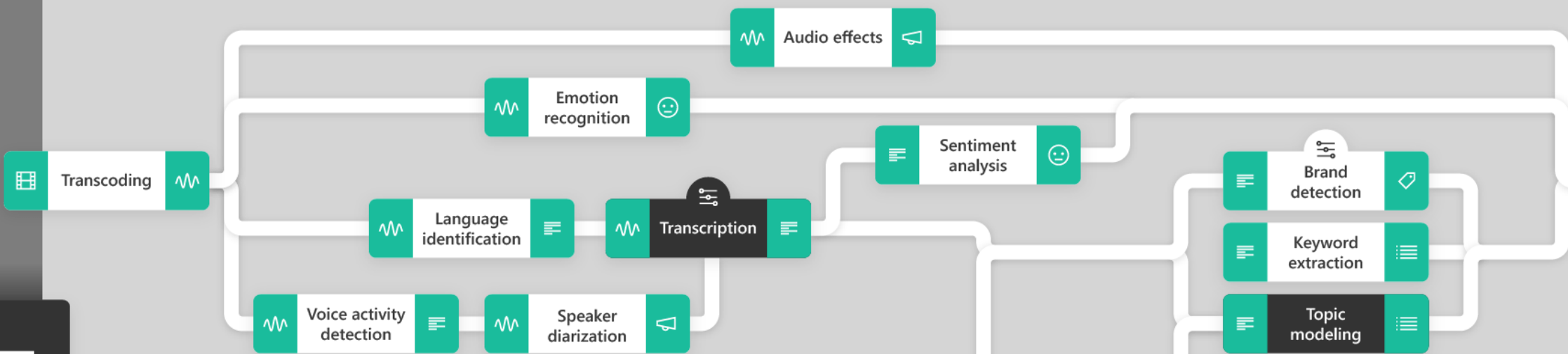


Start

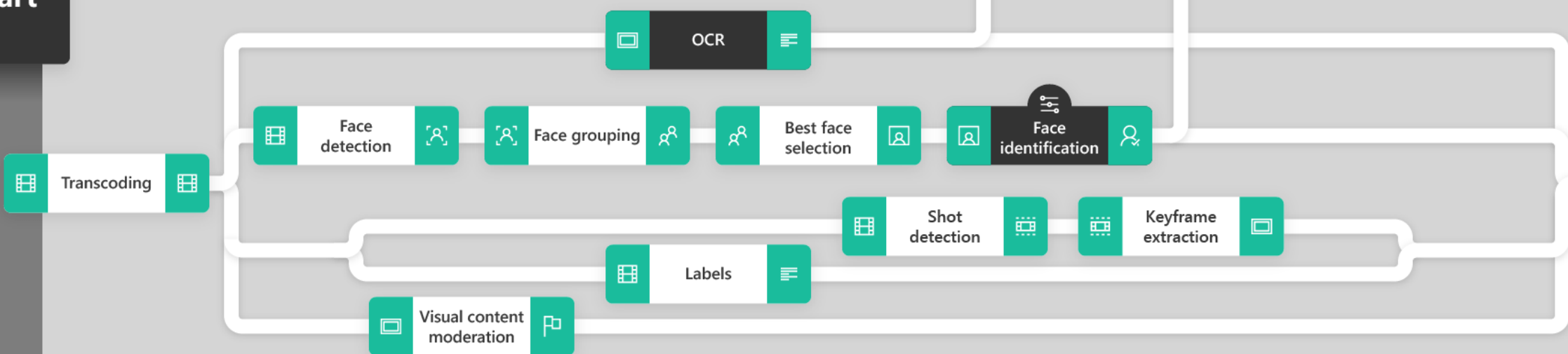


Video index

Audio analysis



Vision analysis

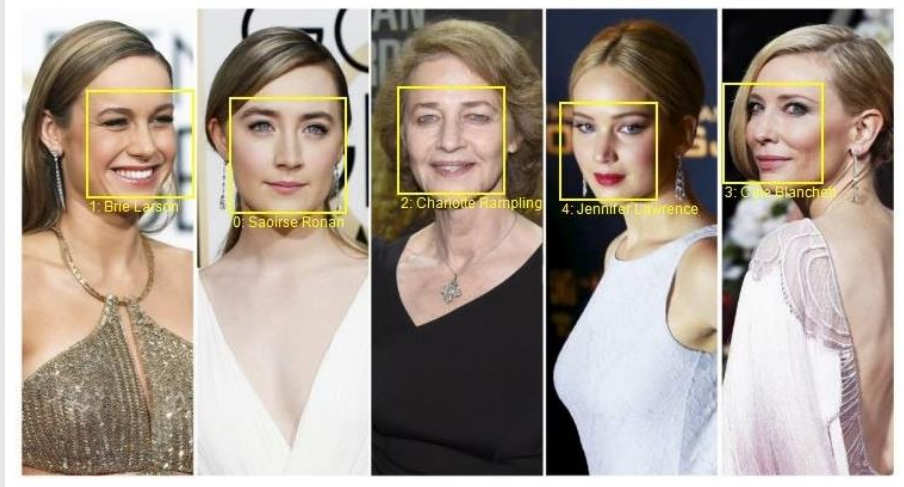


Start

Video index

Input signals

Face recognition - 1Million celebs



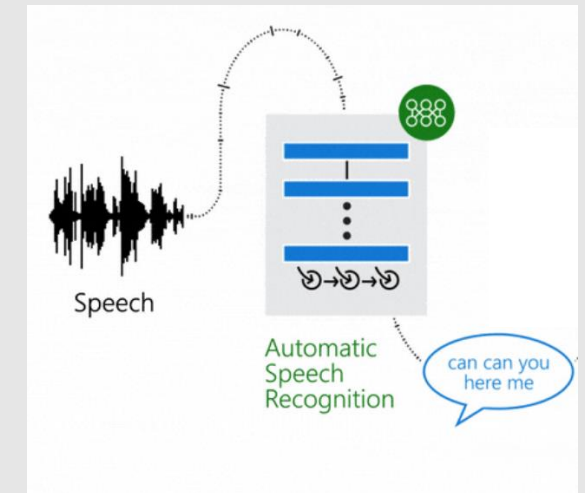
[Yandong, G., Lei, Z., Yuxiao, H., Xiaodong, H., & Jianfeng, G. \(2016\). MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition. ECCV.](#)

Optical Character Recognition



[OCR Video Indexer](#)

Speech to Text



Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M., Stolcke, A., Zweig, G. (2016). Achieving Human Parity in Conversational Speech Recognition.



WIKIPEDIA
The Free Encyclopedia

Article **Talk**

Read | View source | View history

Search Wikipedia

Edward Snowden



From Wikipedia, the free encyclopedia

Edward Joseph Snowden (born June 21, 1983) is an American [computer professional](#), former [Central Intelligence Agency](#) (CIA) employee, and former contractor for the [United States government](#) who copied and leaked classified information from the [National Security Agency](#) (NSA) in 2013. His disclosures revealed numerous global surveillance programs, many run by the NSA and the [Five Eyes Intelligence Alliance](#) with the cooperation of [telecommunication companies](#) and European governments.

In 2013, Snowden was hired by an NSA contractor, [Booz Allen Hamilton](#), after previous employment with [Dell](#) and the CIA.^[1] On May 20, 2013, Snowden flew to [Hong Kong](#) after leaving his job at an NSA facility in Hawaii, and in early June he revealed thousands of classified NSA documents to journalists [Glenn Greenwald](#), [Laura Poitras](#), and [Ewen MacAskill](#). Snowden came to international attention after stories based on the material appeared in *The Guardian* and *The Washington Post*. Further disclosures were made by other publications including *Der Spiegel* and *The New York Times*.

On June 21, 2013, the U.S. Department of Justice unsealed charges against Snowden of two counts of violating the



Edward Snowden

Categories: [Edward Snowden](#) | [1983 births](#) | [Living people](#) | [American computer specialists](#) | [American exiles](#) | [American whistleblowers](#) | [Booz Allen Hamilton people](#) | [Fugitives wanted by the United States](#) | [Fugitives wanted under the Espionage Act of 1917](#) | [National Security Agency people](#) | [People from Elizabeth City, North Carolina](#) | [People of the Central Intelligence Agency](#) | [People of the Defense Intelligence Agency](#) | [People with epilepsy](#) | [Privacy activists](#) | [Refugees in Russia](#) | [American refugees](#)

This page was last edited on 9 November 2018, at 21:36 (UTC).

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.

Input

1. Raw STT transcript
2. OCR

STT: "The intelligence collection that the [NSA](#) gathers, the data itself isn't the most sensitive aspect to it. It's how you get that data and that's what [Snowden](#) revealed."

Output

Wikipedia pages

National Security Agency

From Wikipedia, the free encyclopedia

Edward Snowden

From Wikipedia, the free encyclopedia



ELIS



Transcript, OCR
and Celebs

Entity
Linking

Input

1. Raw STT transcript
2. OCR

STT: "So Fishing is a big issue..."

OCR: "Phishing and Web Exploits Ransomware"

Fishing

From Wikipedia, the free encyclopedia

Phishing

From Wikipedia, the free encyclopedia

Output

Wikipedia pages



ELIS



Transcript, OCR
and Celebs

Entity
Linking

Input

Set of entities
(from transcript,
OCR, and faces)

Edward Snowden

From Wikipedia, the free encyclopedia

Edward Joseph Snowden (born June 21, 1983) is an American [computer professional](#), former [Central Intelligence Agency \(CIA\)](#) employee, and former contractor for the [United States government](#) who copied and [leaked classified information](#) from the [National Security Agency \(NSA\)](#) in 2013. His disclosures revealed numerous [global surveillance](#) programs, many run by the NSA and the

Output

Wikipedia pages



ELIS



Transcript, OCR
and Celebs

Entity
Linking

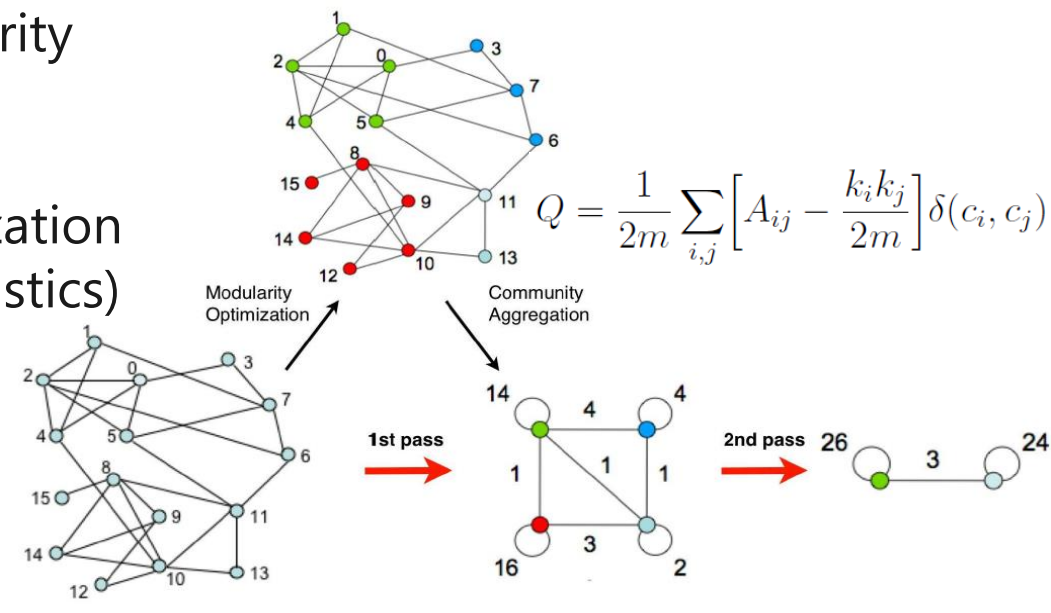
Wikipedia
dump

Input

TFIDF Cosine Similarity
entity-graph

Modularity maximization
(a.k.a. Louvain heuristics)

Find optimal K



Output

Clusters



Transcript, OCR
and Celebs

ELIS



Entity
Linking



Wikipedia
dump



Build and cluster
a graph

Input

Categories

Model: PCA Whitening
+ Random Forest

Data: 10k data points

Topics: 1/2 per cluster
2-8 clusters per video

Abstain policy: < 8 entities

Feature vector: 21 features

- Graph level
- Cluster level e.g. Diameter
- Entity level e.g. Centrality
- Category level e.g. Jaccard IoU

Output

Top ranked
categories per
cluster



ELIS



Transcript, OCR
and Celebs

Entity
Linking

Wikipedia
dump

Build and cluster
a graph

Rank categories
per cluster

Government agencies ^

EDWARD SNOWDEN &
DEFENDING AGAINST THE
INSIDER THREAT



IEEE
computer
society

⏪ Play previous

Play next ⏩



ELIS



Transcript, OCR
and Celebs

Entity
Linking

Wikipedia
dump

Build and cluster
a graph

Rank categories
per cluster

Link
to the timeline

EDWARD SNOWDEN & DEFENDING AGAINST THE INSIDER THREAT



<https://www.youtube.com/watch?v=kQVLoNmYtKA>



**Steven Bay talks
about Edward
Snowden**

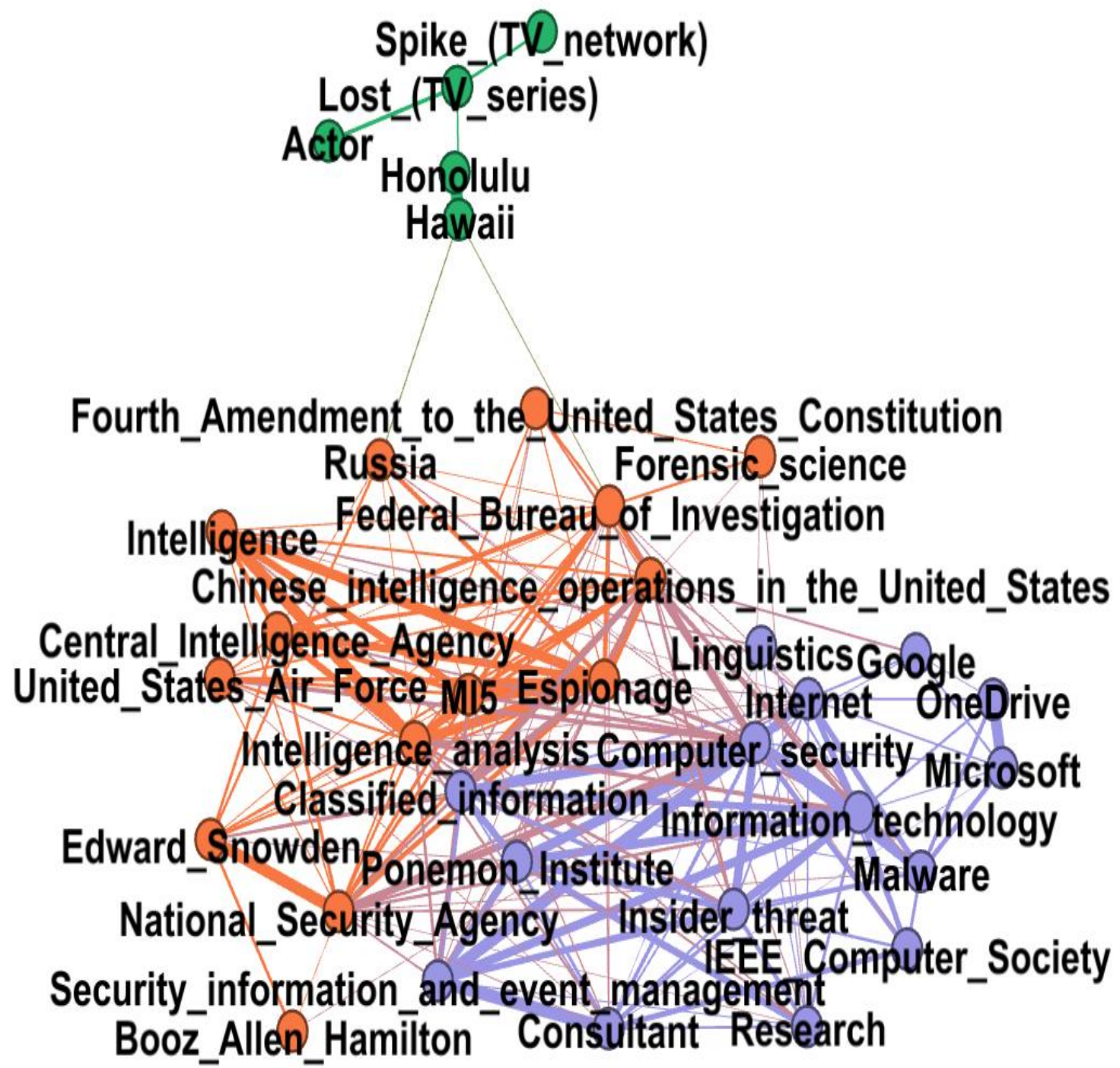
EDWARD SNOWDEN & DEFENDING AGAINST THE INSIDER THREAT



<https://www.youtube.com/watch?v=kQVLoNmYtKA>



Steven Bay talks about Edward Snowden



EDWARD SNOWDEN & DEFENDING AGAINST THE INSIDER THREAT

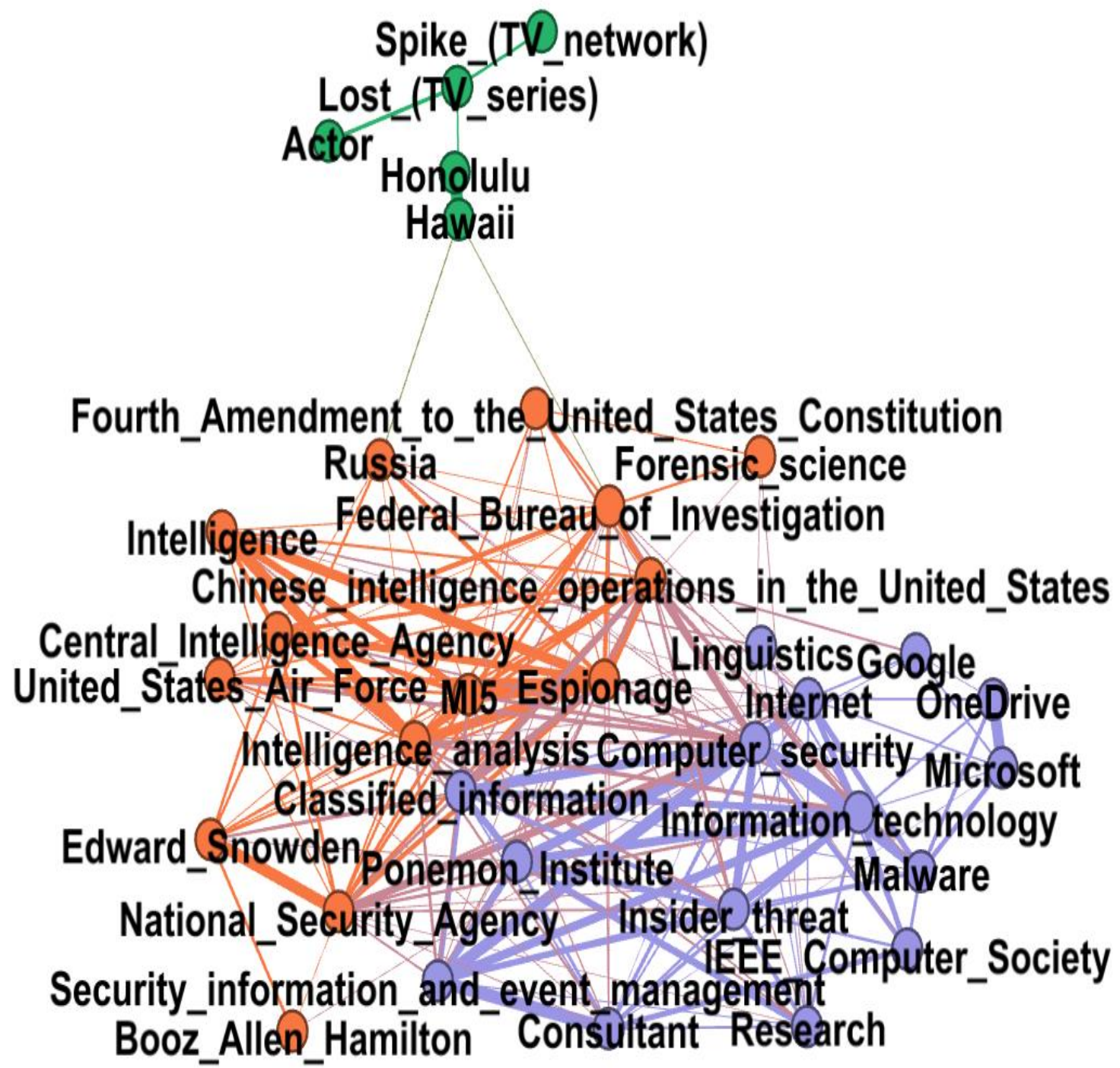


<https://www.youtube.com/watch?v=kQVLoNmYtKA>



Steven Bay talks about Edward Snowden

- Cyberwarfare
- Cryptography
- Espionage
- Government agencies










What is physics good for

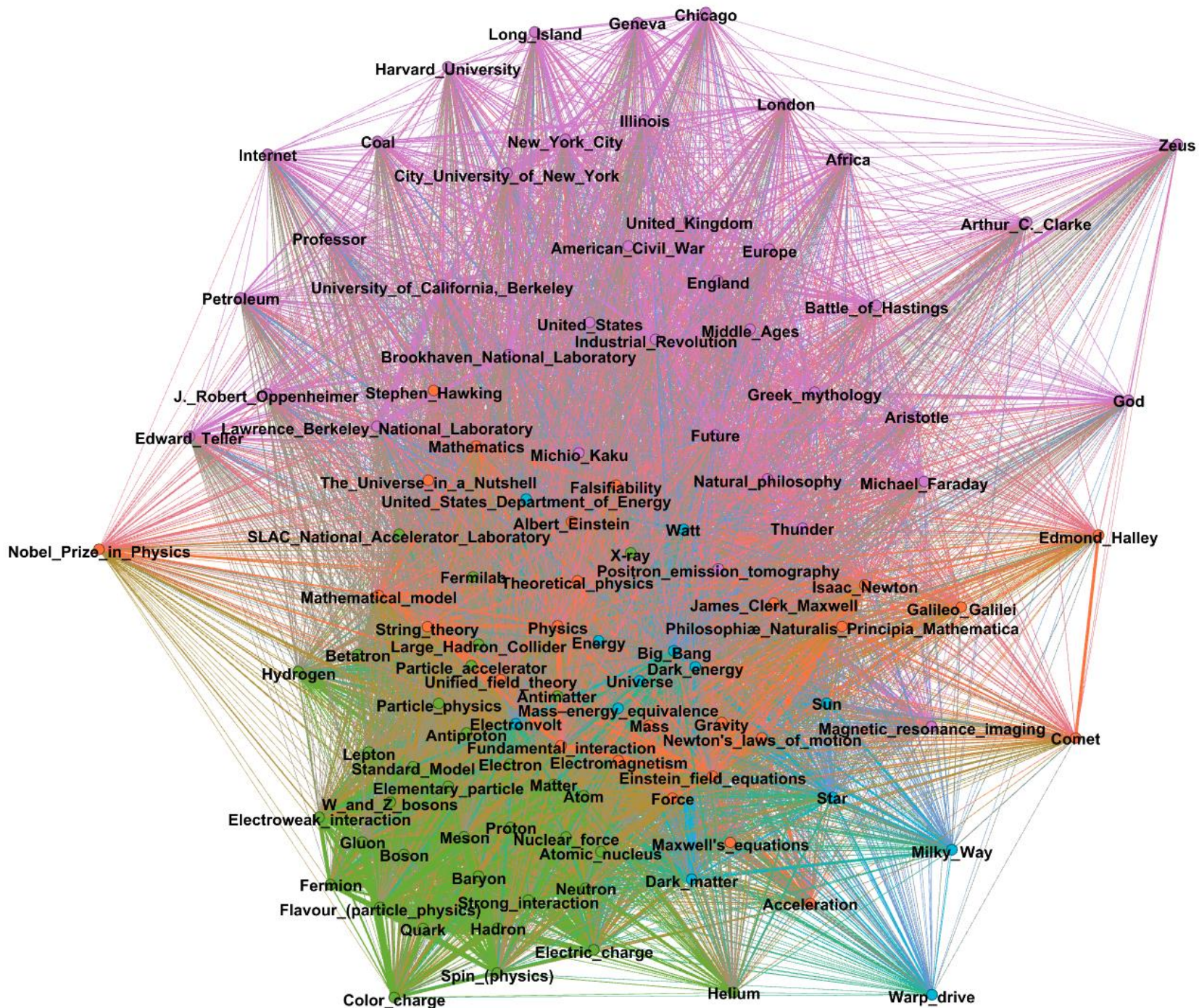
Better TV
Internet reception



<https://www.youtube.com/watch?v=0NbBjNiw4tk>

Michio Kaku The Universe in a nutshell

-  Physics
-  Elementary particles
-  Matter
-  Magnetism
-  Physical cosmology



Model evaluation over 1.7M options(!)

Precision@1 = 95.2%

Precision=71.9%

Summary

Robustness: Multi-channel analysis/fusion

Problem reduction: Knowledge Graph

Self maintained: Wikipedia dump periodic update

Interpretability: Topic Detection and Tracking



Thank you!

Multi-modal Topic inference from video

Oron Nir, Senior Data Scientist, oronnir@microsoft.com

Daniel Nurieli, Data Scientist



Q&A

Multi-modal Topic inference from video

Oron Nir, Senior Data Scientist, oronnir@microsoft.com

Daniel Nurieli, Data Scientist