

Big Brother liest, denkt und wirkt mit

Künstliche Intelligenz
zwischen gesellschaftlichem Nutzen und Schaden

Michael Strube

Heidelberg Institute for Theoretical Studies gGmbH
Heidelberg, Germany

From: [REDACTED]

To: <michael.strube@h-its.org>

CC: [REDACTED]

Subject: Interview in der NN vom 2.9.14

Date: Tue, 2 Sep 2014 08:38:40 +0200

Sehr geehrter Herr Doktor Strube,

könnten sie sich auch vorstellen, dass es Menschen gibt, die sich nicht vor dem Datenmissbrauch der NSA fürchten, weil sie sagen, das ist eine von einer Demokratie beauftragte und kontrollierte Institution und nicht der KGB. Mich stört auch der Datenzugriff von Google und Facebook nicht und auch nicht die Überwachung des öffentlichen Raums durch Kameras. Ich nutze täglich die neuen Möglichkeiten von Google Now und freue mich über die Hinweise, die ich dadurch automatisch bekomme. Es ist mir ein Vergnügen, Google Maps zu nutzen zusammen mit den anderen 900 Mio. Menschen. Ich kann mir gar nicht vorstellen, wie weit sie von der Zukunft entfernt sind. Ich habe mein ganzes Leben in der Computerbranche gearbeitet und Computeranwendungen innerhalb der Betriebsinformatik studiert ([REDACTED], Nürnberg). Also unterstellen sie mir bitte nicht Nichtwissen. Mir tun nur die Studenten leid, die sie so einseitig indoktrinieren.

MfG
[REDACTED]

Von Samsung Galaxy Note 10.1 gesendet[]



Nutzen

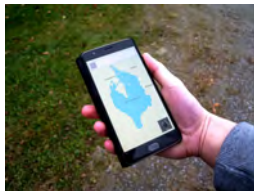
Schaden

Lokalisierung



Nutzen:

dank Google Maps und GPS nie wieder verlaufen, nie wieder verfahren;
Suchergebnisse besser, da lokalisiert

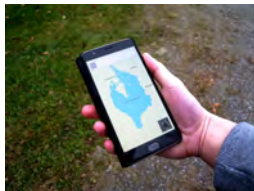


Lokalisierung



Nutzen:

dank Google Maps und GPS nie wieder verlaufen, nie wieder verfahren;
Suchergebnisse besser, da lokalisiert



Schaden:

Smartphone dient als elektronische
Fußfessel

Elektronische Kommunikation



Nutzen:

Email, Instant-Messaging, soziale Netzwerke, ... : umsonst, schnell, bequem, ...



Elektronische Kommunikation



Nutzen:

Email, Instant-Messaging, soziale Netzwerke, ... : umsonst, schnell, bequem, ...



Schaden:

Metadaten und alles Gesendete und Empfangene beim Provider gespeichert

Elektronische Assistenten (Siri, Alexa, ...)



Nutzen:

Smartes Haus; Bedienung des SmartTVs durch Sprache; Bedienung des Smartphones, ohne die Hände vom Lenkrad nehmen zu müssen; ...

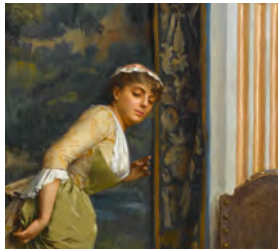


Elektronische Assistenten (Siri, Alexa, ...)



Nutzen:

Smartes Haus; Bedienung des SmartTVs durch Sprache; Bedienung des Smartphones, ohne die Hände vom Lenkrad nehmen zu müssen; ...



Schaden:

ständiges Abhören

Suchmaschinen



Nutzen:

"Information at your fingertips" wurde durch Google zur Realität; alles Wissen der Welt ist nur wenige Mausclicks entfernt



Suchmaschinen



Nutzen:

“Information at your fingertips” wurde durch Google zur Realität; alles Wissen der Welt ist nur wenige Mausclicks entfernt



| User identification | Date | Time | Search_url |
|---------------------------------------|--------------------|---------------------|---|
| ce010604048155087704275089f9eud | 25/Apr/2004 | 04:00:50 | Splatoon Artist. Harvesting - New Jersey - Ranking email@yabov. |
| 38094f346f51137587efba3041a4d35 | 25/Apr/2004 | 04:00:50 | Taijitsu |
| 6be9531e53099a0877732a1a970250a | 25/Apr/2004 | 04:00:54 | 1 personality AND gender AND education |
| 3010406d79032608042c3c77b795e4 | 25/Apr/2004 | 04:00:54 | |
| 25/Apr/2004 | 04:00:54 | dir paonsoni | |
| 8062acc4864e4570b091097694b301 | 25/Apr/2004 | 04:00:55 | Bowtie poems: |
| | 25/Apr/2004 | "Mark Twain" | |
| 3979f5665501380c181835d0c39426 | | 04:00:56 | gay poen |
| a9560248d10d47975ff4556921ca0b | 25/Apr/2004 | 04:00:58 | skin diagnosis: |
| 81347ca595321a15b16c00ba51678e3 | 25/Apr/2004 | 04:00:59 | Pink Floyd CD label cover items |
| 3c5c39912d709742e091ee0e64305464 | 25/Apr/2004 | 04:00:00 | from millard dougand |
| 8aaf120909e0d4f1e084e956ca83740d6 | 25/Apr/2004 | 04:00:00 | Youn Ji |
| 414134432bc18f978abce8c355147266 | 25/Apr/2004 | 04:00:00 | Capability Maturity Model VS. |
| 093488794a6a0891e263e3e8c7231ef | 25/Apr/2004 | 04:00:01 | ato chelonis paulo finisura |

Schaden:

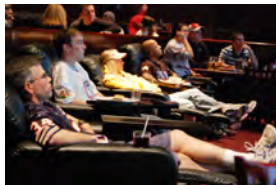
Searchlogs enthalten intimste Information über uns

Informationstechnologie



Nutzen:

Information und Kommunikation
kostengünstig, schnell, bequem, ...

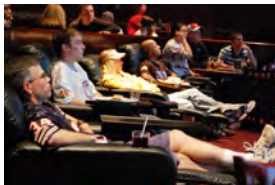


Informationstechnologie



Nutzen:

Information und Kommunikation
kostengünstig, schnell, bequem, ...



Schaden:

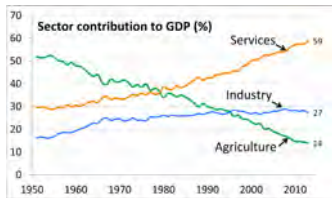
Google, Facebook & Co. sammeln alle Information über uns (Geschäftsmodell der Firmen); es gibt keine Anonymität mehr

Informationstechnologie



Nutzen:

allein Google trägt etwa 2-3% zum US-Bruttosozialprodukt bei: etwa 0.5% direkt durch Google, den Rest durch andere Firmen (Peter Norvig, Director of Research at Google, 2011)

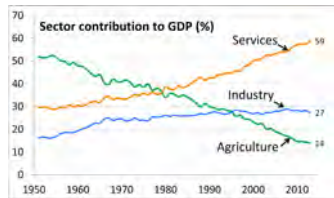


Informationstechnologie



Nutzen:

allein Google trägt etwa 2-3% zum US-Bruttosozialprodukt bei: etwa 0.5% direkt durch Google, den Rest durch andere Firmen (Peter Norvig, Director of Research at Google, 2011)



Schaden:

seit Snowden wissen wir, daß Information über Benutzer nicht bei den Firmen bleibt, sondern an NSA etc. weitergegeben wird

U.S. NEWS

How the NSA Could Get So Smart So Fast

Modern Computing Is Helping Companies and Governments Accurately Parse Vast Amounts of Data in a Matter of Minutes

By MICHAEL HIGGINS

Five years ago it would have been unimaginable for a government agency such as the National Security Agency to efficiently parse trillions of phone, text and online conversations for keywords that could have warned of an impending terrorist attack. Today, a set of new technologies make it relatively affordable and manageable for it to do so.

These technologies can store vastly different types of data in a single database, and can be processed rapidly using inexpensive hardware, without an analyst having to formulate a hypothesis.

It is unclear exactly what type of computing the NSA is using in its data-center facilities around the U.S., or in a \$1.2 billion facility in Utah that will open this fall.

But broadly speaking the technology can be broken down into three categories:

Database systems

Traditional databases, usually written in a language known as SQL (pronounced sequel), store data in tables, columns and rows

TMI?

Monthly global consumer Web, email and data traffic

15 thousand petabytes



EQUIVALENT TO:

The content of 517,300 Libraries of Congress or 156 billion hours of high-quality video.

Sources: Data Systems: EMC
The Wall Street Journal

but are limited when it comes to storing strings of words such as those found in an email or text message. They are also unable to handle pictures or video.

New types of databases that emerged beginning in late 2008, such as MongoDB, Cassandra and

Simple DB, don't have these limitations, and allow analysts to create queries against all these types of data.

Such databases, known collectively as NoSQL (for "not only SQL"), can make a huge difference to companies analyzing very large data sets, even if they are fairly conventional. For example, analysts at risk consultancy Verisk Analytics Inc. run various models and analytics against billions of customer records in order to help identify fraudulent insurance claims.

Perry Rotella, vice president and chief information officer at Verisk, said using a traditional DB2 database from International Business Machines Corp., "would be a six-hour job" that had to run overnight. Analysts would pore over the results and generate new queries that would have to run again. He said it took weeks every time analysts needed to create a new statistical model. The company recently changed to a NoSQL database that allows analysts to run the same types of queries in 30 seconds.

For online businesses like photography marketplace Shutterstock Inc., which store a great variety of file types, it is difficult to imagine life without this technology. Shutterstock has a library of more than 24 million images and adds an additional 10,000 each day, each with associated data to help narrow search results.

Its databases also record what users do on the site—not just decisive actions such as which images they license, but also minute details such as where they place their cursor and how long they hover there.

Natural language processing

In the past analysts needed to have enough understanding of the data to form a hypothesis and then create complex queries to run against the database.

Recently developed programs known as natural language processing and machine learning rely on the computer programs themselves to find patterns and even elucidate the meaning of ambiguous words based on context.

With natural language pro-

cessing, "you could figure out whether a term like 'bomb' is being used to describe a Broadway play versus something a terrorist would use," said Tom Davenport, an expert on analytics and a visiting professor at Harvard Business School.

Hadoop

Until recently, complex computer programs needed to run on expensive hardware, such as enormous mainframe computers.

Today, an open-source software framework called Hadoop—which was developed at Yahoo Inc. with contributions from technology developed by Google Inc. and named after a child's toy elephant—allows queries to be split up by the program.

Different analytic tasks are distributed among seas of inexpensive servers, each of which solves a part of the puzzle, before reassembling the queries when the work is completed. "It's really cheap and really fast," Mr. Davenport said.

The ability to distribute complex queries to a large number of inexpensive computers helps peo-

ple get very quick responses to complicated questions with a large number of variables.

For example, online automotive market Edmunds.com Inc. can help auto dealers predict how long a given car will remain on their lots by comparing car makes, models and other features against the number of days inventory cars at that price point averaged on a lot in a given dealer's region.

The predictions help minimize the number days a car remains unsold—"one of the most important sales metrics for dealers," said Philip Potliel, Edmunds.com's chief information officer.

—Steven Rosenbush
and Clint Boulton
contributed to this article.

Continuing Coverage

Scan this code for continuing coverage of the NSA surveillance controversy or visit WSJ.com/US.





Five years ago it would have been unimaginable for a government agency such as the National Security Agency to efficiently parse millions of phone, text and online conversations for keywords that could have warned of an impending terrorist attack. Today, a set of new technologies make it relatively affordable and manageable for it to do so.

U.S. NEWS

How the NSA Could Get So Smart So Fast

Modern Computing Is Helping Companies and Governments Accurately Parse Vast Amounts of Data in a Matter of Minutes

By MICHAEL HIGGINS

Five years ago it would have been unimaginable for a government agency such as the National Security Agency to efficiently parse trillions of phone, text and online conversations for keywords that could have warned of an impending terrorist attack. Today, a set of new technologies make it relatively affordable and manageable for it to do so.

These technologies can store vastly different types of data in a single database, and can be processed rapidly using inexpensive hardware, without an analyst having to formulate a hypothesis.

It is unclear exactly what type of computing the NSA is using in its data-center facilities around the U.S., or in a \$1.2 billion facility in Utah that will open this fall.

But broadly speaking the technology can be broken down into three categories:

Database systems

Traditional databases, usually written in a language known as SQL (pronounced sequel), store data in tables, columns and rows

TMI?

Monthly global consumer Web, email and data traffic

15 thousand petabytes



EQUIVALENT TO:

The content of 517,300 Libraries of Congress or 156 billion hours of high-quality video.

Sources: Data Systems: EMC
The Wall Street Journal

but are limited when it comes to storing strings of words such as those found in an email or text message. They are also unable to handle pictures or video.

New types of databases that emerged beginning in late 2008, such as MongoDB, Cassandra and

Simple DB, don't have these limitations, and allow analysts to create queries against all these types of data.

Such databases, known collectively as NoSQL (for "not only SQL"), can make a huge difference to companies analyzing very large data sets, even if they are fairly conventional. For example, analysts at risk consultancy Verisk Analytics Inc. run various models and analytics against billions of customer records in order to help identify fraudulent insurance claims.

Perry Rotella, vice president and chief information officer at Verisk, said using a traditional DB2 database from International Business Machines Corp., "would be a six-hour job" that had to run overnight. Analysts would pore over the results and generate new queries that would have to run again. He said it took weeks every time analysts needed to create a new statistical model. The company recently changed to a NoSQL database that allows analysts to run the same types of queries in 30 seconds.

For online businesses like photography marketplace Shutterstock Inc., which store a great variety of file types, it is difficult to imagine life without this technology. Shutterstock has a library of more than 24 million images and adds an additional 10,000 each day, each with associated data to help narrow search results.

Its databases also record what users do on the site—not just decisive actions such as which images they license, but also minute details such as where they place their cursor and how long they hover there.

Natural language processing

In the past analysts needed to have enough understanding of the data to form a hypothesis and then create complex queries to run against the database.

Recently developed programs known as natural language processing and machine learning rely on the computer programs themselves to find patterns and even elucidate the meaning of ambiguous words based on context.

With natural language pro-

cessing, "you could figure out whether a term like 'bomb' is being used to describe a Broadway play versus something a terrorist would use," said Tom Davenport, an expert on analytics and a visiting professor at Harvard Business School.

Hadoop

Until recently, complex computer programs needed to run on expensive hardware, such as enormous mainframe computers.

Today, an open-source software framework called Hadoop—which was developed at Yahoo Inc. with contributions from technology developed by Google Inc. and named after a child's toy elephant—allows queries to be split up by the program.

Different analytic tasks are distributed among seas of inexpensive servers, each of which solves a part of the puzzle, before reassembling the queries when the work is completed. "It's really cheap and really fast," Mr. Davenport said.

The ability to distribute complex queries to a large number of inexpensive computers helps peo-

ple get very quick responses to complicated questions with a large number of variables.

For example, online automotive market Edmunds.com Inc. can help auto dealers predict how long a given car will remain on their lots by comparing car makes, models and other features against the number of days inventory cars at that price point averaged on a lot in a given dealer's region.

The predictions help minimize the number days a car remains unsold—"one of the most important sales metrics for dealers," said Philip Potliel, Edmunds.com's chief information officer.

—Steven Rosenbush
and Clint Boulton
contributed to this article.

Continuing Coverage

Scan this code for continuing coverage of the NSA surveillance controversy or visit WSJ.com/US.





Natural language processing

In the past analysts needed to have enough understanding of the data to form a hypothesis and then create complex queries to run against the database.

Recently developed programs known as natural language processing and machine learning rely on the computer programs themselves to find patterns and even elucidate the meaning of ambiguous words based on context.

With natural language pro-

cessing, “you could figure out whether a term like ‘bomb’ is being used to describe a Broadway play versus something a terrorist would use,” said Tom Davenport, an expert on analytics and a visiting professor at Harvard Business School.

Workshop ...



... Language Analysis in Social Media at NAACL'13

Topical Positioning: A New Method for Predicting Opinion Changes in Conversation

Ching-Sheng Lin¹, Samira Shaikh¹, Jennifer Stromer-Galley^{1,2},
Jennifer Crowley¹, Tomek Strzalkowski^{1,3}, Veena Ravishankar¹

¹State University of New York - University at Albany, NY 12222 USA

²Syracuse University

³Polish Academy of Sciences

In this paper, we described an automated approach to detect participant's Topical Positioning and capture the opinion shifts by Topical Position Maps. This work is still in progress and we intend to process more genres of data, including Twitter and online chat, to confirm effects seen in the data we currently have. . . .

Topical Positioning: A New Method for Predicting Opinion Changes in Conversation

Ching-Sheng Lin¹, Samira Shaikh¹, Jennifer Stromer-Galley^{1,2},
Jennifer Crowley¹, Tomek Strzalkowski^{1,3}, Veena Ravishankar¹

¹State University of New York - University at Albany, NY 12222 USA

²Syracuse University

³Polish Academy of Sciences

In this paper, we described an automated approach to detect participant's Topical Positioning and capture the opinion shifts by Topical Position Maps. This work is still in progress and we intend to process more genres of data, including Twitter and online chat, to confirm effects seen in the data we currently have. . . . Another avenue of future work is to apply proposed model on virtual chatroom agent to guide the discussion and change participants' attitudes.

Überblick



Big Brother liest mit: Computerlinguistik und Wissenschaftsförderung

Big Brother denkt mit: Maschinelles Lernen und Vorurteile

Big Brother wirkt mit: Künstliche Intelligenz und Wir

Unsere Verantwortung als Wissenschaftler

Überblick



Big Brother liest mit: Computerlinguistik und Wissenschaftsförderung

Big Brother denkt mit: Maschinelles Lernen und Vorurteile

Big Brother wirkt mit: Künstliche Intelligenz und Wir

Unsere Verantwortung als Wissenschaftler

Computerlinguistik?





Wie kann ich einen Reifen wechseln?



Sign in

[All](#) Videos Shopping Images Maps More Settings Tools

About 548.000 results (0,68 seconds)

Reifenwechsel: So klappt der Räderwechsel selbst - autob...www.autobild.de/artikel/reifenwechsel-2288727.html • Translate this page

Mar 14, 2016 - Zweimal im Jahr steht für die meisten

Autofahrer der **Reifenwechsel** auf ... Denn ich habebeschlossen: Das **kann** ich selbst, investiere 20 Euro ...**Fachgerecht Räder & Reifen wechseln: So geht's! | Reifen...**<https://www.reifen.de/static/de/tipps.../reifentipps/reifen-wechseln> • Translate this pageWer Räder oder **Reifen wechseln** möchte, **kann** viel

falsch machen. Die folgende Bedienungsanleitung gibt

Tipps für den richtigen **Reifenwechsel**.**Reifenwechsel-Anleitung – ausführlich Schritt für Schritt - ...**www.pkw.de/ratgeber/wartung-pflege/reifenwechsel-anleitung • Translate this pageOct 15, 2015 - **Radwechseln kann** jeder. Das istunbestreitbar, jedoch **kann** man auch hierbei einige Fehlerbegehen. Damit du beim nächsten **Reifenwechsel** ...**ADAC Tipps - Reifen - Der Radwechsel Schritt für Schritt**<https://www.adac.de/infotestrat/reifen/profil-kauf-und.../radwechseln...> • Translate this pageDer Alptraum für jeden Autofahrer – urplötzlich der **Reifen**macht schlapp. **Wechseln** ist also angesagt. Gut, wenn

man vorher mal geübt hat – diese Anleitung ...

Reifen wechseln - aber richtig - Tippscoutwww.tippscout.de • Auto • Translate this pageNov 15, 2010 - Wer die **Reifen** selbst **wechseln** will,



How do I change a tire?



Sign in

All Videos Shopping Images Maps More Settings Tools

About 106.000.000 results (0,88 seconds)

How to Change Tires

1. Find a Safe Location. As soon as you realize you have a flat tire, do not abruptly brake or turn. ...
2. Turn on Your Hazard Lights. ...
3. Apply the Parking Brake. ...
4. Apply Wheel Wedges. ...
5. Remove the Hubcap or Wheel Cover. ...
6. Loosen the lug nuts. ...
7. Place the Jack Under the Vehicle. ...
8. Raise the Vehicle With the Jack.

More items...



[How to Change a Flat Tire | Bridgestone Tires
www.bridgestonetire.com/tread-and-trend/drivers-ed/how-to-change-a-flat-tire](http://www.bridgestonetire.com/tread-and-trend/drivers-ed/how-to-change-a-flat-tire)

About this result Feedback

People also ask

Which way do wheel nuts come off?

How do you change a TYRE?

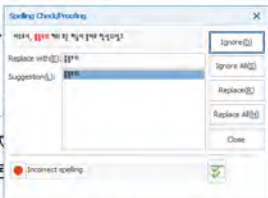
What do you do when you have a flat tire?

What tools do you need to change a tire?

표

심사요지:

이 글은 알렉산더 폰 훔볼트 "1부에 대한 서사이론적 의 관심을 끌게 되었으며, 스티븐 지식인 혹은 이 연구에 매우 비중있게 다루어지기에 필수적인 부분이라고 본



및 그가 생애에 걸친 연구를 집약한 "코
알렉산더 폰 훔볼트는 독일에서도 최근에
되지 않은 저술가이지만, 거의 마지막
천재"라고 할 수 있으며 세계시민으로서
대한 개괄적 소개를 하는 것은 이 논문의

1. 문제: 학술논문에서 어울리지 않는 어법이나 너무나 자주 반복되어 사용된 표현 "마지막 보편적 같은 말은 직접 본문에 붉은 색으로 표시하였으므로, 검토하시고 수정을 요합니다.

2. 각주1) 망각의 이유로서, **훔볼트의** 거의 모든 저술이 불어로 작성되었고 불란서에서 출판된 것 일에서 훔볼트 연구를 방해한 가장 큰 이유에 속한다는 지적도 필요하다.

My First Car

For years I have been driving an old used car with a lot of mileage, and I hate it. It gets me where I need to go, but I'm tired of fixing leaks and broken parts all the time. Its annoying every times I need to take it to the mechanic. Even when they takes care of everything, I know I'll just end up going back there in a few weeks. I have finally decided that I am not going to do it anymore. I have decided to buy a new car! Unfortunately, I have a problem, I have no idea what car to get. Do I want something big? Do I want something stylish? Something economical? I have so many choices that I don't even know where to begin.

I am not sure if I will be able to make a decision on my own. I don't have a lot of money, either, so I probably don't have many options. **After I did some research,** I knew that I would need something reliable. Eventually, I went to a local dealership to check out some options. I talked to the saleswoman and listened at she carefully explained her and professionalism were really impressive. She had given me helpful suggestions and showed myself some safe, affordable options. After a long discussion I finally decided which one I would buy.

did → had done

Incorrect verb tense

It appears that the verb *did* may need to be in the past perfect form. Consider changing the verb form.

✓ MORE

✗ IGNORE

My First Car

For years I have been driving an old used car with a lot of mileage, and I hate it. It gets me where I need to go, but I'm tired of fixing leaks and broken parts all the time. Its annoying every times I need to take it to mechanic. Even when they takes care of everything, I know I'll just end up going back there in a few weeks. I have finally decided that I am not going to do it anymore. I have decided to buy a new car. I have a problem, I have no idea what car to get. Do I want something big? Do I want something stylish? Something economical? I have many **choices** that I don't even know where to begin.

I am not sure if I will be able to make a decision on my own. I don't have a lot of money, either, so I probably don't have many options. After I did some research, I knew that I would need some help. Eventually, I went to a local dealership to check out some cars. I talked to the saleswoman and listened at she carefully. Her advice and professionalism were really impressive. She had some helpful suggestions and showed myself some safe, affordable options. After a long discussion I finally decided which one I would buy.

Synonyms for *choices*

1. That which is chosen
favorites options alternatives picks selections
2. The act or power of choosing
options opportunities decisions voices alternatives
3. A variety from which to choose
supplies varieties funds stocks selections

CLOSE

My First Car

For years I have been driving an old used car with a lot of problems. I hate it. It gets me where I need to go, but I'm tired of the broken parts all the time. Its annoying **every times** I go to the mechanic. Even when they takes care of everything, I have to go up going back there in a few weeks. I have finally decided I'm not going to do it anymore. I have decided to buy a new car. I have a problem. I have no idea what car to get. Do I want something big? Do I want something stylish? Something economical? There are many choices that I don't even know where to begin.

I am not sure if I will be able to make a decision on my own. I don't have a lot of money, either, so I probably don't have a lot of choices. After I did some research, I knew that I would need a car that was reliable. Eventually, I went to a local dealership to check out some cars. I talked to the saleswoman and listened at she cared for her customers. Her care and professionalism were really impressive. She had a lot of helpful suggestions and showed myself some safe, affordable options. After a long discussion I finally decided which one I wanted.

every times → every time

The quantifier *every* may not be used properly in this sentence. Consider changing the noun *times* to the singular, or using a different adjective with the plural noun.

Each and *every* are used as singular quantifiers to singular countable nouns, not plural nouns. *Each* is referring to an individual as one, while *every* is referring to a group lumped together as one. Make sure the correct verb form is used if the noun is the subject.

Incorrect: *Every customers are important.*

Correct: *Every customer is important.*

Correct: *All customers are important.*

With the quantifier *every*, use the singular noun and verb. With a plural noun and verb, choose a plural quantifier.

Incorrect: *He carried one box in each hands.*

Correct: *He carried one box in each hand.*

Correct: *He carried boxes in both hands.*





[http://www.medicalexpo.com/prod/infinitt-north-america/
product-82348-696539.html](http://www.medicalexpo.com/prod/infinitt-north-america/product-82348-696539.html)

WU
WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS





<http://www.museumsyndicate.com/item.php?item=82429>

U.S. Army tests nine-language translation device to support African operations

Alice MacGregor Tue 11 Aug 2015 11:51am

SUBSCRIBE TO OUR NEWSLETTER



h 5 s 2 d j 9 f 3 a 19 SHARES

The U.S. Army is developing a hand-held translation device designed to support military communications in Africa, where several French accents and dialects are spoken.

In a [press release](#) last week, the army said that understanding foreign languages is becoming critical for soldiers. "If we can get

Latest posts

Google mishandled NHS patient data, investigation warns

Australia \$36 billion broadband slower than Kenya

T-Mobile to build new data centre in

Most shared

Adobe develops AI-driven approach that could end the age of the 'green screen' in movies and VR

5K Total Shares

Research: Microsoft

Google

Google Übersetzer



Anmelden

Sofortübersetzung deaktivieren

Englisch Deutsch Chinesisch Sprache er...

Deutsch Englisch Französisch

Übersetzen

嘉義市 (臺語:Ka-gī-çhhī, 鄒語:maibayu), 是中華民國臺灣省 x 的市, 位於臺灣本島西南部的都市, 地處嘉南平原北端, 全市在北回歸線以北 (但極靠近北回歸線 [註 1]), 嘉義都會區的中心城市。全境除東部屬丘陵地帶外, 其餘為平原, 地勢由東向西緩降。外圍由嘉義縣環繞, 古名「諸羅山」, 係源自平埔族原住民洪雅族社名之譯音, 又名「桃城」, 因清治時期興築之古城形如桃而得名 [2]。

嘉義市是阿里山森林鐵路的起點[3]。過往曾因林業的發展而繁榮。由於發展歷史甚早, 市區內留有不少古蹟, 形成重要觀光資源。地方特產方面, 則以方塊酥[4]、阿里山羊羹、雪花餅、小饅頭、甘藷、鳳梨酥、火雞肉飯[5]最為人所知。文化活動上, 日治時期以繪畫最為著名, 曾有「畫都」之美譽, 代表人物有陳澄波等人; 近年來則以國際管樂節最為興盛。在政治上, 市長長年由源於黨外運動的「許家班」掌握[6]:27, 故有「民主聖地」之稱[6]:75。唯隨著中國國民黨籍黃敏惠的當選, 政治勢力亦有所轉變, 升格為省轄市後, 一至八屆市長 (正任) 皆由女性擔任, 在臺灣政壇相當特殊 [7]。直到第九屆市長由民主進步黨籍涂醒哲當選[8], 成為嘉義市升格省轄市後首位男性民選市長 (首位男性市長為接替許世賢擔任代理市長的省府委員江慶林)。

 拼

536/5000

jiāyì shì (tāi yǔ:Ka-gī-çhhī, zōu yǔ:Maibayu), shì

The city of Taiwan Province, located in the southwestern part of the island of Taiwan, is located at the northern end of the Jiaonan Plain. The city is located north of the Tropic of Cancer in the north (only in the north of the Tropic of Capricorn) Close to the Tropic of Cancer [Note 1]), the central city of Chiayi Metropolitan area. In addition to the eastern part of the hilly area, the rest of the plains, the terrain from east to west slow down. Surrounded by the Chiayi County, the ancient name of "Zhu Luoshan", derived from the Pingpo indigenous people Hong Ya family name transliteration, also known as "Taocheng", due to the construction of the ancient city of Qing Dynasty shaped like a peach named [2].

Chiayi City is the starting point of the Alishan Forest Railway [3], in the past because of the development of forestry and prosperity. As the development of very early history, the city left a lot of monuments, the formation of important tourist resources. Local specialty, then the box crisp [4], Ali goat soup, snow cake, small steamed bread, sweet potato, pineapple cakes, turkey rice [5] the most known. Cultural activities, the Japanese rule to the most famous painting, there have been "painting are" reputation, representatives of Chen Chengbo and others; in recent years, the most

مصر للطيران قد تعود غدا الاربعاء رحلتها الى ليبيا

الطائرة 6 - 4 (اف ب) - اعلن مسؤول في شركة الخطوط المصرية للطيران اليوم الثلاثاء ان شركة "مصر للطيران" قد تسلك اعتبارا من يوم غد الاربعاء رحلتها الى ليبيا لتز فرار مجلس الامن الدولي تعليق الحظر المفروض على ليبيا.



Foreign Languages

MT

VOLUME ONE

NUMBER ONE

MARCH, 1954

MECHANICAL TRANSLATION

DEVOTED TO THE TRANSLATION OF LANGUAGES WITH THE AID OF MACHINES

Erwin Reifler 22
Studies in Mechanical Translation, No. 7,
The Mechanical Determination of the Con-
stituents of German Substantive Composita
Mimeographed, 27 pages, September 3, 1952

Detailed discussion of the contents of a
German-English "Capital Memory." Pro-
poses methods by which compound nouns
not entered in "Capital Memory" would be
split up into their constituents.

James W. Perry 23
Machine Techniques for Index Searching
and for Machine Translation
Mimeographed, 6 pages, September 22, 1952

Machine techniques are described for
searching in sequence over an entire in-
dex of recorded information by successive
pattern matching operations, and for the
selection of those patterns which satisfy
certain associative relations between va-
rious parts of the pattern. These techni-

James W. Perry 25
*Machine Translation of Russian Techni-
cal Literature.
Notes on Preliminary Experiments
Mimeographed, 16 pages and appendix,
about October 1952

Words of Russian texts were written on
separate slips of paper, drawn at random,
translated, restored to original order. The
crude word-for-word translation was sub-
mitted to various people for rewriting.
They were able to obtain an insight into
the subject matter, but various mis-
understandings crept in.

James W. Perry 26
* Machine Translation of Russian Techni-
cal Literature.
Notes on Exploitation of the Russian Gram-
mar
Mimeographed, 11 pages, about Nov. 1952

It is proposed that the longest stem entry
which matches the word be looked up first,
then the remaining ending be looked up in
a list selected by a key number which was
obtained from the stem entry.

MECHANICAL TRANSLATION

DEVOTED TO THE TRANSLATION OF LANGUAGES WITH THE AID OF MACHINES

VOLUME TWO, NUMBER THREE

DECEMBER, NINETEEN FIFTY FIVE

COPYRIGHT 1956 BY THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

News

ACLS COMMITTEE

A committee has been formed by the American Council of Learned Societies for the purpose of following progress in mechanical translation and the application of mechanical methods to linguistics in general. The members of the committee are Norman A. McQuown, chairman, Paul Garvin, Martin Joos, and W.F. Twaddell.

MEETING IN CLEVELAND

A number of people who have been active in the field of mechanical translation met in Cleveland January 18. The occasion was a Conference on the Practical Utilization of Recorded Knowledge held under the auspices of the School of Library Science, Western Reserve University. One of the evening meetings was devoted exclusively to

MT IN MOSCOW

We were glad to hear of mechanical translation activity in the Soviet Union. According to newspaper accounts, there is interest in MT by people associated with a large computer at the Institute of Exact Mechanics and Computing Technique of the Soviet Academy of Sciences, Moscow.

The reports indicate that mechanical translation routines have been devised for this computer. Sergei A. Lebedev, the institute's director, is quoted as saying, "The word-for-word translation thus obtained is edited by the machine with an eye to the grammar and syntax of the language."

See Pravda, December 4, 1955 and January 22, 1956, also The New York Times, December 11, 1955 and The Christian Science Monitor, March 21, 1956. See also abstract 64 in this issue.

MECHANICAL TRANSLATION

DEVOTED TO THE TRANSLATION OF LANGUAGES WITH THE AID OF MACHINES

VOLUME THREE, NUMBER TWO

NOVEMBER, NINETEEN FIFTY SIX

COPYRIGHT 1957 BY THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

News

UNIVERSITY OF WASHINGTON
(from the Univ. of Washington News Service)

Research to prepare a machine for the translation of Russian into English is being conducted at the University of Washington under a contract with the International Telemeter Corporation of Los Angeles. The device is being developed by the corporation under a contract with the Rome Air Development Center, Griffis Air Force Base, Rome, N.Y.

BABEL

Babel, an international journal of translation, published by the International Federation of Translators with the assistance of UNESCO, devoted its October, 1956, issue entirely to mechanical translation. Abstracts of the articles can be found in the bibliography section. The journal, a quarterly now in its second year, is devoted entirely to matters of interest to translators. Articles are included on literary

Die 1960er Jahre



- Gründung der *Association for Computational Linguistics (ACL)* 1962, erste *ACL* Konferenz 1963
- formale Sprachen, automatische Syntaxanalyse, Aspekte der maschinellen Übersetzung
- erste *COLING* Konferenz 1965

Die 1960er Jahre



- Gründung der *Association for Computational Linguistics (ACL)* 1962, erste *ACL* Konferenz 1963
- formale Sprachen, automatische Syntaxanalyse, Aspekte der maschinellen Übersetzung
- erste *COLING* Konferenz 1965

1966: KI Winter (folgend auf den *ALPAC* Report)

Die 1970er Jahre



- Arbeiten an maschineller Übersetzung werden abrupt beendet
- US-Förderung verlegt sich auf Firmen (IBM, BBN, GE, TI, ...) und von der Regierung speziell eingerichtete Forschungslabore (SRI, MITRE, ...)
- Forschung konzentriert sich auf formale Grammatiken, Durchbruch im Bereich von Parsing-Algorithmen (Earley algorithm, chart parsing, ...)
- NLP beginnt Methoden der Künstlichen Intelligenz zu verwenden
- auch: frühe statistische Arbeiten zur Spracherkennung

Die 1980er and 1990er Jahre



- das *US Department of Defense (DOD)* fördert NLP wieder
- seit 1987 organisiert das *DOD* die Reihe *Message Understanding Conference (MUC)* – trotz des Names sind es nicht Konferenzen sondern vom DOD organisierte Wettbewerbe (Grishman & Sundheim, 1996)

Die 1980er and 1990er Jahre



We have just completed the sixth in a series of Message Understanding Conferences, which have been organized by NRAD, the RDT&E division of the Naval Command, Control and Ocean Surveillance Center (formerly NOSC, the Naval Ocean Systems Center) with the support of DARPA, the Defense Advanced Research Projects Agency. This paper looks briefly at the history of these Conferences and then examines the considerations which led to the structure of MUC-6.¹

The Message Understanding Conferences were initiated by NOSC to assess and to foster research on the automated analysis of military messages containing textual information. Although called

Die 1980er and 1990er Jahre



TST1-MUC3-0080

BOGOTA, 3 APR 90 (INRAVISION TELEVISION CADENA 1) -- [REPORT] [JORGE ALONSO SIERRA VALENCIA] [TEXT] LIBERAL SENATOR FEDERICO ESTRADA VELEZ WAS KIDNAPPED ON 3 APRIL AT THE CORNER OF 60TH AND 48TH STREETS IN WESTERN MEDELLIN, ONLY 100 METERS FROM A METROPOLITAN POLICE CAI [IMMEDIATE ATTENTION CENTER]. THE ANTIOQUIA DEPARTMENT LIBERAL PARTY LEADER HAD LEFT HIS HOUSE WITHOUT ANY BODYGUARDS ONLY MINUTES EARLIER. AS HE WAITED FOR THE TRAFFIC LIGHT TO CHANGE, THREE HEAVILY ARMED MEN FORCED HIM TO GET OUT OF HIS CAR AND GET INTO A BLUE RENAULT.

HOURS LATER, THROUGH ANONYMOUS TELEPHONE CALLS TO THE METROPOLITAN POLICE AND TO THE MEDIA, THE EXTRADITABLES CLAIMED RESPONSIBILITY FOR THE KIDNAPPING. IN THE CALLS, THEY ANNOUNCED THAT THEY WILL RELEASE THE SENATOR WITH A NEW MESSAGE FOR THE NATIONAL GOVERNMENT.

LAST WEEK, FEDERICO ESTRADA VELEZ HAD REJECTED TALKS BETWEEN THE GOVERNMENT AND THE DRUG TRAFFICKERS.

Die 1980er and 1990er Jahre



- das *US Department of Defense (DOD)* fördert NLP wieder
- seit 1987 organisiert das *DOD* die Reihe *Message Understanding Conference (MUC)* – trotz des Names sind es nicht Konferenzen sondern vom DOD organisierte Wettbewerbe (Grishman & Sundheim, 1996)

- ähnlich: *TIPSTER* Konferenzen (Fokus auf Dokumenterkennung und Informationsextraktion) und *TREC* Konferenzen (Fokus auf Information Retrieval)
- Dialogsysteme: *DARPA Communicator*

Die 2000er Jahre



- *DOD* fördert multilinguale Systeme
 - *Automatic Content Extraction (ACE)* Wettbewerbe für Englisch and Arabisch
 - *OntoNotes* (Weischedel et al., 2011) fügt Chinesisch and Arabisch hinzu
- Renaissance (statistischer) maschineller Übersetzung
- das *World Wide Web* für jeden
- Google wächst und wird zum vielleicht wichtigsten NLP-Forschungsinstitut

Die 2010er Jahre



- noch mehr Forschung über maschinelle Übersetzung
- noch mehr Forschung über multi- und *cross-linguale* Anwendungen
 - unterstützt etwa durch die EU (EU FP7)
- Fokus auf gesprochene arabische Dialekte
- Dialogsysteme funktionieren und werden kommerziell eingesetzt
- Geburt und Wachstum des *Social Web*
- Konvergenz von NLP und der Analyse unstrukturierter Daten

Entity Linking



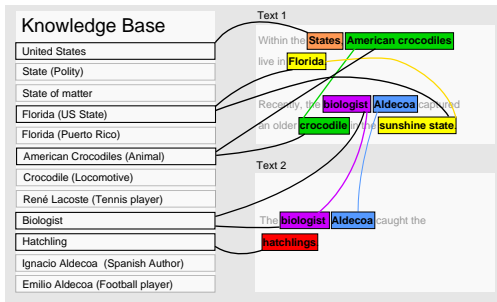
Entity Linking



Aufgabenbeschreibung (*Text Analysis Conference – TAC*):

Eingabe: Anfrage für Person, Organisation oder geo-politischer Entität (ein geographischer Ort mit einer Regierung); die Anfrage besteht aus einer Zeichenkette (Name) und einem Dokument, das die Zeichenkette enthält

Ausgabe: die *ID* des Eintrags in der Wissensdatenbank, auf den sich der Name bezieht; oder *NIL*, wenn es keinen solchen Eintrag gibt





[TAC 2017 Tracks](#)
[KBP Tracks](#)
[Cold Start KB/SF](#)
[EDL](#)
[Event](#)
[BeSt](#)
[Data](#)
[Schedule](#)
[Organizers](#)
[ADR Track](#)
[Call for Participation](#)
[Track Registration](#)
[Reporting Guidelines](#)
[TAC 2017 Workshop](#)



TAC Knowledge Base Population (KBP) 2017

Evaluation: February-November, 2017

Workshop: November 13-14, 2017

Conducted by:
U.S. National Institute of Standards and Technology (NIST)

With support from:
U.S. Department of Defense

Overview

The goal of TAC Knowledge Base Population (KBP) is to develop and evaluate technologies for populating knowledge bases (KBs) from unstructured text. KBP includes component tracks that develop specific components and capabilities for KBP, as well as an end-to-end KB construction task called "Cold Start", which builds a KB from scratch by integrating selected components as their technology matures. The capabilities required in the component tracks can be both "more" and "less" than what is exercised in the Cold Start KB task. The component tracks are "more" than Cold Start in the sense that each track may explore pilot tasks that are not immediately integrated into the Cold Start task; and they are "less" in the sense that integrating the components into a single KB requires additional coordination and reconciliation of mismatches between the various components so that the KB conforms with the KB schema (e.g., the KB cannot assert that an entity is the "PLACE" for an event if it also asserts that the entity is a "PERSON").

Standalone component tasks are offered in the following KBP tracks:

1. Entity Discovery and Linking (EDL): The main EDL task is to extract name and nominal mentions of specific individual person (PER), organization (ORG), geopolitical entity (GPE), location (LOC), and facility (FAC) entities mentioned in the evaluation document collection, and to link each mention to its KB node (either a node in the TAC reference KB, or a newly created NIL node if it doesn't have a corresponding KB entry). Additionally, the 2017 EDL track includes a name tracking pilot (which is not part of the 2017

Entity Linking



Wettbewerb bei der *Text Analysis Conference (TAC)* organisiert von NIST

For additional information on how to use the **tac-kbp@nist.gov** mailing list, send a message to tac-kbp-request@nist.gov, subject = help.

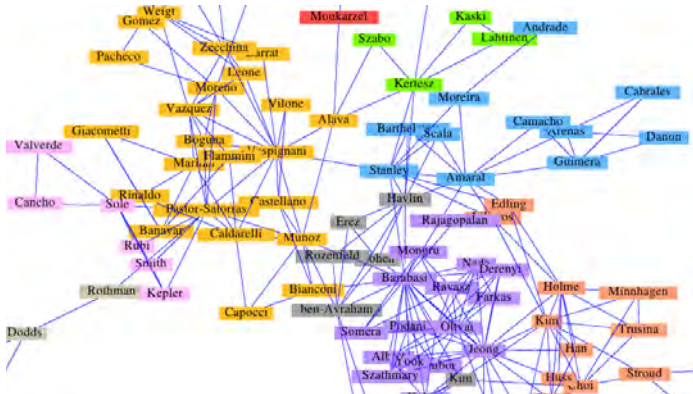
Organizing Committee

Hoa Trang Dang (U.S. National Institute of Standards and Technology)
Jason Duncan (MITRE)
Joe Ellis (Linguistic Data Consortium)
Marjorie Freedman (BBN Technologies)
Ralph Grishman (New York University)
Eduard Hovy (Carnegie Mellon University)
Heng Ji (Rensselaer Polytechnic Institute)
James Mayfield (Johns Hopkins University)
Teruko Mitamura (Carnegie Mellon University)
Boyan Onyshkevych (U.S. Department of Defense)
Shahzad Rajput (U.S. National Institute of Standards and Technology)
Owen Rambow (Columbia University)
Zhiyi Song (Linguistic Data Consortium)
Stephanie Strassel (Linguistic Data Consortium)

NIST is an agency of the
U.S. Department of Commerce

Source: <http://nlp.cs.rpi.edu/kbp/2014/>

Soziale Netzwerke ...



... können leicht aus Metadaten extrahiert werden

Soziale Netzwerke ...



... und unstrukturierte Daten?

Soziale Netzwerke ...



... und unstrukturierte Daten?

Agarwal, Apoorv et al. (2013): “Automatic extraction of social networks from literary text: A case study on *Alice in Wonderland*”. In Proc. IJCNLP-13

Aufgabe: analysiere Text, erkenne *Social Events* und extrahiere soziale Netzwerke **ohne Metadaten**

Soziale Netzwerke und Text



wer interessiert sich für diese Forschung:

Soziale Netzwerke und Text



wer interessiert sich für diese Forschung:

Acknowledgments

This paper is based upon work supported in part by the DARPA DEFT Program. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.



Information Innovation Office



PROGRAM MANAGER

Dr. Boyan Onyshkevych
boyan.onyshkevych@darpa.mil

PROGRAM IMAGES



DEEP EXPLORATION AND FILTERING OF TEXT (DEFT)

Department of Defense (DoD) operators and analysts collect and process copious amounts of data from a wide range of sources to create and assess plans and execute missions. However, depending on context, much of the information that could support DoD missions may be implicit rather than explicitly expressed. Having the capability to automatically extract operationally relevant information that is only referenced indirectly would greatly assist analysts in efficiently processing data.

Source: [http://www.darpa.mil/Our_Work/I2O/Programs/Deep_Exploration_and_Filtering_of_Text_\(DEFT\).aspx](http://www.darpa.mil/Our_Work/I2O/Programs/Deep_Exploration_and_Filtering_of_Text_(DEFT).aspx)

Deep Exploration and Filtering of Text (DEFT)

Dr. Boyan Onyshkevych



Department of Defense (DoD) operators and analysts collect and process copious amounts of data from a wide range of sources to create and assess plans and execute missions. However, depending on context, much of the information that could support DoD missions may be implicit rather than explicitly expressed. Having the capability to automatically extract operationally relevant information that is only referenced indirectly would greatly assist analysts in efficiently processing data.

Source: <http://www.darpa.mil/program/deep-exploration-and-filtering-of-text>

Autoren *Profiling* in *Microblogs*



Autoren *Profiling* in *Microblogs*



*Dirac was one of my boyhood heroes. I'm glad I met him once.
RT Paul Dirac image by artist Eric Handy: [http:...](http://...)*

Autoren *Profiling* in *Microblogs*



Forschungsfragen in Eisenstein et al. (2011)

- Geschlecht
- Alter (in Gruppen)
- Rasse/Ethnizität
- sozioökonomische Information (städtisch/ländlich, Familie(ngröße)?, Mieter/Hausbesitzer, Einkommen, ...)
- politische Einstellung
- ...

Autoren *Profiling* in *Microblogs*



wer interessiert sich für diese Arbeiten:

Autoren *Profiling* in *Microblogs*



wer interessiert sich für diese Arbeiten:

Using Conceptual Class Attributes to Characterize Social Media Users

Shane Bergsma and Benjamin Van Durme

Department of Computer Science and Human Language Technology Center of Excellence
Johns Hopkins University
Baltimore, MD 21218, USA

Autoren Profiling in Microblogs



wer interessiert sich für diese Arbeiten:

Human Language Technology Center of Excellence

- The Johns Hopkins University has been awarded a long-term, multimillion dollar contract to establish and operate a Human Language Technology Center of Excellence near Johns Hopkins' Homewood campus in Baltimore. The center's research will focus on advanced technology for automatically analyzing a wide range of speech, text and document image data in multiple languages.
- The Human Language Technology Center of Excellence is the result of a competitive solicitation from the U.S. Department of Defense. The contract, signed earlier this year, envisions a minimum of \$48.4 million in funding through 2015.
- "This will be the nation's premier language technology center, and it's likely to stimulate further expansion both at Johns Hopkins and at the University of Maryland in this important research field."

Inferring User Political Preferences from Streaming Communications

Svitlana Volkova,¹ Glen Coppersmith² and Benjamin Van Durme^{1,2}

¹Center for Language and Speech Processing,

²Human Language Technology Center of Excellence,

Johns Hopkins University, Baltimore, MD 21218

svitlana@jhu.edu, coppersmith@jhu.edu, vandurme@cs.jhu.edu

Identifying Opinion Subgroups in Arabic Online Discussions

Amjad Abu-Jbara

Department of EECS
University of Michigan
Ann Arbor, MI, USA
amjbara@umich.edu

Ben King

Department of EECS
University of Michigan
Ann Arbor, MI, USA
benking@umich.edu

Mona Diab

Department of Computer Science
George Washington University
Washington DC, USA
mtdiab@gwu.edu

Dragomir Radev

Department of EECS
University of Michigan
Ann Arbor, MI, USA
radev@umich.edu

Discriminating Gender on Twitter

John D. Burger and **John Henderson** and **George Kim** and **Guido Zarrella**

The MITRE Corporation

202 Burlington Road

Bedford, Massachusetts, USA 01730

{john, jhndrsn, gkim, jzarrella}@mitre.org

**Patient Experience in Online Support Forums:
Modeling Interpersonal Interactions and Medication Use**

Annie T. Chen

University of North Carolina, Chapel Hill
School of Information and Library Science
atchen@email.unc.edu

Weshalb bin ich besorgt?



Weshalb bin ich besorgt?



Your personal information will be used by [REDACTED] to provide better services.

[REDACTED] will collect your personal information as necessary to provide the Services in a seamless manner and to improve your user experience. [...] it may be necessary for [REDACTED] to verify your identity information, age, account, contact information and device information. Your personal information may be used to search, register or notify other Service users who are in your contact list or who may be your [REDACTED] friends. [...] In addition, your personal information may be used to perform statistical analysis for new customized services, provide event and advertising information, [...]

Weshalb bin ich besorgt?



██████████ will collect the following information to make our services better.

When you sign up for the Services or in the course of your use of the Services, ██████████ will collect your telephone number, contact information saved on your smartphone or other device (telephone numbers and names of third parties), [...] your status information, name, birth date, ID, photos (including meta-information), service usage history, email address, location information, IP address and cookies through the official website of ██████████, individual applications or programs.

██████████ will also collect shipping information (including the name, mobile phone number and address) to ship your purchases. In addition to the above information, ██████████ may collect your credit card information, [...] or other information required for payment processing when you use paid services.

[...]

Source: <http://www.██████████.com/policy/privacy>

Weshalb bin ich besorgt?



- Datenschutzrichtlinien der koreanischen Firma *KakaoTalk*
(<http://www.kakao.com/policy/privacy>)

Weshalb bin ich besorgt?



Source: <http://reboundmobile.com/6-best-strategies-to-crack-the-south-korean-mobile-game-market/>

Computerlinguistik und Förderung



(dies gilt ebenfalls für die meisten anderen Wissenschaftsgebiete)

- Einzelprojekte, Auftragsarbeit
- Forschungsprogramme
- Infrastruktur: Konferenzen, Wettbewerbe, Daten
- Institutionen



Pattern Analysis



مصر للطيران قد تعود غدا الاربعاء رحلتها الى ليبيا
 القاهرة 6 - 4 (اف ب) - اعلن مسؤولون في شركة
 الخطوط المصرية للطيران اليوم الثلاثاء ان شركة
 "مصر للطيران" قد تختلف اعتبارا من يوم غد
 الاربعاء رحلتها الى ليبيا لتر فرار مجلس الامن
 الدولي تعليق الحظر المفروض على ليبيا.

Foreign Languages

Überblick



Big Brother liest mit: Computerlinguistik und Wissenschaftsförderung

Big Brother denkt mit: Maschinelles Lernen und Vorurteile

Big Brother wirkt mit: Künstliche Intelligenz und Wir

Unsere Verantwortung als Wissenschaftler

Maschinelles Lernen



“Die Technologie des 21. Jahrhunderts!”

Google

Google says machine learning is the future. So I tried it myself

If deep learning will be as big as the internet, it's time for everyone to start looking closely at it



Most popular



Tottenham Hotspur v Manchester United: Premier League – live!



West Ham v Liverpool: Premier League – as it happened



'Accidental hero' halts ransomware attack and warns: this is not over



China's Xi lays out



Why Machine Learning Is the Future

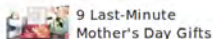
BY MICHAEL J. MILLER NOVEMBER 29, 2016 1 COMMENT

Chip and system makers take modern machine-learning concepts and apply them to supercomputers.

0 SHARES



At this month's **SC16 Supercomputing conference**, two trends stood out. The first is the appearance of Intel's latest Xeon Phi



9 Last-Minute Mother's Day Gifts



8 Creepy Sounds Recorded in Space by NASA



10 Tiny PCs of the 1980s



13 Mother's Day Tech Gifts for Under \$50



Windows 10 Fall Creators Update: 7 Things to Check Out



How a Retired Nurse Provides Her Small Vt. Town With Internet

SEE MORE >>

// DISCOVER...



The Best Laptops of 2017



26 Google Maps Tricks You Need to Try



19 Hidden Roku

Subscribe

PODCASTS MOBILE ARTIFICIAL INTELLIGENCE

In five years, machine learning will be a part of every doctor's job, Vic Gundotra says

Gundotra, a longtime Microsoft and Google executive, is now the CEO of AliveCor, a mobile health-tech startup.

BY ERIC JOHNSON | @HEYHEYESJ | DEC 5, 2016, 6:30AM EST

TWEET SHARE LINKEDIN



TRENDING



Watch this 'SNL' skit about an Amazon Echo Silver for senior citizens



INSIGHTS

USING MACHINE LEARNING TO PREDICT FUTURE TV RATINGS IN AN EVOLVING MEDIA LANDSCAPE

By Jingsong Cui, VP, and Scott Sereday, Manager, Data Science, Nielsen

MEDIA AND ENTERTAINMENT | 10-17-2016

[Twitter](#) [LinkedIn](#) [Share](#) [Submit](#)

Media companies and advertisers rely on TV ratings every day to measure the success of TV shows, verify that their audience size and composition are delivering against media-buy targets, and make-good in case the numbers come up short. From that point of view, TV ratings are metrics that measure the past, or at best the present, of TV viewing.

But media companies are also using ratings to predict the future. Ratings set expectations and affect programming decisions from one season to the next. They also help set advertising rates well in advance of when a campaign might actually air. In the U.S., for

RELATED NEWS

8.5 Generation Consoles Level Up the U.S. Video Game Industry"

Virtual Empathy: How 360-Degree Video Can Boost the Efforts of Non-Profits"

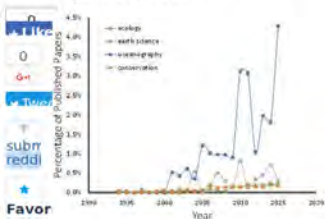
Connecting—and Protecting—the Big Investment: How Consumers Use Technology to Secure Their Homes"

One Bourbon, One Scotch, One Beer: the Adult Beverage Choices of Kentucky Derby Fans"

New Media Scores Big Returns for El Clásico

Machine Learning techniques and the future of Ecology and Earth Science Research

June 27, 2016



This plot shows the proportion of articles about machine learning in four natural science disciplines from 1994 to 2015, illustrating the slow penetration of the method in three of four disciplines. Credit: Anne E Thessen

Increasingly becoming a necessity in

Featured Last Popular
comment



China tests 'Lunar Palace' as it eyes moon mission • May 11, 2017 • 0



Three new sub-species of snow leopard discovered • May 11, 2017 • 0



Hydrogen bonds directly detected for the first time

The Future of Work: Machine Learning and the Professions

19 August 2016 by [Susannah Odell, Policy Adviser](#)

Machine learning - the form of artificial intelligence which allows machines to learn from data to make decisions or predictions - is already permeating our everyday experiences, as we've discussed previously [on this blog](#).



[Much has been written](#) about the impact of machine learning on employment, often focusing on the potential effects of automation. These effects play out differently in different sectors, and for workers with different skills levels.

To explore the impact of machine learning on work in specific sectors, on 13 July the Royal Society held a roundtable with representatives from the legal sector, healthcare, engineering, banking and accountancy. We discussed what machine learning might mean for these professions, considering its implications for business practices, skills, and the place of the professions in society.

This blog post highlights key points from the discussion.

The role, and tasks, of the professions

The professions have traditionally conveyed technical expertise, and judgement on technical matters, to consumers who would otherwise not have access to this expertise. However, technologies like

Welcome to In Verba, the [Royal Society's](#) science policy blog.

[Subscribe to RSS](#)

[Subscribe by email](#)

Please follow our [community guidelines](#).

Categories

- [Agriculture and food](#)
- [Diversity in science](#)
- [Economy and development](#)
- [Education and skills](#)
- [Energy, environment and climate](#)
- [Ethics and conduct](#)
- [Funding, governance and careers](#)
- [Health and wellbeing](#)
- [Industry and innovation](#)
- [International](#)
- [New and emerging technology](#)
- [None](#)
- [Open science and data](#)
- [Security and risk](#)

More blogs

[The Repository](#)

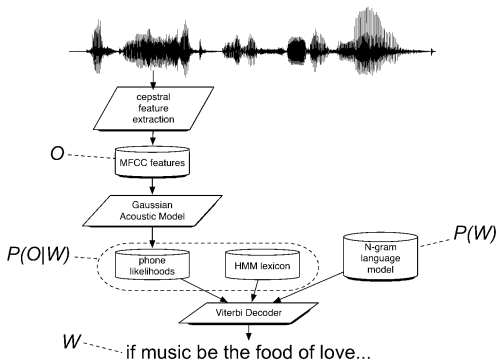
The Royal Society's history of science blog.

[Inside Science](#)

Maschinelles Lernen



- Sensoren zeichnen auf \Rightarrow Rohdaten
- Merkmalsextraktion
- statistischer Zusammenhang und Prädiktion



Neuronale Netzwerke

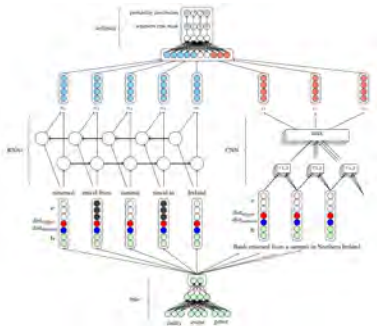


Renaissance neuronaler Netzwerke
mit vielen Ebenen (*Deep Learning*)

erkennen Merkmale selbst,
benötigen keine Domänenexperten

Ergebnisse oftmals besser als mit
traditionellen ML-Methoden

Epochenwechsel in der Künstlichen
Intelligenz



Neuronale Netzwerke



Machine Translation

1. *A Convolutional Encoder Model for Neural Machine Translation*
Authors: Jonas Gehring, Michael Auli, David Grangier and Yann Dauphin
2. *A Teacher-Student Framework for Zero-Resource Neural Machine Translation*
Authors: Yun Chen, Yang Liu, Yong Cheng and Victor O.K. Li
3. *Bandit Structured Prediction for Neural Sequence-to-Sequence Learning*
Authors: Julia Kreutzer, Artem Sokolov and Stefan Riezler
4. *Chunk-based Decoder for Neural Machine Translation*
Authors: Shonosuke Ishiwatari, Jingtao Yao, Shujie Liu, Mu Li, Ming Zhou, Naoki Yoshinaga, Masaru Kitsuregawa and Weijia Jia
5. *Deep Neural Machine Translation with Linear Associative Unit*
Authors: Mingxuan Wang, Zhengdong Lu, Jie Zhou and Qun Liu
6. *Doubly-Attentive Decoder for Multi-modal Neural Machine Translation*
Authors: Iacer Calixto, Qun Liu and Nick Campbell
7. *Improved Neural Machine Translation with a Syntax-Aware Encoder and Decoder*
Authors: Huadong Chen, Shujian Huang, David Chiang and Jiajun Chen
8. *Incorporating Word Reordering Knowledge into Attention-based Neural Machine Translation*
Authors: Jinchao Zhang, Mingxuan Wang, Qun Liu and Jie Zhou



Pattern Analysis

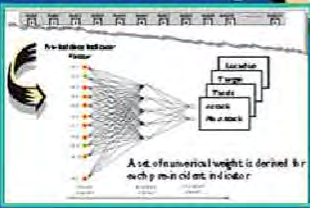
مصر للطيران قد تلوث عذا الاربعاء رحلتها الى ليبيا
 القاهرة 6 - 4 (اف ب) - اعلن مسؤولون في شركة
 الخطوط المصرية للطيران اليوم الثلاثاء ان شركة
 "مصر للطيران" قد تمكثف اعتبارا من يوم غد
 الاربعاء رحلتها الى ليبيا لتر فرار مجلس الامن
 الدولي تخليق الخطر المروض على ليبيا.

Foreign Languages



Collaboration

Predictive Modeling



Wo liegt das Problem?

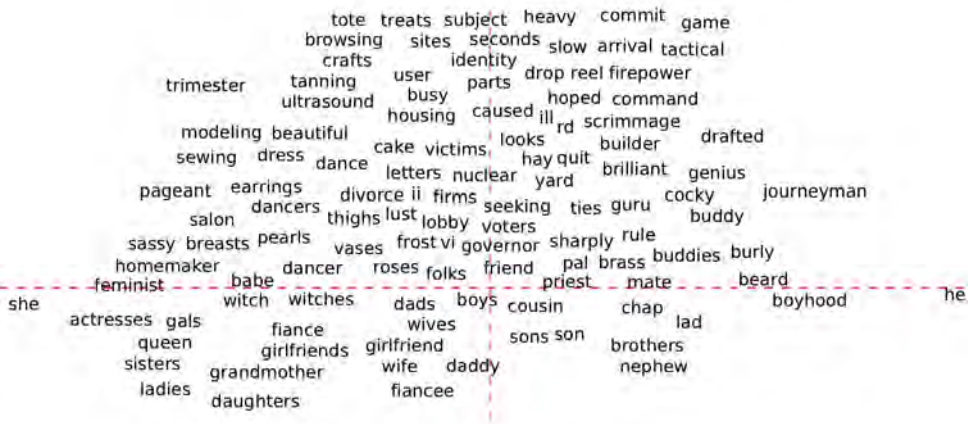


maschinelles Lernen
benötigt große Mengen an
Trainingsdaten

nimmt man Zeitungstexte,
trainiert man ein Modell, daß
*“nichts kann. Außer
Hochdeutsch.”*

Wo liegt das Problem?





Source: Bolukbasi et al. (2016)

$$\overrightarrow{\text{man}} - \overrightarrow{\text{woman}} \approx \overrightarrow{\text{king}} - \overrightarrow{\text{queen}}$$

$$\vec{\text{man}} - \vec{\text{woman}} \approx \vec{\text{king}} - \vec{\text{queen}}$$

$$\vec{\text{man}} - \vec{\text{woman}} \approx \vec{\text{computer programmer}} - \vec{\text{homemaker}}.$$

There have been hundreds or thousands of papers written about word embeddings and their applications, from Web search [27] to parsing Curriculum Vitae [16]. However, none of these papers have recognized how blatantly sexist the embeddings are and hence risk introducing biases of various types into real-world systems.

A word embedding that represent each word (or common phrase) w as a d -dimensional *word vector* $\vec{w} \in \mathbb{R}^d$. Word embeddings, trained only on word co-occurrence in text corpora, serve as a dictionary of sorts for computer programs that would like to use word meaning. First, words with similar semantic meanings tend to have vectors that are close together. Second, the vector differences between words in embeddings have been shown to represent relationships between words [32] [26]. For example given an analogy puzzle, “man is to king as woman is to x ” (denoted as *man:king :: woman:x*), simple arithmetic of the embedding vectors finds that $x=queen$ is the best answer because:

$$\vec{\text{man}} - \vec{\text{woman}} \approx \vec{\text{king}} - \vec{\text{queen}}$$

Similarly, $x=Japan$ is returned for *Paris:France :: Tokyo:x*. It is surprising that a simple vector arithmetic can simultaneously capture a variety of relationships. It has also excited practitioners because such a tool could be useful across applications involving natural language. Indeed, they are being studied and used in a variety of downstream applications (e.g., document ranking [27], sentiment analysis [18], and question retrieval [22]).

However, the embeddings also pinpoint sexism implicit in text. For instance, it is also the case that:

$$\vec{\text{man}} - \vec{\text{woman}} \approx \vec{\text{computer programmer}} - \vec{\text{homemaker}}.$$

Semantics derived automatically from language corpora contain human-like biases

Aylin Caliskan,^{1*} Joanna J. Bryson,^{1,2*} Arvind Narayanan^{1*}

Machine learning is a means to derive artificial intelligence by discovering patterns in existing data. Here, we show that applying machine learning to ordinary human language results in human-like semantic biases. We replicated a spectrum of known biases, as measured by the Implicit Association Test, using a widely used, purely statistical machine-learning model trained on a standard corpus of text from the World Wide Web. Our results indicate that text corpora contain recoverable and accurate imprints of our historic biases, whether morally neutral as toward insects or flowers, problematic as toward race or gender, or even simply veridical, reflecting the status quo distribution of gender with respect to careers or first names. Our methods hold promise for identifying and addressing sources of bias in culture, including technology.

Table 1. Summary of Word-Embedding Association Tests. We replicated eight well-known IAT findings using word embeddings (rows 1 to 3 and 6 to 10); we also help explain prejudiced human behavior concerning hiring in the same way (rows 4 and 5). Each result compares two sets of words from target concepts about which we are attempting to learn with two sets of attribute words. In each case, the first target is found compatible with the first attribute, and the second target with the second attribute. Throughout, we use word lists from the studies we seek to replicate. N , number of subjects; N_T , number of target words; N_A , number of attribute words. We report the effect sizes (d) and

P values (P , rounded up) to emphasize that the statistical and substantive significance of both sets of results is uniformly high; we do not imply that our numbers are directly comparable with those of human studies. For the online IATs (rows 6, 7, and 10), P values were not reported but are known to be below the significance threshold of 10^{-2} . Rows 1 to 8 are discussed in the text; for completeness, this table also includes the two other IATs for which we were able to find suitable word lists (rows 9 and 10). We found similar results with word2vec, another algorithm for creating word embeddings, trained on a different corpus, Google News (see the supplementary materials).

| Target words | Attribute words | Original finding | | | | Our finding | | | |
|--|----------------------------------|------------------|-----|------|----------------|---------------|---------------|------|-----------|
| | | Ref. | N | d | P | N_T | N_A | d | P |
| Flowers vs. insects | Pleasant vs. unpleasant | (5) | 32 | 1.35 | 10^{-8} | 25×2 | 25×2 | 1.50 | 10^{-7} |
| Instruments vs. weapons | Pleasant vs. unpleasant | (5) | 32 | 1.66 | 10^{-10} | 25×2 | 25×2 | 1.53 | 10^{-7} |
| European-American vs. African-American names | Pleasant vs. unpleasant | (5) | 26 | 1.17 | 10^{-5} | 32×2 | 25×2 | 1.41 | 10^{-8} |
| European-American vs. African-American names | Pleasant vs. unpleasant from (5) | (7) | | | Not applicable | 16×2 | 25×2 | 1.50 | 10^{-4} |
| European-American vs. African-American names | Pleasant vs. unpleasant from (9) | (7) | | | Not applicable | 16×2 | 8×2 | 1.28 | 10^{-3} |
| Male vs. female names | Career vs. family | (9) | 39k | 0.72 | $<10^{-2}$ | 8×2 | 8×2 | 1.81 | 10^{-3} |
| Math vs. arts | Male vs. female terms | (9) | 28k | 0.82 | $<10^{-2}$ | 8×2 | 8×2 | 1.06 | .018 |
| Science vs. arts | Male vs. female terms | (10) | 91 | 1.47 | 10^{-24} | 8×2 | 8×2 | 1.24 | 10^{-2} |
| Mental vs. physical disease | Temporary vs. permanent | (23) | 135 | 1.01 | 10^{-3} | 6×2 | 7×2 | 1.38 | 10^{-2} |
| Young vs. old people's names | Pleasant vs. unpleasant | (9) | 43k | 1.42 | $<10^{-2}$ | 8×2 | 8×2 | 1.21 | 10^{-2} |

Wo liegt das Problem?



maschinelles Lernen
benötigt große Mengen an
Trainingsdaten

nimmt man Zeitungstexte,
trainiert mein ein Modell, daß
“nichts kann. Außer
Hochdeutsch.”

nimmt man Texte aus dem
Web, dann reproduziert man
getreu alle Vorurteile, die im
Web liegen: “Stecken wir
Vorurteile rein, kommen
Vorurteile raus.”

Dunkle Algorithmen

Computer lernen neuerdings, den Sinn von Sprache und Texten
zu verstehen. Doch das birgt enorme Risiken.
Maschinen kopieren dadurch auch menschliche Schwächen
und verinnerlichen längst überholte Vorurteile

beachten den Tag mehr vor den Tieren als die
Vorurteile sind dokumentiert und bekann-
ten. Unsere Forschung wird einbe-
traucht, darüber müssen wir sprechen!
Je mehr Publikationen es in den folgen-
den Wochen hat, umso wahrscheinlicher wird
das sein. Er kommt nicht mehr vor, sondern
wie wichtig die Computer sein, schon waren,
und wieder psychologischen Hintergründen
wie aus dem Texten der Menschen herauszu-
lesen können. Einen Aufsatz zur Frage, wie
man sich über den Wert von Bewusstsein, las
er nicht mehr vor, wie immer Forscherinnen

Erkenntnis, insbesondere. Psychologica
kann nachweisen, dass das Programm:
Menschen mit dunkler Hautfarbe härter
bestrafen würde. „Da lag der Bias bereits
in den Daten“, sagt Hovy. Die Vorurteile
von Richtern aus früheren Zeiten hatten
sich nicht im Algorithmus eingemistet.

Neben der Dual-Use-Problematik sieht
Hovy solche mit Vorurteilen infizierten Da-
ten als eine der größten Gefahren in der
Computerlinguistik. Die meisten neuen Algori-
thmen lernen auf der Grundlage von al-
ten Trainingsdaten und reproduzieren ein-
fach alle Muster, ebenso dann Schwächen beifol-
gender Sprache erfullen. Einmalig ganz ein-
fach, sagt Margaret Mitchell von Google
Research: „Stecken wir Vorurteile ein,
bestehen Vorurteile raus.“ Diese werden ab-
hängig häufig gar nicht bemerkt werden.
„Wir haben keine mächtige Technologie“,
sagt Mitchell – und damit stehen sich
zwei Fragen, deren Klärung wird klar, weil
oben Kardinal: solche Algorithmen auf die
Gesellschaft haben können. „Treueren
in den Daten werden man hat es durch

Langen sind solche Phänomene nicht be-
achtet worden, weil man genau hinscha-
wen muss, erklärt Hanna Wallach von Miroso-
ft Research: „Es ist ein großer Unter-
schied zwischen einem Modell, das für alle
Benutzergruppen 95 Prozent genau
ist und einem, das nur für den 95 Prozent
genau ist für weiße Männer, aber nur 90 Pro-
zent Genauigkeit erreicht, wenn es um
Frauen oder Minderheiten geht.“ Hoch
bedenklich stellt es vor, Sprachen wie Tamil
oder andere Sprachen des indischen Sub-
kontinents, die teilweise 99 Millionen

Das Programm würde
Häftlinge mit dunkler
Hautfarbe härter bestrafen.
Es hätte das so von allen
Gerichtsurteilen gelernt

Es gibt Programme,
die Tweets daraufhin
untersuchen,
ob jemand selbst
oder depressiv ist

Überblick



Big Brother liest mit: Computerlinguistik und Wissenschaftsförderung

Big Brother denkt mit: Maschinelles Lernen und Vorurteile

Big Brother wirkt mit: Künstliche Intelligenz und Wir

Unsere Verantwortung als Wissenschaftler

Künstliche Intelligenz



Association for the Advancement of Artificial Intelligence (AAAI):

*AAAI advances the understanding of the mechanisms underlying **thought** and intelligent **behavior** and their **embodiment** in machines.*

Artificial Intelligence

Artificial Intelligence is Completely Transforming Modern Healthcare

Stanford EdTech/Flickr

IN BRIEF

AI in medicine is changing healthcare as we know it. The introduction of deep learning systems is only possible by powerful computing capabilities; capabilities that Nvidia has made possible with their graphic processors.

A NEW AGE OF HEALTHCARE

[Artificial intelligence](#) is slowly making its way into the realm of modern healthcare. Google's [DeepMind](#) is [revolutionizing eye care](#) in the United Kingdom, and [IBM's Watson](#) is tackling cancer diagnostics on par with human physicians. Both AI systems use deep learning, a concept loosely mirroring how our own brains work by having AI software analyze exorbitant amounts of data and uncover patterns — which is particularly applicable in diagnostics.

SHARE



WRITTEN BY

Neil C. Bhavsar
@Neilcogen



Abby Norman



April 2017

© 2017 Futurism Inc. All rights reserved.
Privacy Policy | Terms of Service



[http://www.dlr.de/dlr/en/desktopdefault.aspx/tabid-10081/
151_read-18222/#/gallery/23380](http://www.dlr.de/dlr/en/desktopdefault.aspx/tabid-10081/151_read-18222/#/gallery/23380)

TECH

TECH MOBILE SOCIAL MEDIA ENTERPRISE CYBERSECURITY TECH GUIDE

AI to become main way banks interact with customers within three years: Accenture



947 SHARE

Tuesday, 28 Mar 2017 | 12:00 PM ET

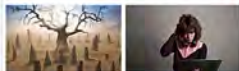
REUTERS



Kentoh | Getty Images

Artificial intelligence (AI) will become the primary way banks interact with their customers within the next three years, according to three quarters of

FROM THE WEB Links by Taboola



This game w... 5 Tricks To L... Throne: Free Babbel

MORE FROM CNBC by Taboola

Warren Buffett's one-word answer for why he hasn't purchased Amazon shares

This 30-year-old owes almost

THE WALL STREET JOURNAL

Subscribe Now Sign In

\$12 for 12 Weeks

Home World U.S. Politics Economy **Business** Tech Markets Opinion Arts Life

Real Estate



...



Atlanta Makes ...



Vodacom to Buy ...



>



CIO JOURNAL

Deutsche Bank Deploys Artificial Intelligence to Help Meet Demands of Regulatory Compliance

The bank is using AI to reduce costs and improve accuracy as it parses voice and video recordings of bankers at work.

By KIM S. NASH

Apr 18, 2017 5:09 pm ET

Deutsche Bank AG, facing a vast body of financial regulations, is turning to artificial intelligence as it sorts through volumes of voice and video recordings to make sure that the bank's professionals are complying with the

... The bank is using AI to reduce costs and improve accuracy as it parses voice and video recordings of bankers at work.

Recommended Videos

1. Trump's Post-Comey Approval Ratings Unchanged, But Fallout Continues



2. Trump's Tweets Intensify Interest in Russia Investigation



To Read the Full Story, Subscribe or Sign In

THE WALL STREET JOURNAL

NEWS

Investment banks to choose artificial intelligence over humans

Artificial intelligence could replace humans at the heart of the way capital markets work

By Samuel Agini

March 8, 2017 Updated: 12:41 p.m. GMT



Under constant pressure to slash costs and boost returns, investment banks are set to replace humans with artificial intelligence.

Stamford, Connecticut-based research firm Greenwich Associates thinks investment banks will seek to deploy artificial intelligence across research, sales, trading, and compliance.

Already used to detect fraud, deliver credit ratings and provide robotic financial advice in the retail sectors, artificial intelligence could be set to overhaul the way wholesale markets work.

Richard Johnson, vice president of market structure and technology at Greenwich, said: "In today's environment of continued cost pressure and low margins in many businesses,

FILTER BY TOPIC

Asset Management

Investment Banking

Hedge Funds

Trading

Private Equity

Regulation

Politics

Fintech

Lists

EDITOR'S PICKS

1. Minds over matter: City urged to change approach to mental health
2. Five simple ways to avoid getting Staley-ed
3. Barclays mustn't blow the chance to make an example of star man Staley
4. Frankfurt vows no gold-plated rules for Brexit refugees



Home » How will artificial intelligence affect the legal profession in the next decade?

Search this site... GO

Law and Technology:

How will artificial intelligence affect the legal profession in the next decade?



Illustration by Carl Wiens / www.i2iart.com

One of the hottest topics in the legal community of late is the expected impact on their profession of super-fast computers with the capacity to simulate human intelligence and decision-making, a.k.a. "artificial intelligence" (AI). Ever since Queen's Law Professor Hugh Lawford initiated the computerized database QUIC/LAW in the 1960s, AI has been creeping into the legal field. The latest example of this is Watson, an attempt by IBM to create a "cognitive computing system" for use in medicine, data analysis ... and law. *Queen's Law Reports* struck a panel of alumni with varying perspectives on the issue to discuss the influence AI is likely to have on the practice of law. **Professor Art Cockfield, Law'93**, was moderator, with participants **Janet Fuhrer, Law'85**, Canadian Bar Association President and law firm partner; **Jordan Furlong, Law'93**, strategic consultant; and **Jeff Fung, Law'08**, and **Addison Cameron-Huff, Law'12**, both lawyers and online entrepreneurs.

How do you think AI will be used in the legal profession over the next 10 years?

FURLONG: I think we're going to see the adoption of AI in the legal market, more broadly speaking, rather than in the legal profession for quite some time to come. Lawyers are sort of naturally disinclined, for cultural reasons, to disrupt the way they work and go about their jobs. Technology tends to generate that aversion.



EDUCATION DISRUPTED
How Google Took Over the...



Hacking Attack
Has Security Experts...



Uber Has a Union of Sorts,
but Faces...



BITS Farhad's and Mike's Week in...



China Hack
by Hacki Attack as

TECHNOLOGY

A.I. Is Doing Legal Work. But It Won't Replace Lawyers, Yet.

By STEVE LOHR MARCH 19, 2017



James Yoon, a partner at Wilson Sonsini Goodrich & Rosati in Palo Alto, Calif., says people are willing to pay for his experience. "What clients don't want to pay for is *new routine work.*"

Jason Times

9 ARTICLES

Get 40% off for one year on a Times

RELATED COVERAGE



A Lesson of Tesla Crashes?
Computer Vision Can't Do It All Yet SEPT. 19, 2016



The Promise of Artificial Intelligence Unfolds in Small Steps FEB. 28, 2016



Robots Will Take Jobs, but Not as Fast as Some Fear, New Report Says JAN. 12, 2017

SEE MY



How Artificial Intelligence Is Changing Talent Acquisition

By JJ-A Min November 11, 2016 TLNT

5

1.0K

194

680



AI for recruiting is on everyone's mind these days with a lot of talk on how it's going to transform recruiting. Artificial intelligence for recruiting is the next generation of software designed to improve or automate some part of the recruiting workflow.

How will AI and Machine Learning impact the future of recruitment?

10 February 2017 • Categories: [Artificial Intelligence \(AI\)](#), [Blog](#), [Machine Learning](#), [News](#), [Partners](#), [Technology](#)



Bullhorn published a [Q&A session](#) with Textkernel's CEO Jakub Zavrel. You can read the full interview about the future of recruitment and how AI and Machine Learning will have an impact on the HR industry below in Textkernel's blog post.

Interview with Jakub Zavrel, CEO at [Textkernel](#)

There's been a lot of talk about AI and Machine Learning in the staffing and recruiting industry, but let's define the problem space. If you think about the entire lifecycle of the recruiting process, in the next 5-10 years where will machines really make a difference and where will humans continue to be able to add value? Will recruiters have jobs in 2025?



To me, it is without a doubt by now, that AI and Machine Learning will have a huge impact on all areas of business.

Newsletter

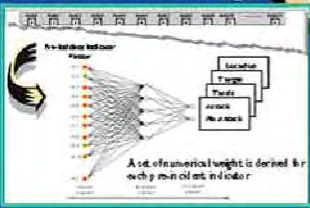
Categories

- > [Active Sourcing](#)
- > [Anniversary](#)
- > [Annual review](#)
- > [Artificial Intelligence \(AI\)](#)
- > [Awards](#)
- > [Big Data](#)
- > [Blog](#)
- > [Candidate experience](#)
- > [Clients](#)
- > [CV parsing](#)
- > [Events](#)
- > [Harvester](#)
- > [Innovation week](#)
- > [Jobfeed](#)
- > [Labour market analysis](#)
- > [Language models](#)
- > [Machine Learning](#)
- > [Match!](#)
- > [Mobile](#)
- > [New release](#)



Pattern Analysis

Predictive Modeling

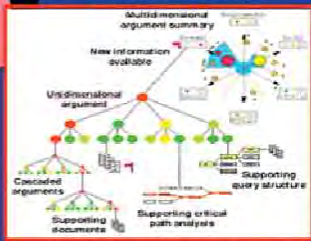


Collaboration Tools

مصر للطيران قد تعود غدا الاربعاء رحلتها الى ليبيا
 القاهرة 6 - 4 (اف ب) - اعلن مسؤولون في شركة
 الخطوط المصرية للطيران اليوم الثلاثاء ان شركة
 "مصر للطيران" قد تمكثف اعتبارا من يوم غد
 الاربعاء رحلتها الى ليبيا اثر قرار مجلس الامن
 الدولي تعليق الحظر المفروض على ليبيا.

Foreign Languages

Analysis Tools and Decision Aides



Überblick



Big Brother liest mit: Computerlinguistik und Wissenschaftsförderung

Big Brother denkt mit: Maschinelles Lernen und Vorurteile

Big Brother wirkt mit: Künstliche Intelligenz und Wir

Unsere Verantwortung als Wissenschaftler

Weshalb ich besorgt bin?





*Interpol und Deutsche Bank, FBI und Scotland Yard
Interpol und Deutsche Bank, FBI und Scotland Yard
Flensburg und das BKA, Haben unsere Daten da
Flensburg und das BKA, Haben unsere Daten da
Zahlen, Handel, Leute
Nummern, Zahlen, Handel, Leute
Computerwelt Computerwelt
Denn Zeit ist Geld Denn Zeit ist Geld*

...

*Automat und Telespiel
Leiten heute die Zukunft ein
Computer für den Kleinbetrieb
Computer für das eigene Heim
Reisen, Zeit, Medizin, Unterhaltung
Reisen, Zeit, Medizin, Unterhaltung
Computerwelt Computerwelt
Denn Zeit ist Geld Denn Zeit ist Geld*



[http://www.badische-zeitung.de/deutschland-1/
fast-schon-ziviler-ungehorsam--9222957.html](http://www.badische-zeitung.de/deutschland-1/fast-schon-ziviler-ungehorsam--9222957.html)

Weshalb ich besorgt bin?



Source: <http://reboundmobile.com/6-best-strategies-to-crack-the-south-korean-mobile-game-market/>

Unsere Verantwortung ...

... als Wissenschaftler



Unsere Verantwortung ...



... als Wissenschaftler

- neben dem unbestreitbaren Nutzen vieler Anwendungen aus Computerlinguistik/Maschinellem Lernen/Künstlicher Intelligenz, können diese auch für fragwürdige Zwecke (etwa Überwachung, militärischer Nutzen, etc.) eingesetzt werden \Rightarrow **Dual Use-Problematik**
- Anwendungen können auch unerwartete **Nebenwirkungen** haben: MP3 (Musikindustrie), Scannerkassen (soziale Beziehungen), Empfehlungssysteme (Filterblase), ...
- **Dual Use plus Nebenwirkungen – Überwachung plus Vorurteile in der Auswertung ...**

Unsere Verantwortung ...



... als Wissenschaftler

- um **Dual Use** und **Nebenwirkungen vorzubeugen**, sollten wir uns gut überlegen, was wir tun, bevor wir es tun
- wir sollten nicht blind nach Projektförderung streben, sondern Fördermaßnahmen kritisch hinterfragen

Unsere Verantwortung ...



... **als Wissenschaftler** sollten wir sprechen

- mit Kollegen: *First Workshop on Ethics in Natural Language Processing*

Unsere Verantwortung ...



... **als Wissenschaftler** sollten wir sprechen

- mit Kollegen: *First Workshop on Ethics in Natural Language Processing*
- mit Studenten: Seminar *The Dark Side of NLP* an der Universität Heidelberg

Unsere Verantwortung ...



... als **Wissenschaftler** sollten wir sprechen

- mit Kollegen: *First Workshop on Ethics in Natural Language Processing*
- mit Studenten: Seminar *The Dark Side of NLP* an der Universität Heidelberg
- mit der Öffentlichkeit: am 25.3. in der Süddeutschen Zeitung, am 22.4. auf dem *Science March*, heute hier ...



#AfterTheMarch



Science March - und jetzt?

Das Bild des Wissenschaftlers in der Öffentlichkeit

- Podiumsdiskussion -

MARCH FOR SCIENCE
- HEIDELBERG -



Das Bild von Wissenschaft existiert multimedial, wie sie begegnet wird und wie mit ihr umgegangen wird. Von Bürgern, von der Politik und von Einflussgruppen. Die Macht fließt auf die Wissenschaft zurück zu reflektieren oder die sogar steuern. Wie wird sie wahrgenommen? Aber wie wird Wissenschaft wahrgenommen in der Öffentlichkeit und die Verantwortung des Einzelnen? Und wie wird sich die Wissenschaft selbst der Wissenschaftler selbst? Nicht weniger wichtig ist die Frage: wie sieht das aus? Wie findet Wissenschaft statt? Und wie geht es? Und wie sieht die Frage: Wie muss sich ändern damit sich das Bild der Wissenschaft ändert? Müssen Bürger ihre Sicht ändern, muss Wissenschaft anders kommunizieren werden, muss sich die Wissenschaft selbst in Teilen ändern?

Diese Fragen können im Mittelpunkt der ersten öffentlichen Podiumsdiskussion des March for Science Heidelberg. Die Veranstaltung richtet sich an interessierte Bürger und Wissenschaftler.

Die Veranstaltung ist öffentlich, der Eintritt ist frei.

Moderation: Carsten Könneker

- Podiumsteilnehmer -

Theresia Bauer

Michael Baumann

Michael Strube

Matthias Hentze

Amardeo Sarma

18. Mai 2017 - 19:30 - 21:30

ORT: DKFZ Kommunikationszentrum

Im Neuenheimer Feld 280, 69120 Heidelberg, Hörsaal 1



www.facebook.com/ScienceMarchHD/

www.marchforscience.de/auff-in-dieser-stadt-heidelberg/



Und wer fördert meine Arbeit?



Und wer fördert meine Arbeit?



Klaus Tschira Stiftung
Deutsche Forschungsgemeinschaft
Leibniz Gemeinschaft

Vielen Dank!



Contact: `michael.strube@h-its.org`

Papiere und weitere Informationen: `http://www.h-its.org/nlp`

Internetbenutzung ohne Spuren



Empfehlungen (ohne Garantie):

1. Firefox verwenden

1.1 *Einstellungen* → *Erweitert* → *Datenübermittlung*: alles deaktivieren

1.2 Add-Ons entrümpeln

1.3 der *private Modus* bringt nichts

1.4 empfohlene Add-Ons:

- HTTPS-Everywhere
- Self-Destructing Cookies
- uBlock oder Privacy Badger

2. Verhaltensweisen:

- **nie** auf einem Browser einloggen (Google, Facebook, etc.) und suchen, surfen, etc.
- entweder mehrere Instanzen von Firefox oder verschiedene Browser für verschiedene Dienste verwenden
- Lokalisierung auf Smartphone und Tablet abstellen

Internetbenutzung ohne Spuren



Empfehlungen (ohne Garantie):

3. alternative Dienste verwenden:

- Suchmaschine: `startpage.com`, `duckduckgo.com`
- Online-Karten: `openstreetmap.org`
- Messenger: *Signal*
- Email-Anbieter: `mailbox.org`, `posteo.de`

4. einmal im Jahr eine Crypto Party besuchen:

`cryptoparty-hd.de`

References

- Abu-Jbara, Amjad, Ben King, Mona Diab & Dragomir Radev (2013).
Identifying opinion subgroups in Arabic online discussions.
In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Sofia, Bulgaria, 4–9 August 2013, pp. 829–835.
- Agarwal, Apoorv, Anup Kotalwar & Owen Rambow (2013).
Automatic extraction of social networks from literary text: A case study on Alice in Wonderland.
In *Proceedings of the 6th International Joint Conference on Natural Language Processing*, Nagoya, Japan, 14–18 October 2013, pp. 1202–1208.
- Bergsma, Shane & Benjamin Van Durme (2013).
Using conceptual class attributes to characterize social media users.
In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Sofia, Bulgaria, 4–9 August 2013, pp. 710–720.
- Bolukbasi, Tolga, Wei Kai-Chang, James Zou, Venkatesh Saligrama & Adam Kalai (2016).
Man is to computer programmer as woman is to homemaker? Debiasing word embeddings.
arXiv:1607.06520.
- Burger, John D., John Henderson, George Kim & Guido Zarrella (2011).
Discriminating gender on Twitter.
In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, Edinburgh, Scotland, U.K., 27–29 July 2011, pp. 1301–1309.
- Caliskan, Aylin, Joanna J. Bryson & Arvind Narayanan (2017).
Semantics derived automatically from language corpora contain human-like biases.
Science, 356(6334):183–186.
- Chen, Annie (2013).
Patient experience in online support forums: Modeling interpersonal interactions and medication use.
In *51st Annual Meeting of the Association for Computational Linguistics: Proceedings of the Student Research Workshop*, Sofia, Bulgaria, 5–7 August 2013, pp. 16–22.
- Eisenstein, Jacob, Noah A. Smith & Eric P. Xing (2011).
Discovering sociolinguistic associations with structured sparsity.
In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Portland, Oreg., 19–24 June 2011, pp. 1365–1374.
- Grishman, Ralph & Beth Sundheim (1996).

Message Understanding Conference-6: A brief history.

In *Proceedings of the 16th International Conference on Computational Linguistics*, Copenhagen, Denmark, 5–9 August 1996.

Jurafsky, Daniel & James H. Martin (2008).

Speech and Language Processing (2nd. ed.).

Upper Saddle River, N.J.: Prentice Hall.

Lin, Ching-Sheng, Samira Shaikh, Jennifer Stromer-Galley, Jennifer Crowley, Tomek Strzalkowski & Veena Ravishankar (2013).

Topical positioning: A new method for predicting opinion changes in conversation.

In *Proceedings of the Workshop on Language Analysis in Social Media*, Atlanta, Georgia, 13 June 2013, pp. 41–48.

Popp, Robert, Thomas Armour, Ted Senator & Kristen Numrzch (2004).

Countering terrorism through information technology.

Communications of the ACM, 47(3):36–43.

Sundheim, Beth M. (1991).

Overview of the third Message Understanding and Evaluation Conference.

In *Proceedings of the 3rd Message Understanding and Evaluation Conference*, San Diego, Cal., 21-23 May 1991, pp. 3–16.

Volkova, Svitlova, Glen Coppersmith & Benjamin van Durme (2014).

Inferring user political preferences from streaming communications.

In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Baltimore, Md., 22–27 June 2014, pp. 186–196.

Weischedel, Ralph, Eduard Hovy, Mitchell Marcus, Martha Palmer, Robert Belvin, Sameer Pradhan, Lance Ramshaw & Nianwen Xue (2011).

OntoNotes: A large training corpus for enhanced processing.

In J. Olive, C. Christianson & J. McCary (Eds.), *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitations*, pp. 54–63. Heidelberg, German: Springer.