

# OVERVIEW OF NATURAL LANGUAGE PROCESSING (NLP) TECHNOLOGIES

---

Presented by: WTI  
[www.wti-solutions.com](http://www.wti-solutions.com)  
703.286.2416



# LEGAL DISCLAIMER

The entire contents of this informational publication is protected by the copyright laws of the United States and other jurisdictions. You may print a copy of any part of this publication for your own personal, noncommercial use, but you may not copy any part of the publication for any other purposes, nor may you modify any part of the publication. Inclusion of any part of the content of this publication in another work, whether in printed or electronic, or other form, or inclusion of any part hereof in another publication or web site by linking, framing, or otherwise without the express written permission of WTI is hereby prohibited.

WTI's informational publications are made available for educational purposes only as well as to give you general information and a general understanding of technology, not to provide technical advice. By reading our informational publications you understand that there is no contractual relationship created between you and WTI. Although the information in our informational publications is intended to be current and accurate, the information presented herein may not reflect the most current technical developments. These materials may be changed, improved, or updated without notice. WTI is not responsible for any errors or omissions in the content of this publication or for damages arising from the use or performance of this publication under any circumstances. We encourage you to contact us or other technology companies for specific technical advice as to your particular matter.

# WHAT IS NLP AND WHY IS IT IMPORTANT?

- **Definition** - Natural language processing (NLP) is the merging of computer processing, and computational linguistics DO1 artificial intelligence to create functions that allow the processing of large amounts of human produced data (written DO2 poken) into machine understandable data sets for analytics.
- **Importance** - The challenges it attempts to solve are natural language understanding, natural language generation, connecting language to machine perception, and creating dialog from that.

### Slide 3

---

**DO1** corrected typo - Linguistics

Debra O'Connell, 6/22/2017

**DO2** should be lower case "w"

Debra O'Connell, 6/22/2017

# NLP ORIGINS DO4

- First established as a field in “Computing Machinery and Intelligence” by Alan Turing in 1950, who outlined the possibilities of what would be known as the Turing Test - in which a computer could understand questions and response well enough to fool a human tester in believing they were speaking to a live human rather than a machine.
- Work began DO5 create systems to automatically translate documents between different languages.
- A decade later the concept of the technology would be popularized by television programs such as “Star Trek” and movies like 2001: A Space Odyssey which showed computers understanding and replying in natural fashion. DO16

<sup>1</sup> Deangelis, Stephen F. (Feb 2014) The Growing Importance of Natural Language Processing. Retrieved from <https://www.wired.com/insights/2014/02/growing-importance-natural-language-processing/> DO3

## Slide 4

---

- DO3** Is this from an article, book or report - this is not a full reference  
Debra O'Connell, 6/22/2017
- DO4** I turned all the incomplete sentences below into complete sentences/bulltes  
Debra O'Connell, 6/22/2017
- DO5** work began by who?  
Debra O'Connell, 6/22/2017
- DO16** I inserted teh citation superscript as required  
Debra O'Connell, 6/22/2017

# NLP EVOLUTION

While a popular science fiction concept, the technology failed to gain real advancement for many decades. Concerned by slow progress in 1964, U.S. <sup>DO8</sup> government sponsors commissioned the Automatic Language Processing Advisory Committee (ALPAC) which concluded in 1966 and reported, short-sightedly, <sup>DO9</sup> that it was too slow and too costly compared to human workers so “there is no immediate or predictable prospect of useful machine translation.” <sup>DO17</sup> Its report chilled research in human language machine learning for decades.

<sup>2</sup> Hutchins, John (Nov <sup>DO6</sup> 2005) The History of Machine Learning in a Nutshell. Retrieved from <http://www.hutchins<sup>DO10</sup>.uk/Nutshell-2005.pdf>

## Slide 5

---

- DO6** This is not a complete citation  
Debra O'Connell, 6/22/2017
- DO8** U.S. not US  
Debra O'Connell, 6/22/2017
- DO9** typo - removed "I" from sightedly was slightedly ; )  
Debra O'Connell, 6/22/2017
- DO10** Also it has a 1 vice a 5  
Debra O'Connell, 6/22/2017
- DO17** this was re-numbered to a 2  
Debra O'Connell, 6/22/2017



# NLP REBIRTH

The late 1990s saw the rebirth of NLP with the rise of the World Wide Web and the need for real-time information processing became a mass market need. This pushed interest in NLP from a mostly research topic to a focus on practical applications<sup>3</sup>, such as automated translation, and search engines.

Around the same time on the research side, several breakthroughs combining old rule based systems with newer statistical and example based systems increased the complexity that NLP technologies might be able to handle.

<sup>3</sup> Hutchins DO7h (Nov 2005) The History of Machine Learning in a Nutshell. Retrieved from <http://www.DO18insweb.me.uk/Nutshell-2005.pdf>

## Slide 6

---

**DO7**

Citation needed

Debra O'Connell, 6/22/2017

**DO18**

This is the third citation. The number 3 in superscript needs to be inserted in the text as appropriate

Debra O'Connell, 6/22/2017

# NLP IN THE NEWS

## IBM Watson wins Jeopardy!

Watson, a primarily NLP engine, defeats the top human champions in 2011. Starting as a stand-alone supercomputer it DO11 since evolved into a powerful cloud-based DO12 engine capable of outperforming doctors at medical diagnosis or legal paperwork processing.<sup>4</sup>



<sup>4</sup> Captain, Sean (Jan 2017) Can IBM's Watson do it all? Retrieved from <https://www.fastcompany.com/3065339/can-ibms-watson-do-it-all>

## Slide 7

---

**DO11** I should be lower case  
Debra O'Connell, 6/22/2017

**DO12** inset dash  
Debra O'Connell, 6/22/2017

# NLP IN THE NEWS

## NLP Teaching Assistant Fools Students

Utilizing the later cloud-based Watson technology, an artificial intelligence professor at Georgia Tech created a virtual teaching assistant to help respond to the over 10,000 inquiries received by students. It did a good enough job, most did not realize their answers came from a computer.<sup>5</sup>

DO14  
DO19

DO15



<sup>5</sup> Bussing, Kim (Sept 2016) Georgia Tech's Teaching Assistant "Jill Watson" Turns Out To Be A Robot! Retrieved from <https://www.dogonews.com/2016/9/10/georgia-techs-teaching-assistant-jill-watson-turns-out-to-be-a-robot>

DO120m  
DO20

## Slide 8

---

**DO13** this is not a complete citation and not numbered correctly

Debra O'Connell, 6/22/2017

**DO14** insert the correct superscript citation number

Debra O'Connell, 6/22/2017

**DO15** removed word "thy"

Debra O'Connell, 6/22/2017

**DO19** I changed the number to correct it, as 5

Debra O'Connell, 6/22/2017

**DO20** I corrected

Debra O'Connell, 6/22/2017

# NLP IN THE NEWS

## Advertisements activate home devices

Taking advantage of the rise of home assistants from Google and Amazon, advertisers create commercials intended to activate Google home DO21 devices and use them to continue giving product information beyond the commercial's length. Google quickly updated the devices to prevent the specific ad's command, and prevention of unwanted activations remains an ongoing challenge.<sup>6</sup> DO22



<sup>6</sup> Maheshwari, Sapna (Apr 2017) Burger King 'O.K. Google' Ad Doesn't Seem O.K. With Google. Retrieved from <https://www.nytimes.com/2017/04/12/business/burger-king-tv-ad-google-home.html>

## Slide 9

---

**DO21**

lower case for "home

Debra O'Connell, 6/22/2017

**DO22**

Renumbered the citation, and this is not a complete citation - please correct it

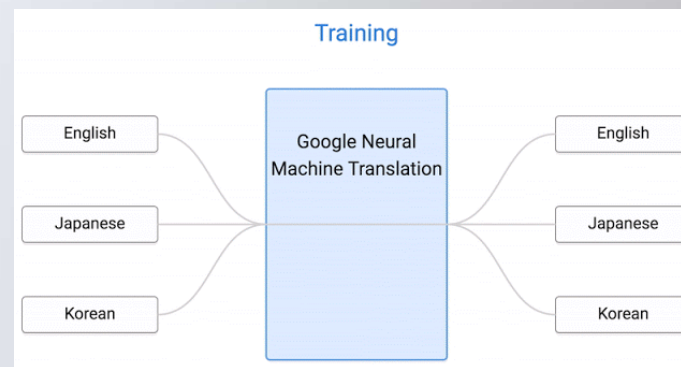
Debra O'Connell, 6/22/2017



# NLP IN THE NEWS

## Google Translate invents its own language

An exaggerated headline to explain a still-impressive achievement in NLP, the system behind Google translate was testing a new “neural machine translation” and the system, unprompted, determined the best way translate two languages that did not have direct correlations was to create an “interlingua” using data from a third language to assist in figuring out the correct translation. <sup>7</sup>



<sup>7</sup> Reback, Gedalyah (Jan 2017) No, Google Translate did not invent its own language called 'interlingua'. Retrieved from <https://www.geektime.com/2017/01/23/no-google-translate-did-not-invent-its-own-language-called-interlingua/>

# NLP EVERYDAY

NLP technologies are also moving beyond large dataset crunching and becoming part of our everyday lives. For example, DO24 is used in everyday DO25 interfaces such as:

- Apple's Siri
- Google's "Ok Google"
- Automated support DO26 chat boxes
- Automated phone answering services
- Automated journalism DO27 articles
- Optical Character Recognition technologies<sup>8</sup>

<sup>8</sup> "Outlook on Artificial Intelligence in the Enterprise 2016", Narrative Science in partnership with National Business Research Institute, Page 6, retrieved from <http://www.datascienceassn.org>

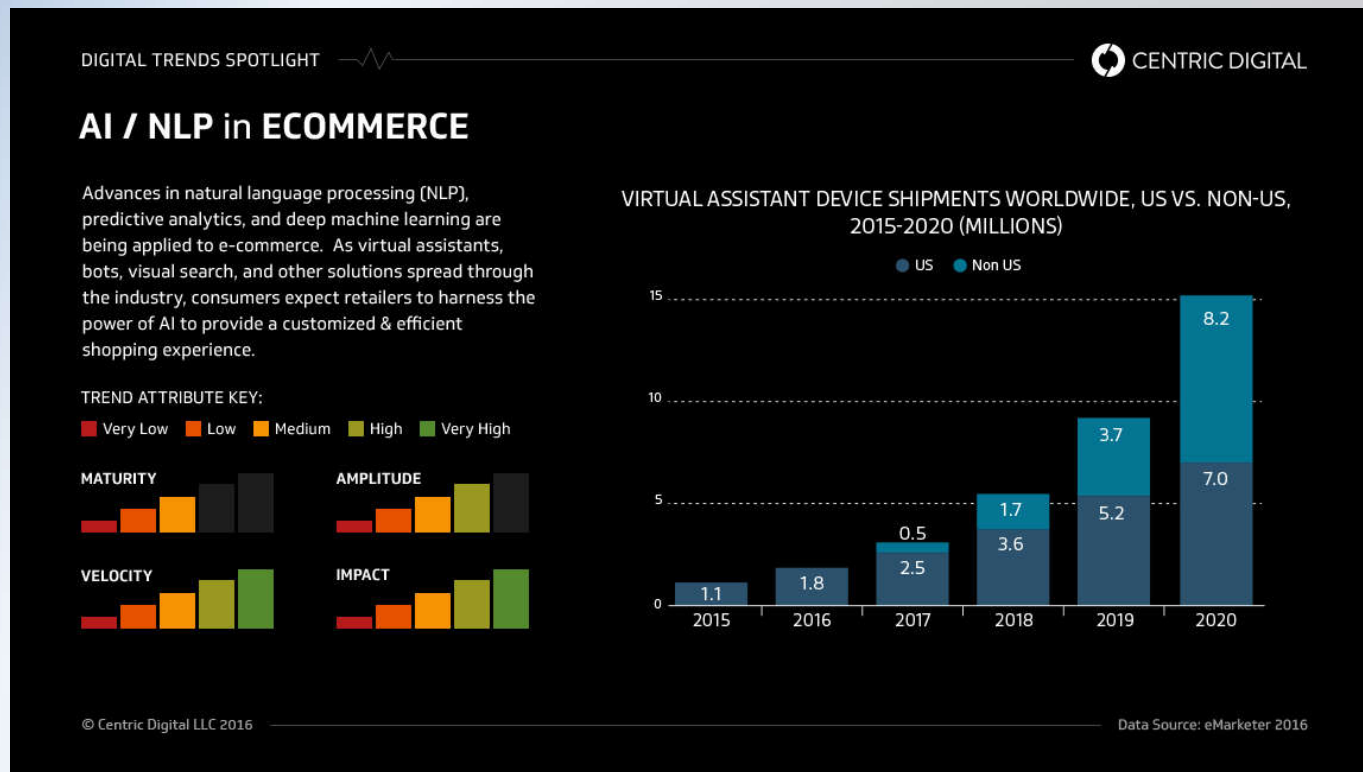
## Slide 11

---

- DO24** insert comma  
Debra O'Connell, 6/22/2017
- DO25** remove space  
Debra O'Connell, 6/22/2017
- DO26** lower case font  
Debra O'Connell, 6/22/2017
- DO27** lower cae font  
Debra O'Connell, 6/22/2017

# NLP FOR VIRTUAL ASSISTANTS DO29

As the usefulness of NLP functions increase, consumer demand for DO30 al assist DO31 is expected to rise sharply.



<sup>9</sup> <https://centricdigital.com/blog/artificial-intelligence/how-natural-language-process-artificial-intelligence-are-changing-e-commerce/> DO28 DO32

## Slide 12

---

**DO28** This is not a proper citation - please correct, I renumbered in text and bottom

Debra O'Connell, 6/22/2017

**DO29** added nusubg S

Debra O'Connell, 6/22/2017

**DO30** bolded tex

Debra O'Connell, 6/22/2017

**DO31** changed teh word AI to NLP

Debra O'Connell, 6/22/2017

**DO32** Debra O'Connell, 6/22/2017

## HOW DOES NLP DIFFER

Use of DO35 is different from prior process for older systems, in which the exact rules had to be entered to create the program. NLP systems are trained using machine learning to analyze example data sets and make inferences. Examples include:

- Summarizing DO36 blocks of text to extract the most important and central ideas while ignoring irrelevant information.
- Creating a chat bot that uses Point-of-Speech tagging.
- Automatically generating keyword tags from content which leverages LDA, a technique that discovers topics contained within a body of text.
- Using Sentiment Analysis to identify the sentiment of DO33 of text, from very negative to neutral to very positive.<sup>10</sup>

<sup>10</sup> Kiser, Matt (Aug 2016) Introduction to Natural Language Processing (NLP DO34). Retrieved from <https://blog.algorithmia.com/introduction-natural-language-processing-nlp/>.

## Slide 13

---

**DO33**

corrected citation number

Debra O'Connell, 6/22/2017

**DO34**

this is not a proper citation - needs update

Debra O'Connell, 6/22/2017

**DO35**

changed from "are" to "is"

Debra O'Connell, 6/22/2017

**DO36**

changed tense to "ing" words

Debra O'Connell, 6/22/2017

# WHY DOES NLP MATTER TO GOVERNMENT

The Government has need for rapid processing of large text datasets. NLP can reduce the workload and speed response. DO37

Examples include:

- Civic Government - Sorting public comments to understand policy response.
- Emergency Services - Insurance and disaster response.
- Healthcare - Paperwork processing and diagnostic assistance.
- Law enforcement – Crime solving and prevention by analyzing multiple case files for patterns and connections.<sup>11</sup> DO38

<sup>11</sup> Murrow, Brian (Mar 2014) Extracting Meaningful Analysis from Data Using Natural Language Processing. Retrieved from <https://www.ibm.com/blogs/insights-on-business/government/extracting-meaningful-analysis-from-data-using-natural-language-processing/> DO39





## Slide 14

---

**DO37** change in grammar  
Debra O'Connell, 6/22/2017

**DO38** I inseted superscript citation number and corrected number to 11  
Debra O'Connell, 6/22/2017

**DO39** this is not full citation - needs correction  
Debra O'Connell, 6/22/2017

# EXAMPLES OF WTI NLP WORK

WTI works with Government agencies to provide innovative NLP solutions that: BR1

- Process data from the web and other sources to extract information about the world in order to persist the data and query it.
- Integrate commercial products such as Basistech Rosette to extract “entities” from text.
- Extract entities using existing ontologies that describe military equipment.
- Develop processes to create relationships between entities that were extracted using both commercial and home grown NLP.
- Develop an Object Store to persist the results of NLP and perform queries.
- Develop ontologies using RDF to classify the entities that are extracted by NLP.

## Slide 15

---

**BR1**

These come from Ryan Lantzy. Too many or too wordy for one page so I'll the bullets other recommend.

Brent Reitze, 6/22/2017

# CONTACT INFORMATION

WTI is a CMMI ® Level 3 appraised and SBA 8(a) Certified, Economically Disadvantaged Woman-Owned Small Business based in the National Capital Region, committed to helping the government and industry harness the power of actionable information through intelligence and IT solutions.

WTI's work in NLP spans multiple Federal Government executive departments and a number of independent Federal agencies. To find out more, contact:

## **Brent Reitze**

### **Technical Engineer, Corporate Development**

703.286.2416 ext. 119

571.345.5366 (mobile)

703.997.7227 (fax)

[brent.reitze@wti-solutions.com](mailto:brent.reitze@wti-solutions.com)

## **Debra O'Connell**

### **Director of Strategic Relationships**

703.286.2416 ext. 105

301.312.1263 (mobile)

703.997.7227 (fax)

[debra.oconnell@wti-solutions.com](mailto:debra.oconnell@wti-solutions.com)