



BASIS
TECHNOLOGY

GRAHAM MOREHEAD

CHRIS BIOW

Linguistic Extraction for Prediction

English

What is Big Text?

Pronoun

Verb

Name

Concept

It's huge volumes of multilingual, unstructured

Pronoun+Verb

Adjective

Noun

Prep.

Adjective

Adjective

text that must be processed to deliver insights

Noun

Rel. Pronoun

Verb

Verb

Verb

Inf.

Verb

Noun

and build connections. It's President Clinton

Conjunction

Verb

Noun

Pronoun+Verb

Person

Title

Name

helping Malawi. Secretary Clinton in اسلام آباد.

Verb

Name

Title

Name

Prep.

Name

Urdu: "Islamabad"

The 福島 meltdown. This is Big Text.

Determiner

Place

福島

Japanese:
"Fukushima"

Noun

Pronoun

Verb

Concept

Big Text

美国《华盛顿邮报》25日称，“小布什总统中心”南方卫理公会大学举行落成仪式。24日，小布什夫妇接他(杰布)选择这样一位总统。

《华盛顿邮报》人认为，希拉里美联社分析称，关系将他成为和成长于墨西哥制度。不过，他拙劣工作所玷污需求。那几年，女儿诺伊尔2002

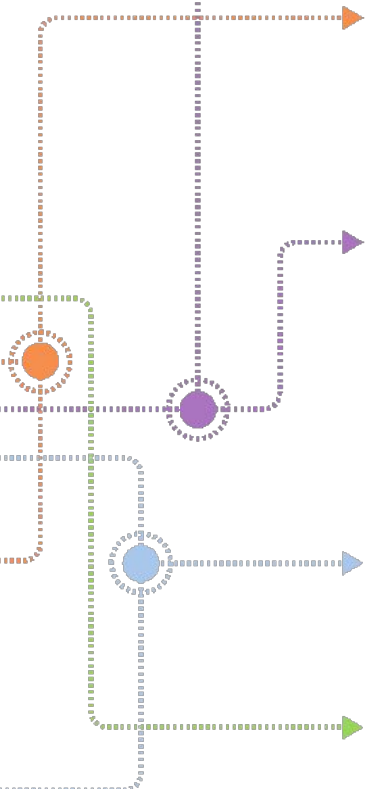


FISHING NEWS

KENNEBUNKPORT, Maine — Three generations of an American political dynasty went fishing off the Maine coast. The family set out together in the morning on their new powerboat, built by **Penobscot Boat Builders**. On board were the first **President Bush**; the second **President Bush** and his niece, **Noelle Bush**.

Un es
durar
pone
Slate
de un
nume
la ma

paracaídas. Slate señaló que se sintió inspirado por el **expresidente de Estados Unidos Goerge H.W. Bush**, quien realizó un salto con motivo de su 85 cumpleaños en 2009. Tras pisar tierra, el nonagenario dijo que esperaba poder repetir en su 95 y 100 cumpleaños.



1 Identifying Known Entities WIKIPEDIA



George H. W. Bush

41st U.S. President
31 January 1917
2018 - June 10, 1998

George Herbert Walker Bush is an American politician who served as the 41st President of the United States.



George W. Bush

43rd U.S. President
29 July 1946
2018 - July 6, 2001

George Walker Bush is an American oil billionaire businessman who was the 43rd President of the United States of America from 2001 to 2009 and the 71st Governor of Texas from 1995 to 2000.

2 Inferring New Entities



Unknown Person

Noelle Bush

ID: 30237552



Unknown Organization

Penobscot Boat Builders

ID: 11267723



EMBERS :

- Linguistic-based Prediction System for *Civil Unrest*
- 24x7 Continuous, Automated (*No human intervention!*)
- Across 10 Countries in Latin America
- Extracts intelligence from: *blogs, tweets, news, economic indicators*
- Processes 20GB/day
- Funder: IARPA Open Source Indicators (OSI) program
- began April 2012
- Partners:

CACI

 VirginiaTech


Discovery Analytics Center

GRAHAM MOREHEAD 617-970-9688

gmorehead@basistech.com

CHRIS BIOW

cbiow@basistech.com

2553 Dulles View Dr. #450 Herndon, VA

www.BasisTech.com

More about EMBERS:

<http://people.cs.vt.edu/naren/papers/kddindg1572-ramakrishnan.pdf>