

# NLP

# Introduction to NLP

## *Summarization*

# Text Summarization

## Health Benefits

- Eating a diet rich in vegetables and fruits as part of an overall healthy diet may reduce risk for heart disease, including heart attack and stroke.
- Eating a diet rich in some vegetables and fruits as part of an overall healthy diet may protect against certain types of cancers.
- Diets rich in foods containing fiber, such as some vegetables and fruits, may reduce the risk of heart disease, obesity, and type 2 diabetes.
- Eating vegetables and fruits rich in potassium as part of an overall healthy diet may lower blood pressure, and may also reduce the risk of developing kidney stones and help to decrease bone loss.
- Eating foods such as vegetables that are lower in calories per cup instead of some other higher-calorie food may be useful in helping to lower calorie intake.

## Nutrients

- Most vegetables are naturally low in fat and calories. None have cholesterol. (Sauces or seasonings may add fat, calories, or cholesterol.)
- Vegetables are important sources of many nutrients, including potassium, dietary fiber, folate (folic acid), vitamin A, and vitamin C.
- Diets rich in potassium may help to maintain healthy blood pressure. Vegetable sources of potassium include sweet potatoes, white potatoes, white beans, tomato products (paste, sauce, and juice), beet greens, soybeans, lima beans, spinach, lentils, and kidney beans.
- Dietary fiber from vegetables, as part of an overall healthy diet, helps reduce blood cholesterol levels and may lower risk of heart disease. Fiber is important for proper bowel function. It helps reduce constipation and diverticulosis. Fiber-containing foods such as vegetables help provide a feeling of fullness with fewer calories.
- Folate (folic acid) helps the body form red blood cells. Women of childbearing age who may become pregnant should consume adequate folate from foods, and in addition 400 mcg of synthetic folic acid from fortified foods or supplements. This reduces the risk of neural tube defects, spina bifida, and anencephaly during fetal development.
- Vitamin A keeps eyes and skin healthy and helps to protect against infections.
- Vitamin C helps heal cuts and wounds and keeps teeth and gums healthy. Vitamin C aids in iron absorption.

# Summary

Eating vegetables is healthy.

# News Summarization

NewsInEssence: Web-based News Summarization - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites History Print Mail News RSS Feeds

Address http://www.newsinessence.com/nle.cgi

...www...NewsInEssence...com...

## Interactive Multi-source News Summarization

Home  
Current Clusters  
Create Cluster  
Summarize Cluster  
Track Cluster  
User Cluster Archive  
CIDR Cluster Archive  
Google Cluster Archive

Help  
About NewsInEssence  
Contact Us

CLAIR  
MEAD  
summarization.com

### Pressure grows on Bush to globalise Iraq effort

Bangkok Post Friday 1 August, 2003 - US sets mid-2004 target for Iraq elections, two soldiers killed. BBC NEWS World Middle East Bush under fire over Iraq. President George W Bush is coming under increasing pressure from his own Republican Party to disclose how much financing will be required to cover the costs of occupying and reconstructing Iraq.

[7 Articles from 6 Sources] [5 Summaries]

Recent NIE News Clusters (more)

- 'Ananova - Tensions high as cortege approaches Najaf'  
24 articles, 4 summaries: 09/02, 6:10 AM
- 'Bush Makes Push for Manufacturing Jobs'  
18 articles, 4 summaries: 09/02, 6:10 AM
- 'Israeli Strike Kills Hamas Member September 1, 2003 21:44:37'  
12 articles, 4 summaries: 09/02, 6:10 AM

Recent User News Clusters (more)

- 'Japan launches asteroid probe'  
4 articles, 3 summaries: 09/02, 11:28 AM
- 'Death tax tops 50 more'  
1 article, 3 summaries: 09/02, 10:32 AM
- 'Spam peddlers hijack computers'  
7 articles, 3 summaries: 09/02, 10:23 AM

NIE Headlines  
Build your own cluster of articles.

NewsTroll from URL:  
URL must be from CNN, Yahoo!, MSNBC, BBC, or USA Today.  
http://www.hindoonnet.c  
Search

NewsTroll from query:  
luger bremer bush iraq  
Search

Advanced Options

---

NIE News Clusters (Archive)

- 'Ananova - Tensions high as cortege approaches Najaf'  
24 articles, 4 summaries: 09/02, 6:10 AM
- 'Bush Makes Push for Manufacturing Jobs'  
18 articles, 4 summaries: 09/02, 6:10 AM
- 'Israeli Strike Kills Hamas Member September 1, 2003 21:44:37'  
12 articles, 4 summaries: 09/02, 6:10 AM
- 'Taliban ambush two Afghan patrols'  
9 articles, 4 summaries: 09/02, 6:10 AM
- 'FOXNews.com'  
7 articles, 4 summaries: 09/02, 6:10 AM
- 'BBC SPORT Cricket Cricket to return to Kashmir'  
7 articles, 4 summaries: 09/02, 6:10 AM
- 'Vendredi mullins Bronfman, GE offers: could decide Tuesday'

### Pressure grows on Bush to globalise Iraq effort

produced on 09/02, 6:10 AM

2% Summary

Bangkok Post Friday 1 August, 2003 - US sets mid-2004 target for Iraq elections, two soldiers killed (1.1) BBC NEWS World Middle East Bush under fire over Iraq (2.1) President George W Bush is coming under increasing pressure from his own Republican Party to disclose how much financing will be required to cover the costs of occupying and reconstructing Iraq. (2.2) The Bush administration, which is already spending \$4bn a month of US taxpayers' money on the military costs in Iraq alone, looks to be gearing up for a big international appeal for contributions towards the cost of rebuilding Iraq. (2.10)

Bush is facing growing calls from within his own party and from men running for his job to bring more international troops into Iraq amid mounting US casualties and costs. (4.3) A peacekeeping battalion from Kazakhstan has been deployed

# Book Summaries

- Cliff's notes
- Book a minute (<http://www.rinkworks.com/bookamminute/>)
  - Which book is this?

*(Some BOYS crash on an ISLAND.)*

**Ralph** We need a fire. *(They make a fire. It goes out.)*

**Ralph** We need a fire. *(They make a fire. It goes out.)*

**Ralph** We need a fire.

**Jack** Forget the fire. Let's kill each other.

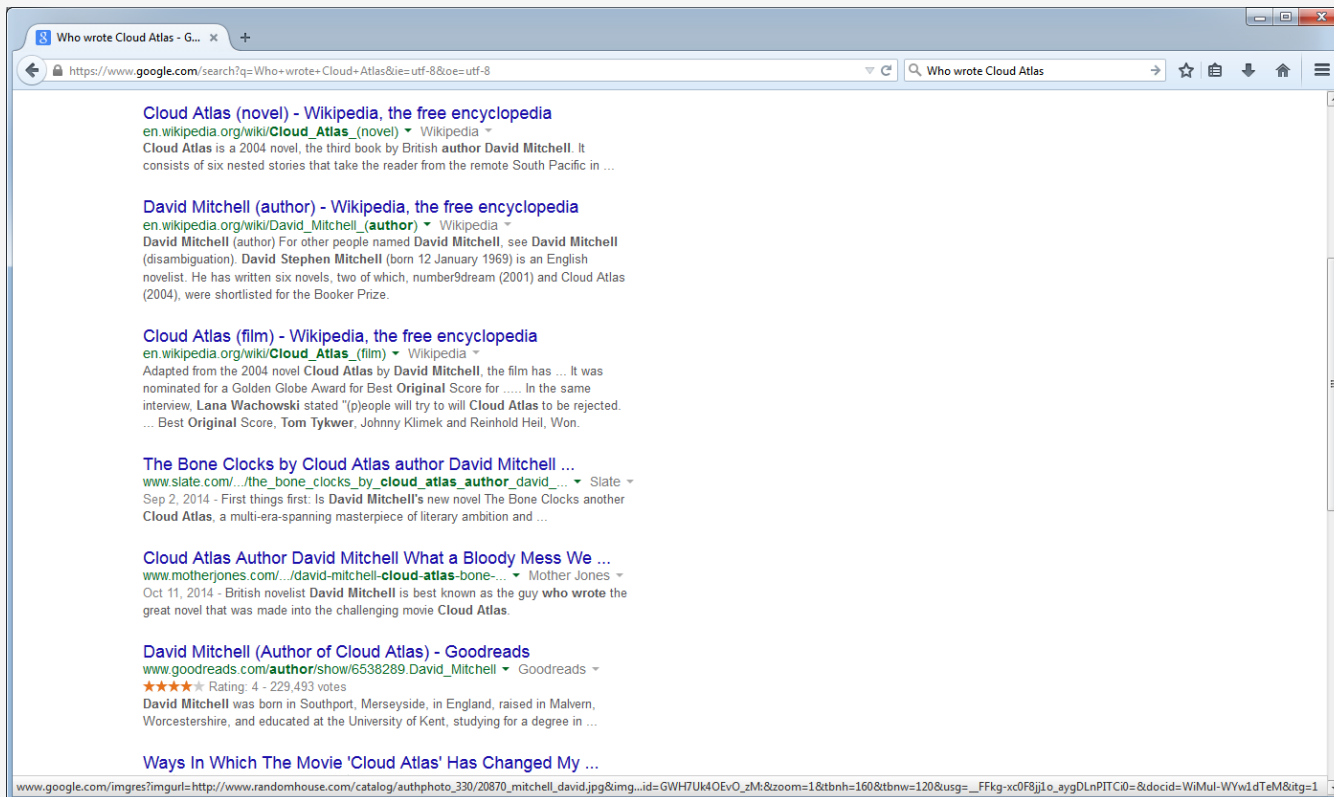
**Other Boys** Yeah! *(They do.)*

**THE END**

# Movie Summaries

- Titanic
  - Beginning with genuine footage of the departure of the Titanic on its fateful voyage, this epic movie tells the events of that tragic night from the perspective of fictional survivor Rose. As an old lady of 100 she recounts her story of duty, love and disaster to a salvage crew searching for a lost gem.
  - Winner of 11 Oscars, James Cameron's effects-driven blockbuster puts a human face on a tragedy of epic proportions by wedding the historical tale of the doomed ocean liner with a fictional romance between two of the ship's ill-fated passengers. Society girl Rose Dewitt Bukate and penniless artist Jack Dawson struggle to survive both the sinking ship -- and the wrath of Rose's wealthy fiancé.

# Search Engine Snippets





## Genres

- headlines
- outlines
- minutes
- biographies
- abridgments
- sound bites
- movie summaries
- chronologies, etc.

[Mani and Maybury 1999]

# Types of Summaries

- Input
- Output
- Purpose
  - Indicative, informative, and critical summaries
- Form
  - Extracts (representative paragraphs/sentences/phrases)
  - Abstracts: “a concise summary of the central subject matter of a document” [Paice90].
- Dimensions
  - Single-document vs. multi-document
- Context
  - Query-specific vs. generic

# Stages

- Three stages (typically)
  - content identification
  - conceptual organization
  - realization

**BAGHDAD, Iraq (CNN) 6 July 2004** -- Three U.S. Marines have died in al Anbar Province west of Baghdad, the Coalition Public Information Center said Tuesday.

According to CPIC, "Two Marines assigned to [1st] Marine Expeditionary Force were killed in action and one Marine died of wounds received in action Monday in the Al Anbar Province while conducting security and stability operations."

Al Anbar Province -- a hotbed for Iraqi insurgents -- includes the restive cities of Ramadi and Fallujah and runs to the Syrian and Jordanian borders.

Meanwhile, officials said eight people died Monday in a U.S. air raid on a house in Fallujah that American commanders said was used to harbor Islamic militants.

A senior U.S. military official told CNN the target was a group of people suspected of planning suicide attacks using vehicles.

The strike was the latest in a series of raids on the city to target what U.S. military spokesmen have called safehouses for the network led by fugitive Islamic militant leader Abu Musab al-Zarqawi.

A statement from Allawi said: "The people of Iraq will not tolerate terrorist groups or those who collaborate with any other foreign fighters such as the Zarqawi network to continue their wicked ways.

"The sovereign nation of Iraq and our international partners are committed to stopping terrorism and will continue to hunt down these evil terrorists and weed them out, one by one. I call upon all Iraqis to close ranks and report to the authorities on the activities of these criminal cells."

American planes dropped two 1,000-pound bombs and four 500-pound bombs on the house about 7:15 p.m. (11:15 a.m. ET), according to a statement from the U.S.-led Multi-National Force-Iraq.

At least four previous air raids have targeted suspected Zarqawi safehouses in Fallujah.

**BAGHDAD, Iraq (CNN) 6 July 2004** -- Three U.S. Marines have died in al Anbar Province west of Baghdad, the Coalition Public Information Center said Tuesday.

According to CPIC, "Two Marines assigned to [1st] Marine Expeditionary Force were killed in action and one Marine died of wounds received in action Monday in the Al Anbar Province while conducting security and stability operations."

Al Anbar Province -- a hotbed for Iraqi insurgents -- includes the restive cities of Ramadi and Fallujah and runs to the Syrian and Jordanian borders.

Meanwhile, officials said eight people died Monday in a U.S. air raid on a house in Fallujah that American commanders said was used to harbor Islamic militants.

A senior U.S. military official told CNN the target was a group of people suspected of planning suicide attacks using vehicles.

The strike was the latest in a series of raids on the city to target what U.S. military spokesmen have called safehouses for the network led by fugitive Islamic militant leader Abu Musab al-Zarqawi.

A statement from Allawi said: "The people of Iraq will not tolerate terrorist groups or those who collaborate with any other foreign fighters such as the Zarqawi network to continue their wicked ways.

"The sovereign nation of Iraq and our international partners are committed to stopping terrorism and will continue to hunt down these evil terrorists and weed them out, one by one. I call upon all Iraqis to close ranks and report to the authorities on the activities of these criminal cells."

American planes dropped two 1,000-pound bombs and four 500-pound bombs on the house about 7:15 p.m. (11:15 a.m. ET), according to a statement from the U.S.-led Multi-National Force-Iraq.

At least four previous air raids have targeted suspected Zarqawi safehouses in Fallujah.

# Human Summarization and Abstracting

- What professional abstractors do
- Ashworth (1973):
  - “To take an original article, understand it and pack it neatly into a nutshell without loss of substance or clarity presents a challenge which many have felt worth taking up for the joys of achievement alone. These are the characteristics of an art form”.

# Extractive Summarization

- Selecting units of the original text
  - Usually sentences
  - No simplification
  - No rewriting
- Baseline
  - Extract the first few sentences

# NLP