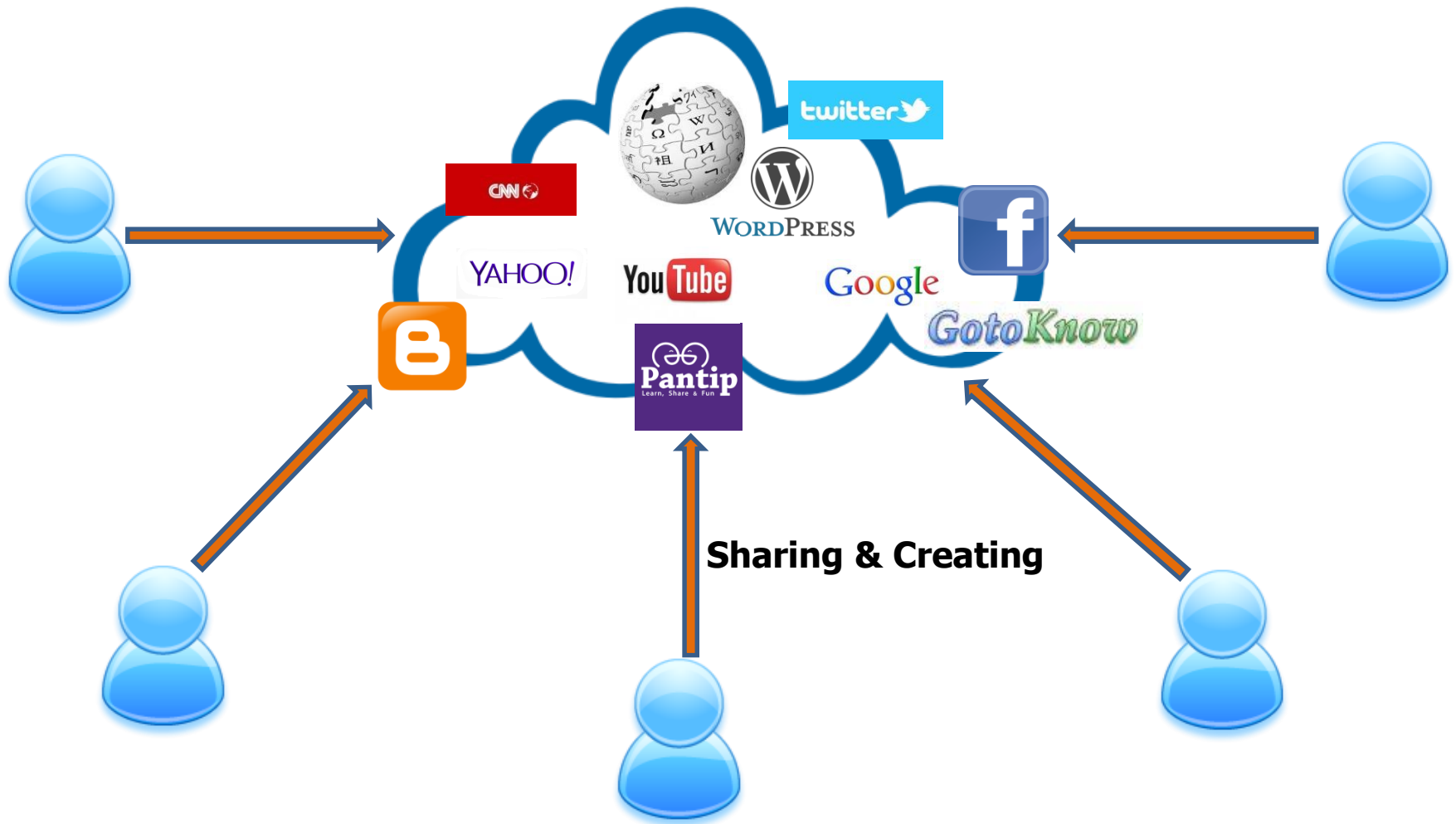


## **CopyAlert**

### **“Monitoring and Protecting Your Valuable Contents”**

**Santipong Thaiprayoon, Kanokorn Trakultaweekoon  
and Alisa Kongthon**

Speech and Audio Technology Laboratory (SPT)  
National Electronics and Computer Technology Center (NECTEC)  
Thailand Science Park



# John's Diary



# Bob's Diary



# John's Diary



# Bob's Diary

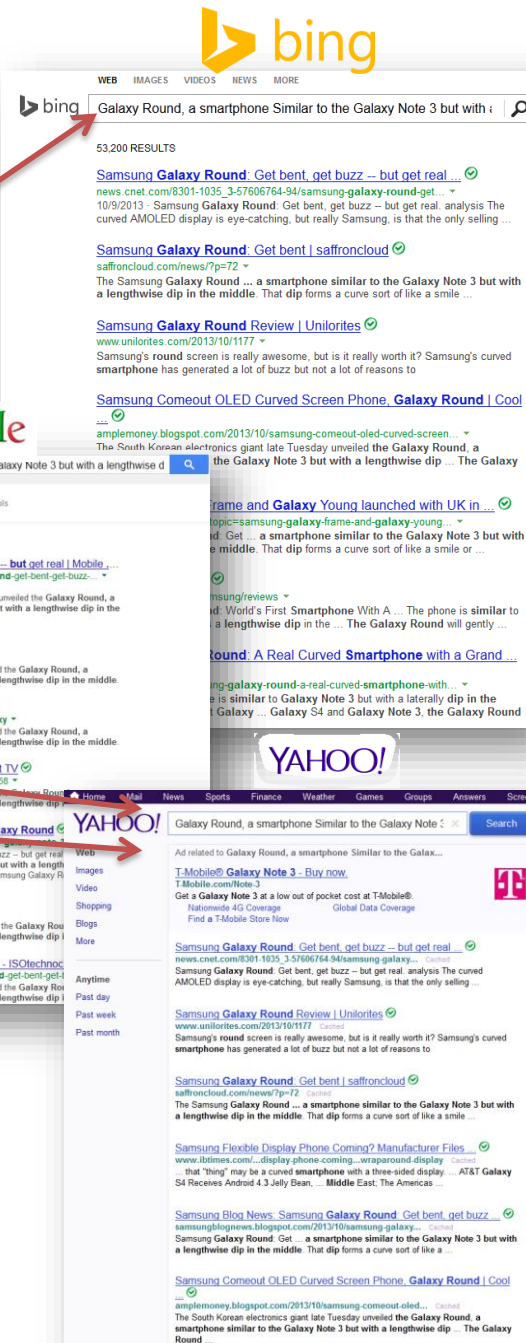
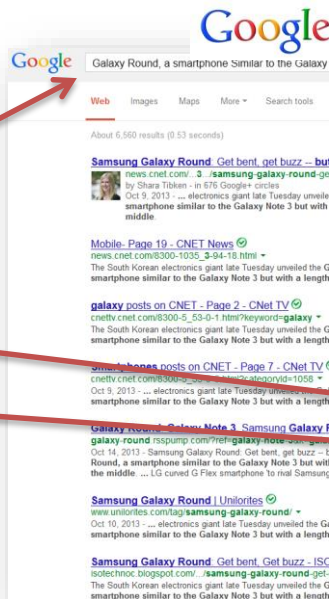


**Sentence#1**

**Sentence#2**

**Sentence#3**

**Sentence#4**



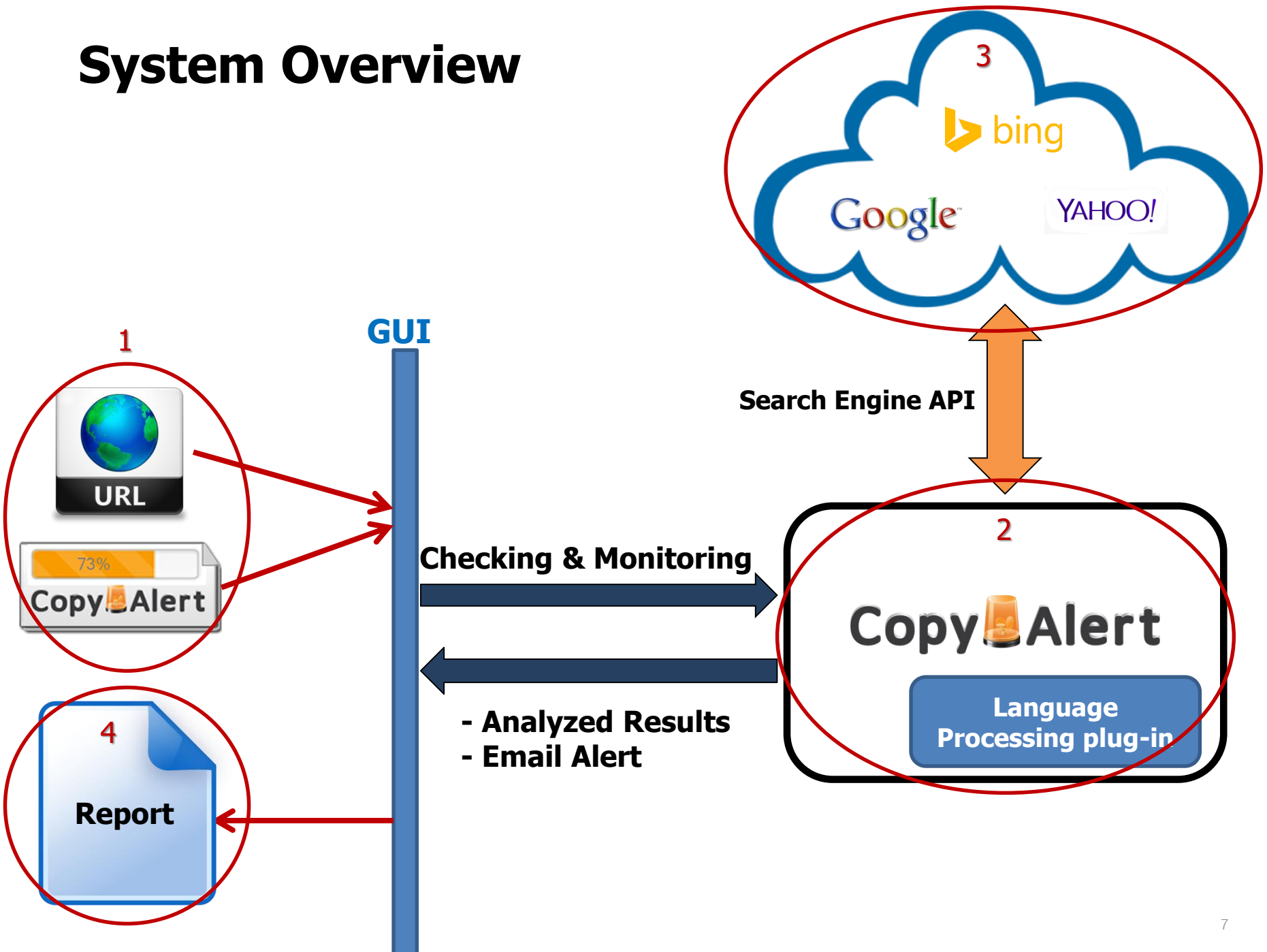
CopyAlert



Thanks to  
CopyAlert

CopyAlert

# System Overview

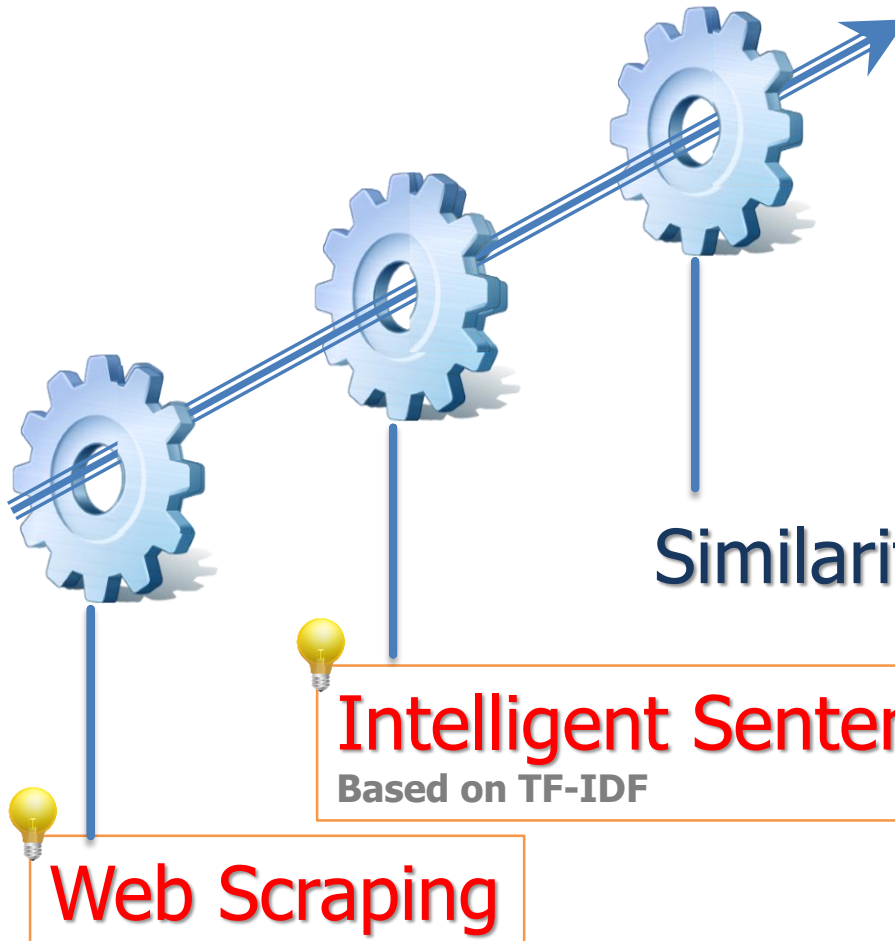


# Process of CopyAlert

## Analyzed Results



Webpage

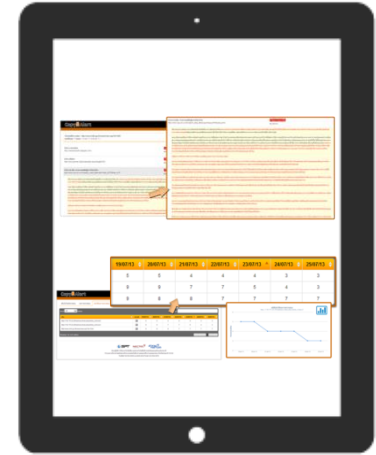


Web Scraping

Intelligent Sentence Selection

Based on TF-IDF

Similarity Calculation



# Webpage

The screenshot shows the TECHINASIA website with several elements highlighted for scraping:

- Navigation Bar:** Includes links for STARTUP, MOBILE, SOCIAL, E-COMMERCE, OPINION, ASIA, and GAMES IN ASIA. Social media icons for RSS, Twitter, Facebook, LinkedIn, and Google+ are also present.
- Header:** Features a "Save money on data charges" banner and a "GET SITE-SIZED NEWS @TECHINASIA" button.
- Main Article:** Titled "Google Play's app vending machine could work well in Southeast Asia" by Willis Wee, dated October 2, 2013. The article discusses the Google Play Android app vending machine in Tokyo, Japan, and its potential for Southeast Asia.
- Image:** A photo of a vending machine with colorful app icons (C3, BUY APPS HERE!) is shown.
- Sidebars:** The right sidebar contains sections for "Tech In Asia Online Community" (with a newsletter sign-up), "Featured Posts" (including "Scratch cards and gift cards: is Asia missing out on a huge online payment opportunity?" and "The necessary evils of fixing Singapore's engineering talent shortage"), "Upcoming Event" (VN Premiere: Crocodile in the Yangtze (The Alibaba Story)), and "Trending Posts" (including "The necessary evils of fixing Singapore's engineering talent shortage", "10 startups that can beat PayPal in Southeast Asia", and "Huawei is quietly taking over the world, minus the US").
- Footer:** Includes a "Jump To Comments" button and a subscription form.

## Clean Context

### Google Play's app vending machine could work well in Southeast Asia

This article was *originally posted* on *Games in Asia*, our new sister site that's dedicated to bringing you all the best in Asia games and gaming. You should check it out.

Yesterday, we saw Google rolling out its Google Play Android app vending machine in Tokyo, Japan. Unlike the conventional vending machines, Google Play's machine can only be accessed through NFC-enabled smartphones. For the moment, that might mean it only works in highly mobile and developed countries like Japan.

I'm thinking. If these vending machines could accept cash or coins as payment, this concept could work very well in fast-developing nations like Indonesia where paying for stuff online is *still not widely accepted* and people don't always trust e-payment systems. Solid cash is still the main way people pay for things.

It's common to hear game and app developers in Southeast Asia complain about the broken e-payment infrastructure. Google isn't lacking for users given that Android is *already the top smartphone OS* in Indonesia. But all those users can't pay, unfortunately, because the Google Play store mainly accepts credit card payments, and many users in developing countries don't have credit cards.

So if Google could build Android app store vending machines that accept printed notes and/or coins and allow folks to download apps through a connecting cable, it could really help create a trustworthy channel for users to pay and gain access to games and apps. That would be great for the mobile games industry in Southeast Asia.

If folks were willing to buy good content through pre-paid phone credits a decade ago, then I think it is only fair to assume that people are willing to pay for good games and content on their Android devices. The only problem is that without an alternative to the Google Play store's restrictive payment system, they can't.

# Scraping

# Intelligence Text Selection Algorithm

We apply the intelligent text selection algorithm to extract only informative sentence to reduce the number of request and reduce execution time.

- Term weighting calculation based on TF-IDF
- Removing highly similar and redundant sentences based on Jaccard Similarity.

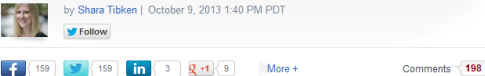
## Publication:

Thaiprayoon S., Haruechaiyasak C., "Web Plagiarism Detection Based on Search Result Snippets," Proceedings of the 28th International Technical Conference on Circuits/Systems, Computers and Communications, 2010.

# Intelligence Text Selection Algorithm (cont'd)

## Samsung Galaxy Round: Get bent, get buzz -- but get real

analysis The curved AMOLED display is eye-catching, but really Samsung, is that the only selling point? Good luck with that.



The Samsung Galaxy Round? Yeah, really curved. (Credit: Samsung)

Samsung's curved smartphone has generated a lot of buzz but not a lot of reasons to buy it -- other than the novelty effect.

The South Korean electronics giant late Tuesday unveiled the Galaxy Round, a smartphone similar to the Galaxy Note 3 but with a lengthwise dip in the middle. That dip forms a curve sort of like a smile or a hodge-bun shape, allowing the phone to fit the contours of a person's face or gently rock when placed on a flat surface. Samsung has released slightly curved phones in the past, such as the Galaxy Nexus, but this is the first device on the market to curve a "super flexible AMOLED" display, the company says. That doesn't mean the device can bend; rather, the Galaxy Round's curved display is fixed in place by the phone's rigid body.

Samsung has long followed the strategy of trying as many features and designs as possible to see what takes off. It also will include technology in devices simply for technology's sake, not because consumers really need those features. Samsung's top executives likely don't expect the Galaxy Round to be a huge seller -- hence its release only in Korea at this time -- but in order to attract more customers over the long run, they will have to articulate the benefits of such a design. So far, the only selling point for Samsung's curved smartphone appears to be that curve, not what that design actually allows users to do.

- Related stories:**
- CNET's take on the Samsung Galaxy Round
  - Samsung Galaxy Round brings curve to smartphones
  - Galaxy Gear: Samsung's big move to one-up Apple
  - How Samsung cranked out the Galaxy Gear: The inside story

buying mobile devices but not at the rapid clip of past quarters. That means companies like Samsung must find other ways to wow buyers and get them to fork over the money for new devices. Companies are releasing unique designs -- such as curved displays and buttons on the back of the phones (hello, LG G2) -- to differentiate themselves from the herd. But many of those tweaks are gimmicky at best.

**The real test of flexibility**

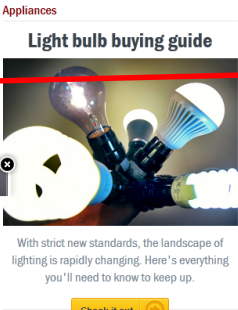
What may really catch consumers' attention -- and their dollars -- is a phone with a flexible device that actually could bend or roll up. It's not clear just what the first flexible devices may eventually be like, but that capability, unlike a slight curve, could really change the mobile device market. Samsung has shown concept devices with truly flexible screens, and it's undoubtedly prepping one for release at some point. But as CNET has predicted, consumers will just have to wait a bit longer. The Galaxy Round is not that sort of device.



- Featured lists**
- GTA V smashes several Guinness World Records
  - Google promotion offers free HP Chromebooks
  - Samsung Galaxy Round: Get bent, get buzz -- but get real
  - Apple's iPad event reportedly set for October 22
  - Steve Jobs action figure sports youthful '70s mop top

- Most Popular**
- T-Mobile to offer free unlimited international data, texts
  - Smartphones with long battery life (rounding)
  - Nexus 5 16GB version reportedly will sell for \$299

- Connect With CNET**
- Facebook Like Us
  - Twitter Follow Us
  - Google+ Add Us
  - RSS Subscribe
  - Mobile Download Now
  - Newsletters Subscribe Now



- Sentence#1**
- This is another factor that's hard to determine before the device hits the market.
- Sentence#2**
- Samsung has show concept devices with truly flexible screens, and it's undoubtedly prepping one for release at some point.
- Sentence#3**
- One feature, the "roll effect," lets users tilt the phone on its axis to trigger a glance screen with the time, date, missed call, and battery information.
- Sentence#4**
- The South Korean electronics giant late Tuesday unveiled the Galaxy Round, a smartphone similar to the Galaxy Note 3 but with a lengthwise dip in the middle.
- Sentence#5**
- The South Korean electronics giant late Tuesday unveiled the Galaxy Round, a smartphone similar to the Galaxy Note 3 but with a lengthwise dip in the middle.

## Term weighting calculation based on TF-IDF (Term frequency-inverse document frequency)

### Sentence#1

This is another factor that's hard to determine before the device hits the market.

Information score  
0.5

### Sentence#2

Samsung has show concept devices with truly flexible screens, and it's undoubtedly prepping one for release at some point.

Information score  
0.95

### Sentence#3

One feature, the "roll effect," lets users tilt the phone on its axis to trigger a glance screen with the time, date, missed call, and battery information.

Information score  
2.78

### Sentence#4

The South Korean electronics giant late Tuesday unveiled the Galaxy Round, a smartphone similar to the Galaxy Note 3 but with a lengthwise dip in the middle.

Information score  
1.89

### Sentence#5

The South Korean electronics giant has launched the Galaxy Round, a variant of the Galaxy Note smartphone but with a lengthwise dip in the middle.

Information score  
1.56

# Removing highly similar and redundant sentences

## Sentence#1

This is another factor that's hard to determine before the device hits the market.

Information score  
0.1

## Sentence#2

Samsung has show concept devices with truly flexible screens, and it's undoubtedly prepping one for release at some point.

Similarity score  
0.2

## Sentence#3

One feature, the "roll effect," lets users tilt the phone on its axis to trigger a glance screen with the time, date, missed call, and battery information.

Similarity score  
0.1

## Sentence#4

The South Korean electronics giant late Tuesday unveiled the Galaxy Round, a smartphone similar to the Galaxy Note 3 but with a lengthwise dip in the middle.

Similarity score  
1.0

## Sentence#5

~~The South Korean electronics giant has launched the Galaxy Round, a variant of the Galaxy Note smartphone but with a lengthwise dip in the middle.~~

~~Similarity score  
1.0~~

# Features



**Inserting Small Script**



**Email Alert**



**Thai and English Languages**



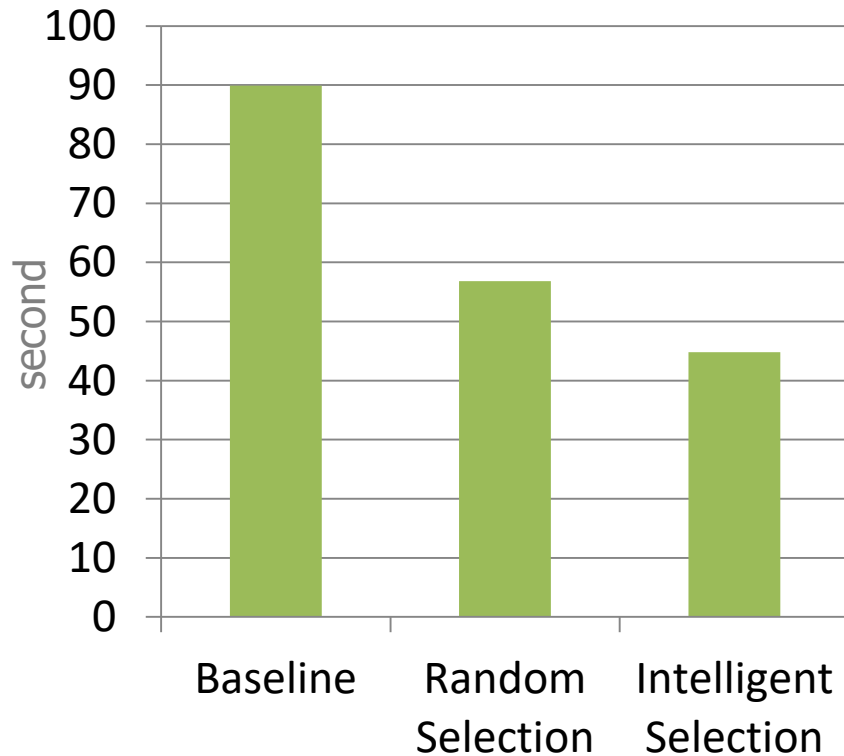
**Monitoring Frequency**



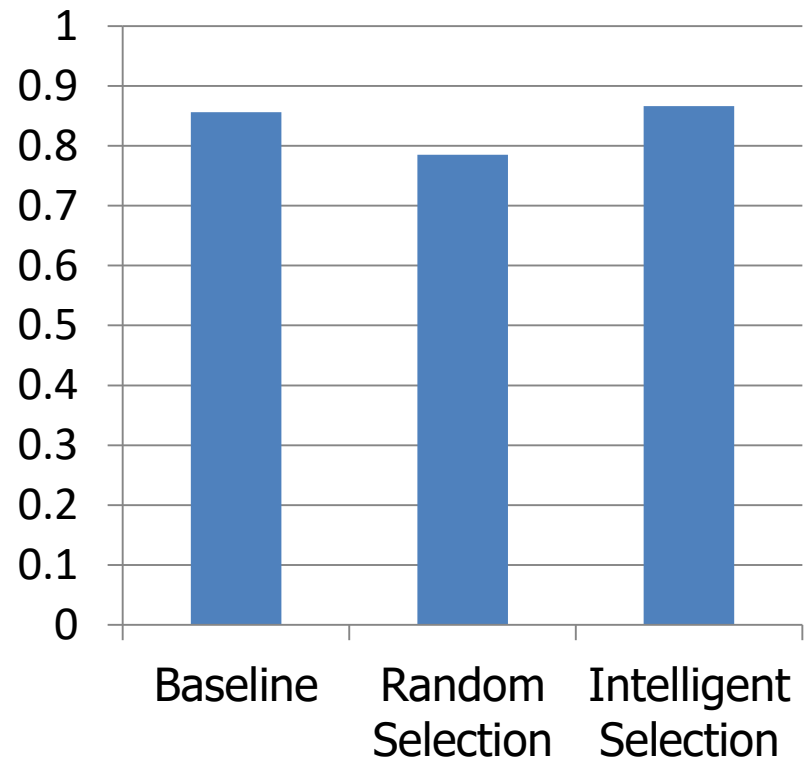
**Report Detail Similarity**

# Performance Evaluation

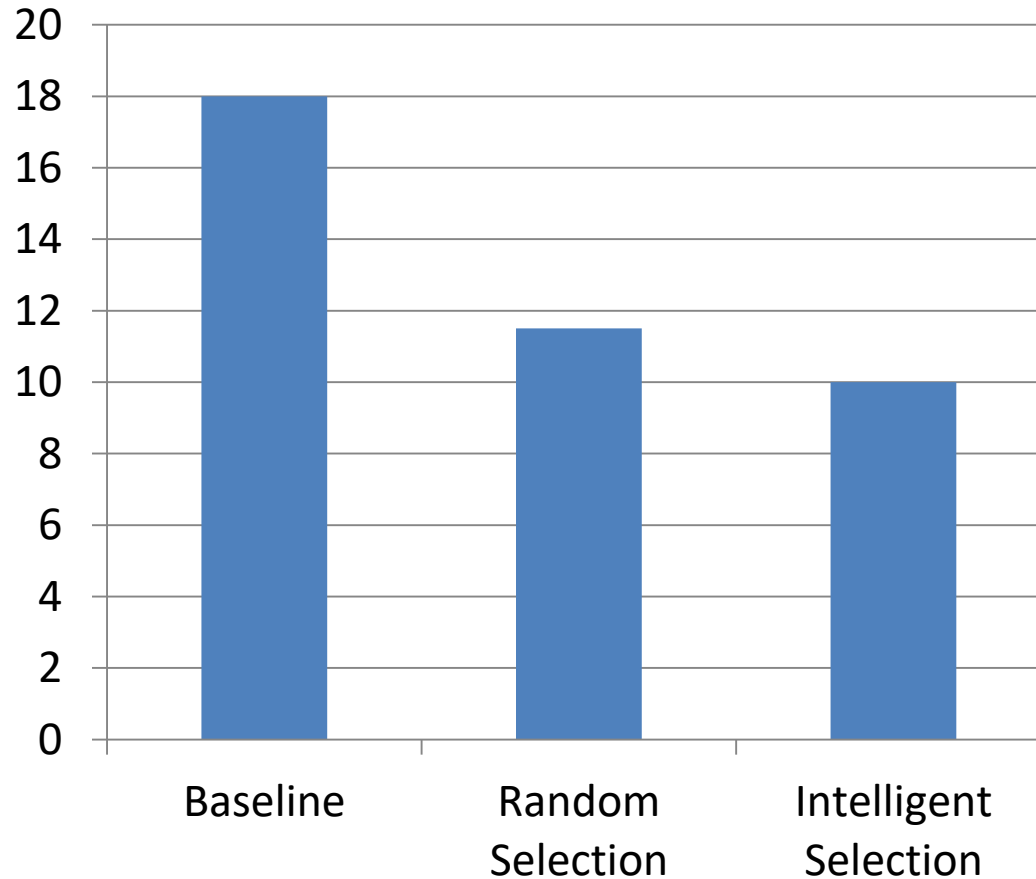
## Response Time



## F-measure



























# Cost Estimation



\*Cost estimation based on Google Search API:  
\$5 per 1000 search queries

# Comparison of Web Plagiarism Detection Tools

Products/Features	CopyAlert 	Copyscape	CopyGator 	PLAG SPOTTER 
Thai-English Language Support				
Automatically Monitor Pages				
Send Notifications by Email				
Embeddable Tag				
Document Highlight and Similarity Score				
Intelligence Text Selection		--	--	--

## Strength

- Support Thai language
- Support segmented and unsegmented languages
- Apply the intelligent text selection algorithm

## Opportunity

- CopyAlert is based on our ongoing R&D project named **CopyCat** which has been successfully applied to detect plagiarism
- Potential customers of CopyAlert include copyrighted content owners
- CopyAlert helps build public awareness of copyright protection and plagiarism problem

# Commercial Opportunity

## Copy👤Alert Standard

- ✓ Check your pages every month
- ✓ Alert email
- ✓ Similar text highlight
- ✓ Similarity score
- ✓ Overall statistics



### Potential customer

- Freelance writer
- Blogger



## Copy👤Alert Professional

- ✓ Check your whole site
- ✓ Customize monitoring frequency
- ✓ Alert email
- ✓ Similar text highlight
- ✓ Similarity score
- ✓ Overall statistics



### Potential customer

- Publishing companies
- News provider



## Copy👤Alert API

- ✓ Customize request
- ✓ Return output results in JSON, XML or HTML formats.



### Potential customer

- Software integrator



**Thank you for your attention**

**Questions / Comments ?**