DaaS (Data as a Service) for National Assembly Libraries

Prof. Sam Oh
Sungkyunkwan University(SKKU) iSchool, LIS and Data Science
iSchool Caucus Chair-Elect, JTC1/SC34 Chair

2017 **APLAP** Conference

Convergence

and

Connectivity

Generating creative ideas

National Assembly Library:

Quality Data

- Approximately 5.9 million books
- Over 200 million pages of original texts
- Legal Deposit Library (Article 7 of the National Assembly Library Act)
- Most Extensive National Authority File (NAF)
- International Standard Name Identifier (ISNI)-RAG (Integrating ISNI into NAF)
- Ongoing thesaurus development
- Various national data systems, including reference services for legislative and legal information

- **New services** generated through convergence and **connectivity between library data and social data**
- Subsequent enhancement of entry points and increase in utilization

No service meets every user’s needs

Target Users Through User-Specific Services

General Public

Assembly Members

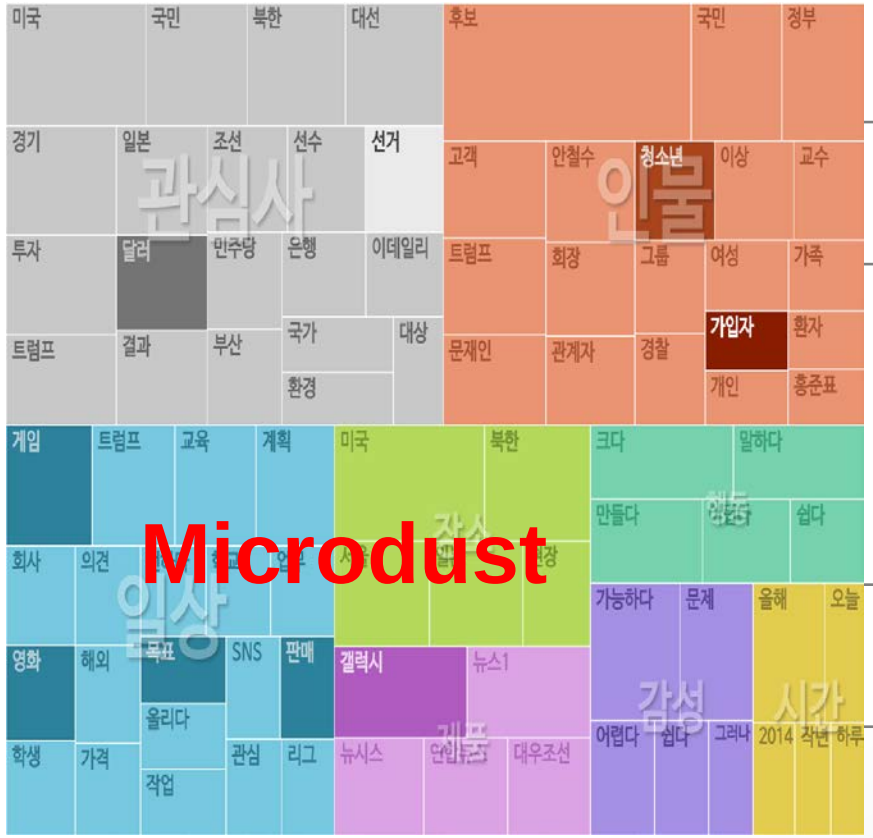
Researchers

What kind of DaaS
for
the General Public?

Social Media Analysis Services

General Public

Particularly on Political Matters

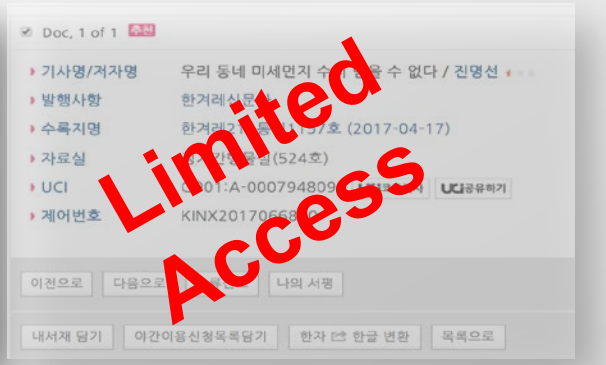


Real-time keyword analysis

View social data

Links to **social data** on a specific keyword

View library data



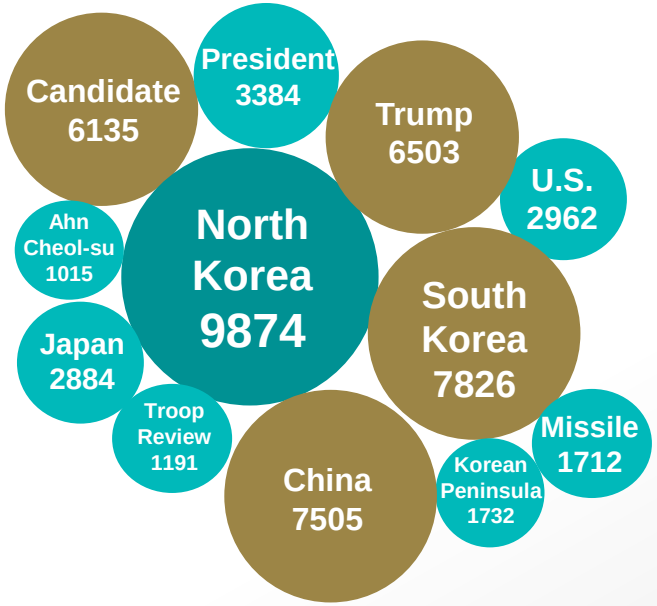
Links to **library data** on a specific keyword

Text Analysis Services

General Public

Sourced from library collections

Keywords extracted from full text and analyzed by subject



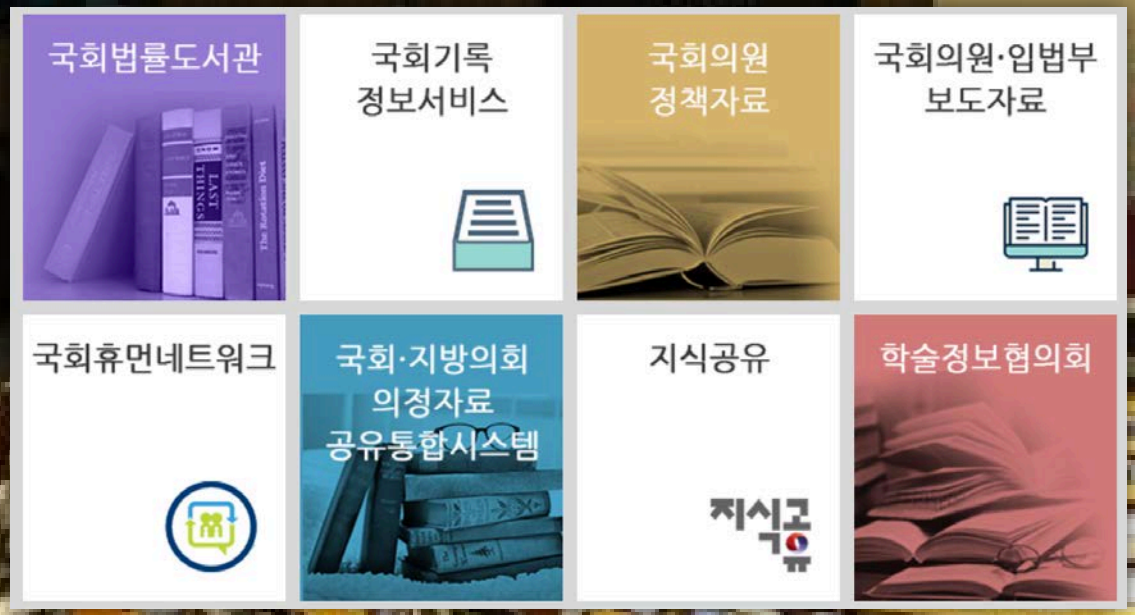
* Numbers derived from occurrences on Odpia

Text Search and Linkage

National Assembly Library

General Public

Unique Data



- Assembly member information
- National Assembly minutes
- Policy data (by member, party, subject, or National Assembly meeting)
- Seminar data (by date or venue)
- Assembly member news releases

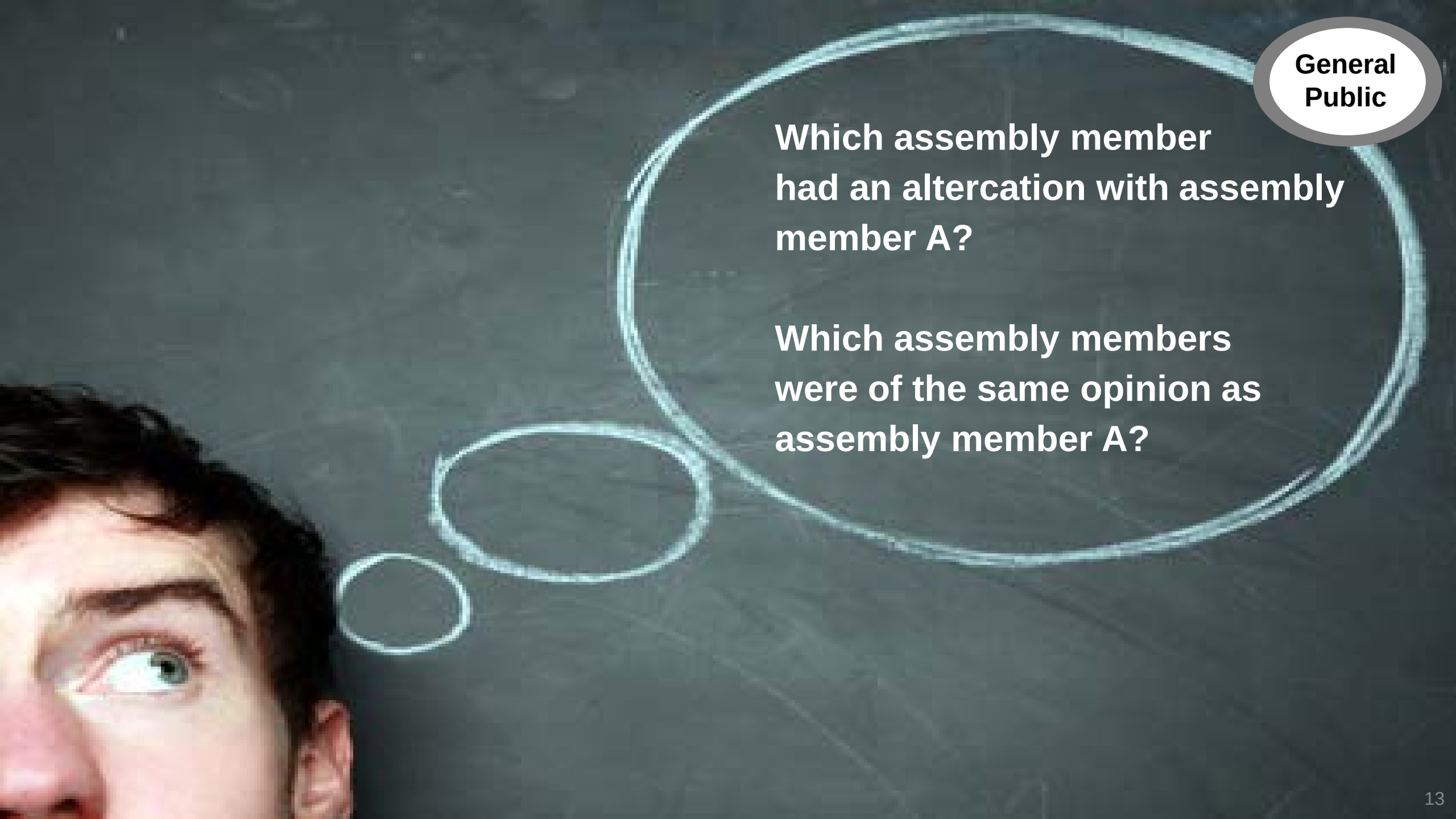
General Public

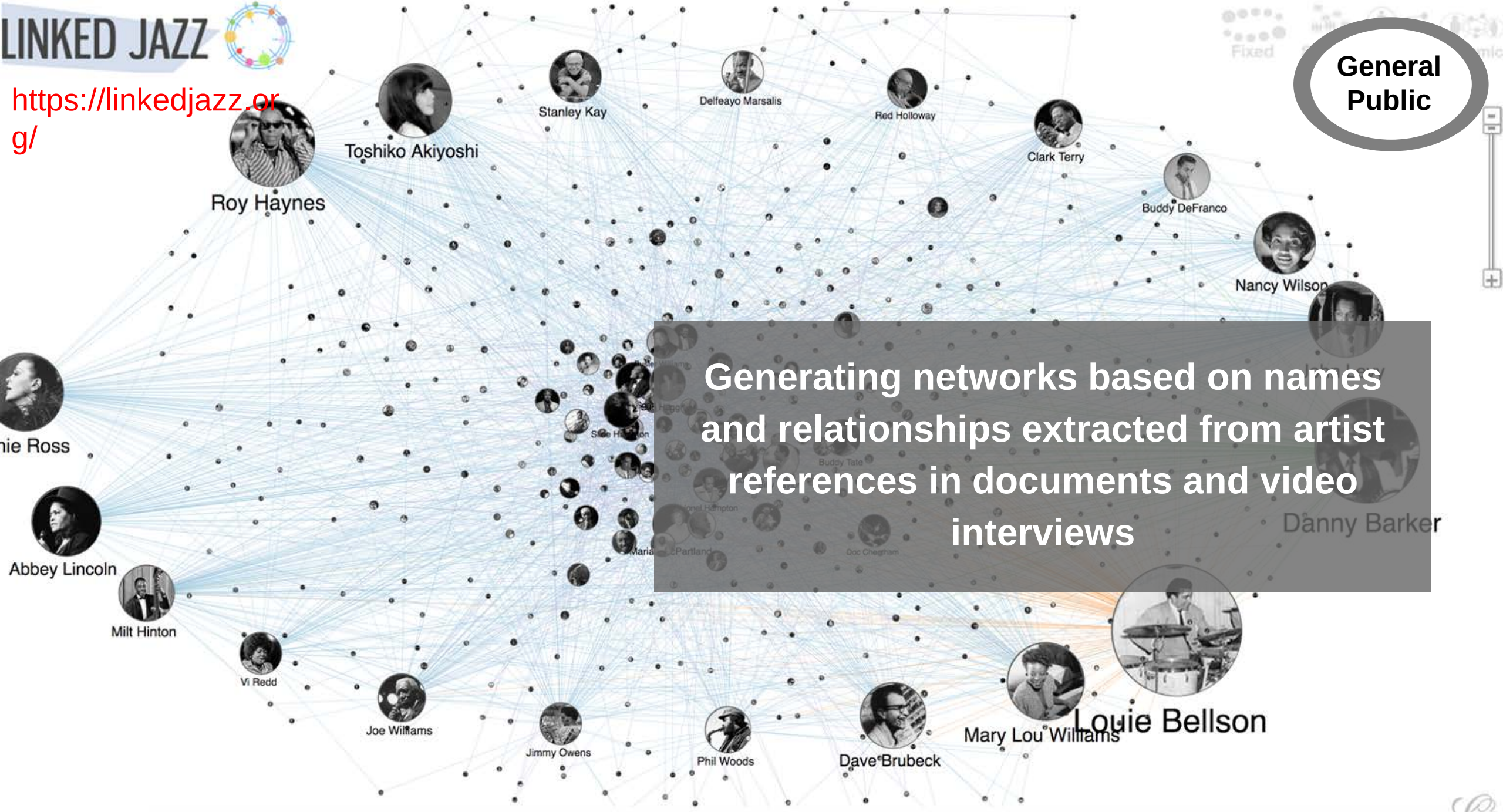
Human Network

Beyond Biographical Data

General Public

Generating networks out of the highly organized meta information provided by the National Assembly library


General Public
Which assembly member had an altercation with assembly member A?
Which assembly members were of the same opinion as assembly member A?


LINKED JAZZ
https://linkedjazz.org/
General Public
Roy Haynes
Toshiko Akiyoshi
Stanley Kay
Delfeayo Marsalis
Red Holloway
Clark Terry
Buddy DeFranco
Nancy Wilson
Generating networks based on names and relationships extracted from artist references in documents and video interviews
Abbey Lincoln
Milt Hinton
Vi Redd
Joe Williams
Jimmy Owens
Phil Woods
Dave Brubeck
Mary Lou Williams
Louie Bellson
Danny Barker

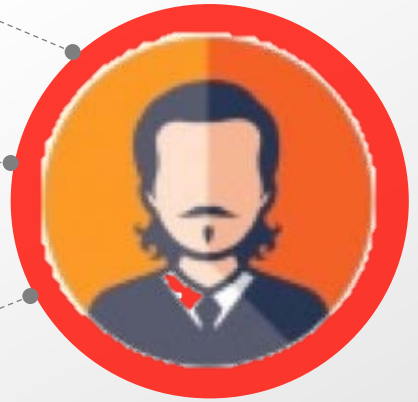
Legislative data analysis

General Public

Supported Bill A

Altercation over Bill D

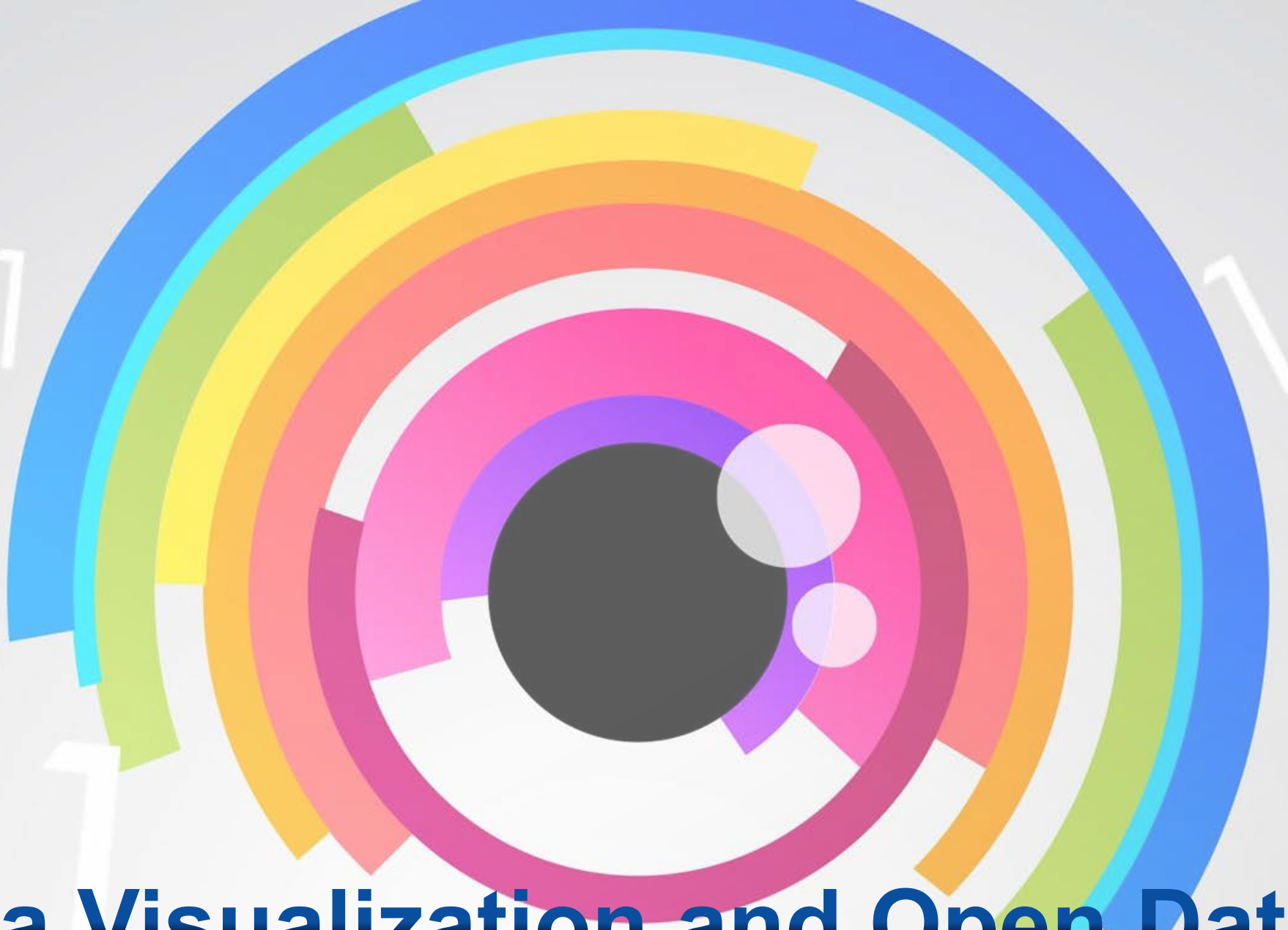
Abstained on Bill C



Supported Agenda E

Opposed to Bill B

General Public



Data Visualization and Open Data

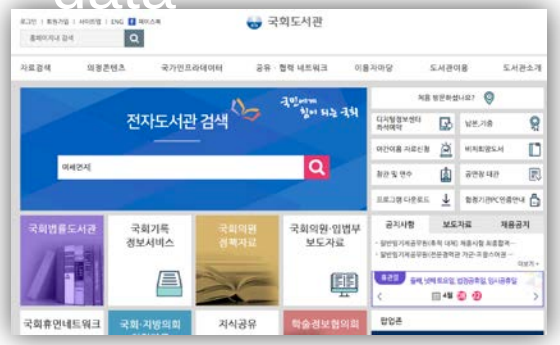
Example of Viz&Open Data - CodeNamu

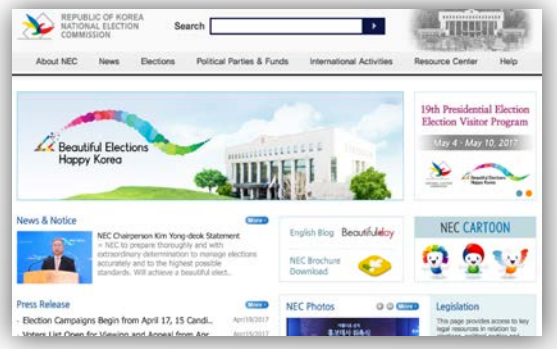
(http://codenamu.org)

General Public

Civic Hacking Community

Data source: Public open data

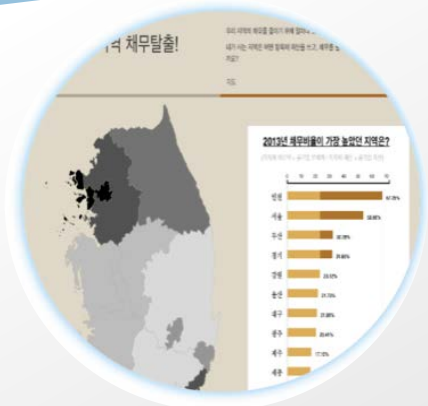




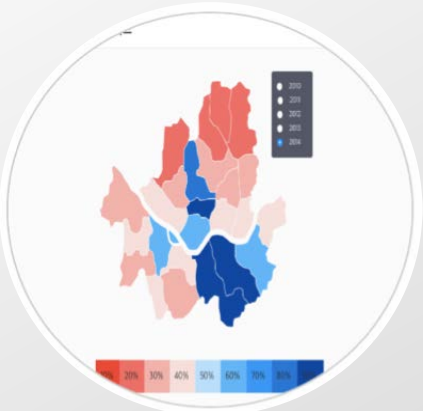
Data refinement and visualization

Assets of high-ranking government officials

Report cards for members of the 19th National Assembly



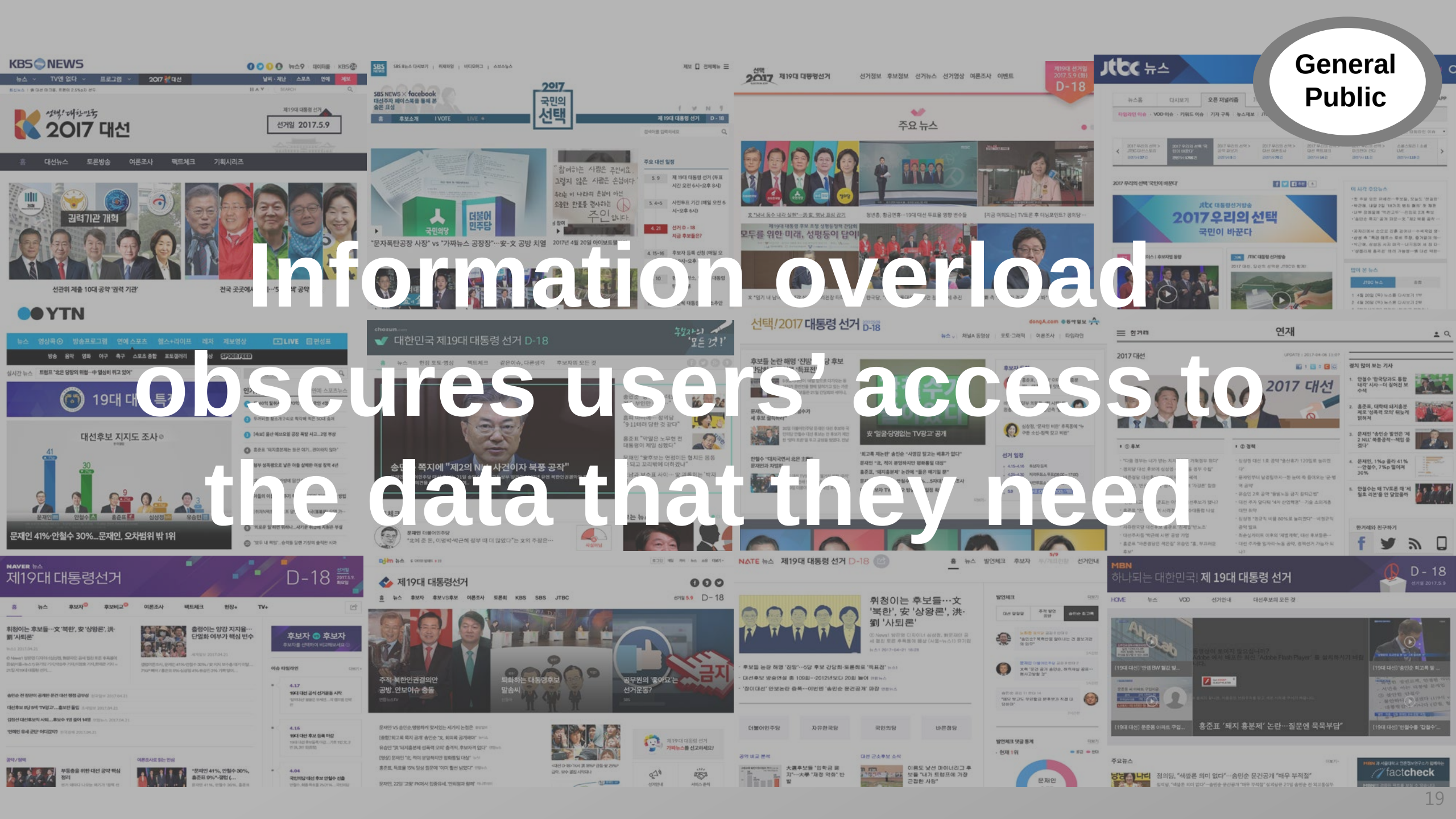
Regional debt settlement



Financial independence in the provinces

General Public

DaaS for Elections


General Public
Information overload
obscures users' access to
the data that they need

Reliable Data on Candidates

True or False on a Certain Issue

General Public

Standard information

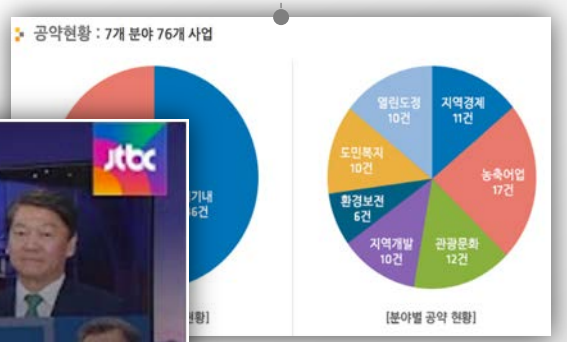
Related articles

Minutes

Legislation

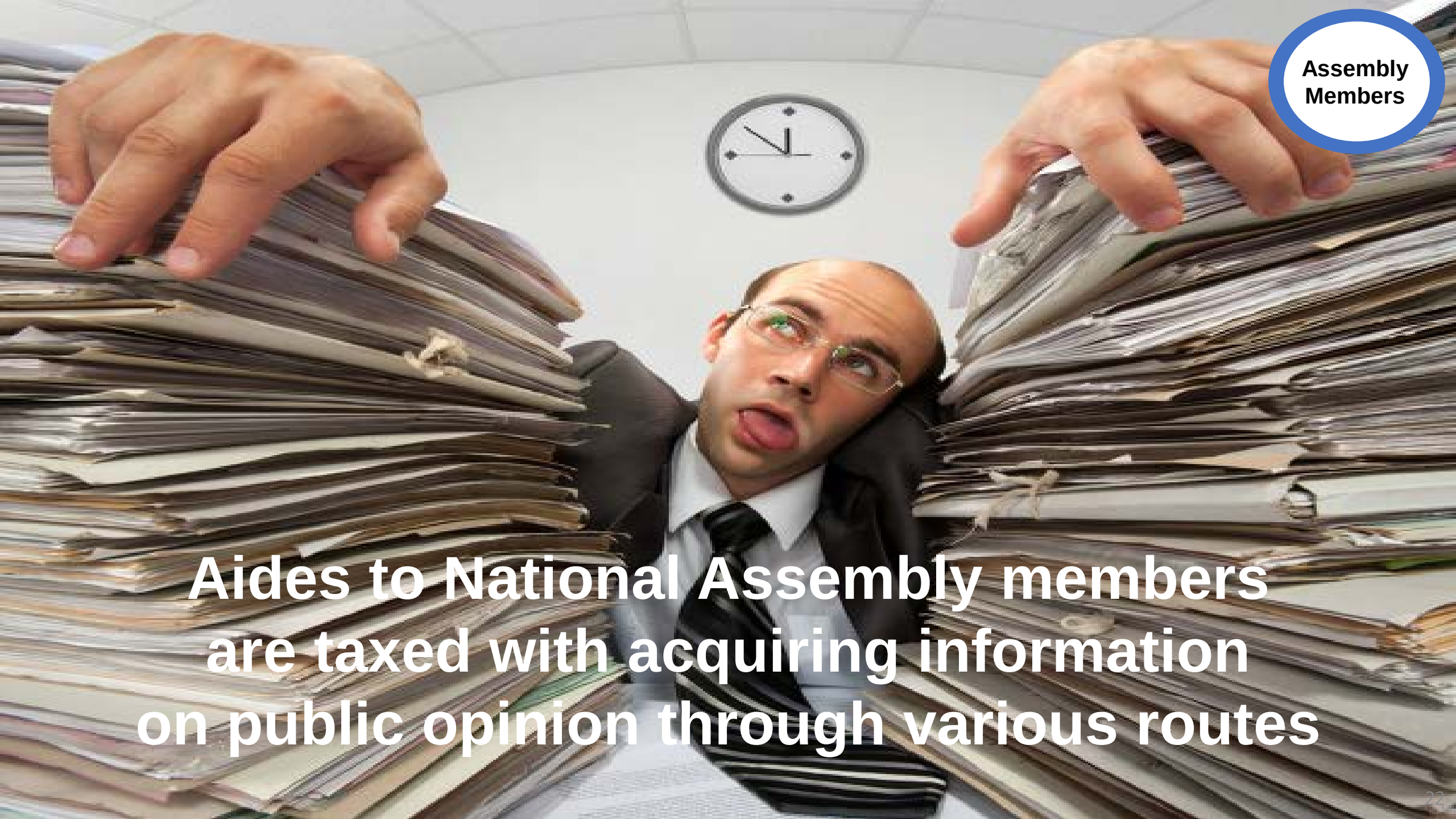
Fulfillment of campaign pledges

Related persons



Assembly Members

What Kind of DaaS for Assembly Members?


Assembly Members
Aides to National Assembly members are taxed with acquiring information on public opinion through various routes

Public Concerns in a Nutshell

Assembly Members

Clustering social data through appropriate classification systems

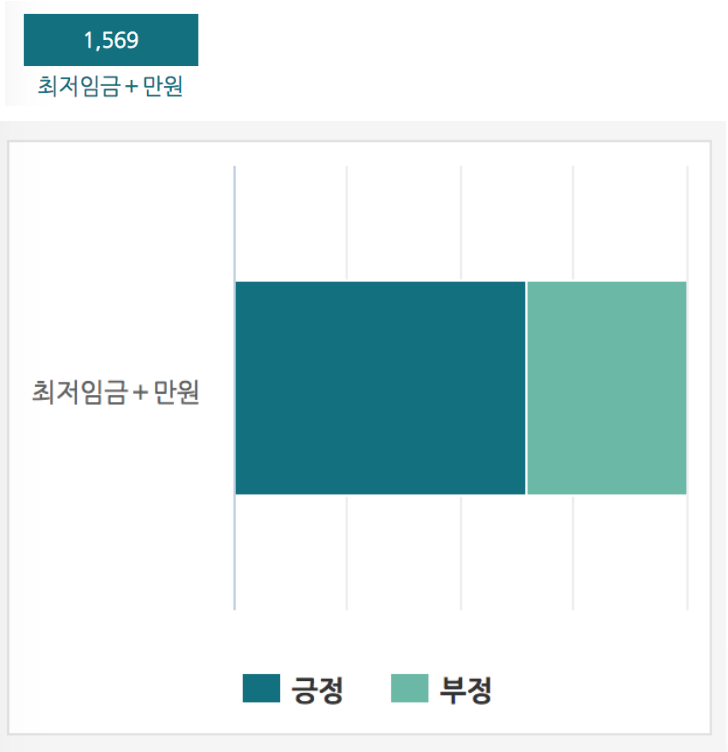
Discovering key issues by subject within each subfield

Speedy keyword-based access to original texts

Monitoring Public Opinion in Real Time

Assembly Members

Feedback on new law initiatives and policy implementations
Identification of the public's needs



긍정 연관어

순위	키워드	문서수	키워드 빈도	중요도 TF-IDF
1	책임	800	946	-195
2	강력	603	619	211
3	가능하다	350	883	1,012
4	필요하다	264	795	1,195
5	자유	227	535	901
6	필요	209	307	547
7	협의	179	248	486

부정 연관어

순위	키워드	문서수	키워드 빈도	중요도 TF-IDF
1	위반	749	862	-56
2	제재	737	2,100	-67
3	신고	732	802	-14
4	불편	713	754	26
5	징역	697	1,335	105
6	적발	675	1,298	181
7	제재하다	664	679	115

Raising the **minimum wage** to 10,000 won has become a widely debated issue

Public opinion monitoring indicates a preference for more qualitative and long-term approaches such as **tackling unemployment**

What Kind of DaaS for Researchers?

Initiative for Researcher Information Sharing

Researchers

Mega-Convergence Linkage Services

Providing a foundation for data convergence
Sharing researchers' authority files

Formulating author authority data in accordance with international standards

Linking to governmental academic data via cloud
The country's largest joint integration of academic data (National Assembly library, Ministry of Science, NIA)
LOD publication, big data analysis

Establishing a system for researcher information sharing
Linking the author data of partner institutions
Author matching identification
Issuing identification numbers (ISNI)

Establishing collaborative bodies
Developing models for sharing/management
Research on the use of author authority data
Researcher support

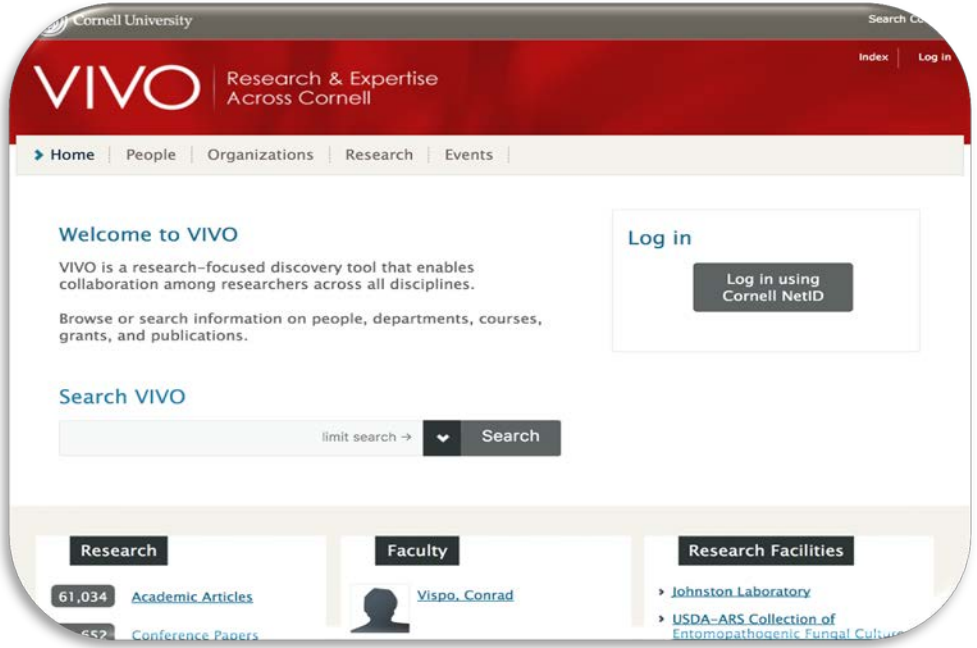
Researchers

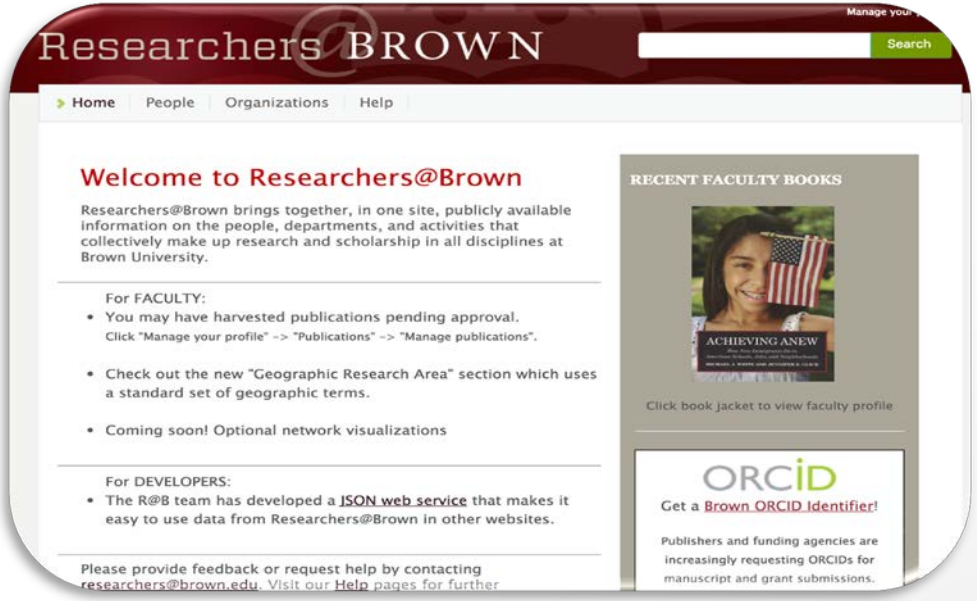
VIVO

A research-focused discovery tool that enables collaboration among researchers across all disciplines

Connect Share Discover

Researchers





Researchers

National Assembly libraries already have raw data that facilitates the extraction of research keywords, tables, and graphs

Researchers

Collaborations and top keyword concepts used within the past 5 years.

Click dots and donuts to bring up details.

Providing new networks by sharing researcher/research information

EPUBLIC OF

Korea University
64 shared Research Output

Seoul National University
38 shared Research Output

Yonsei University
22 shared Research Output

Korea Research Institute of Chemical Technology
20 shared Research Output

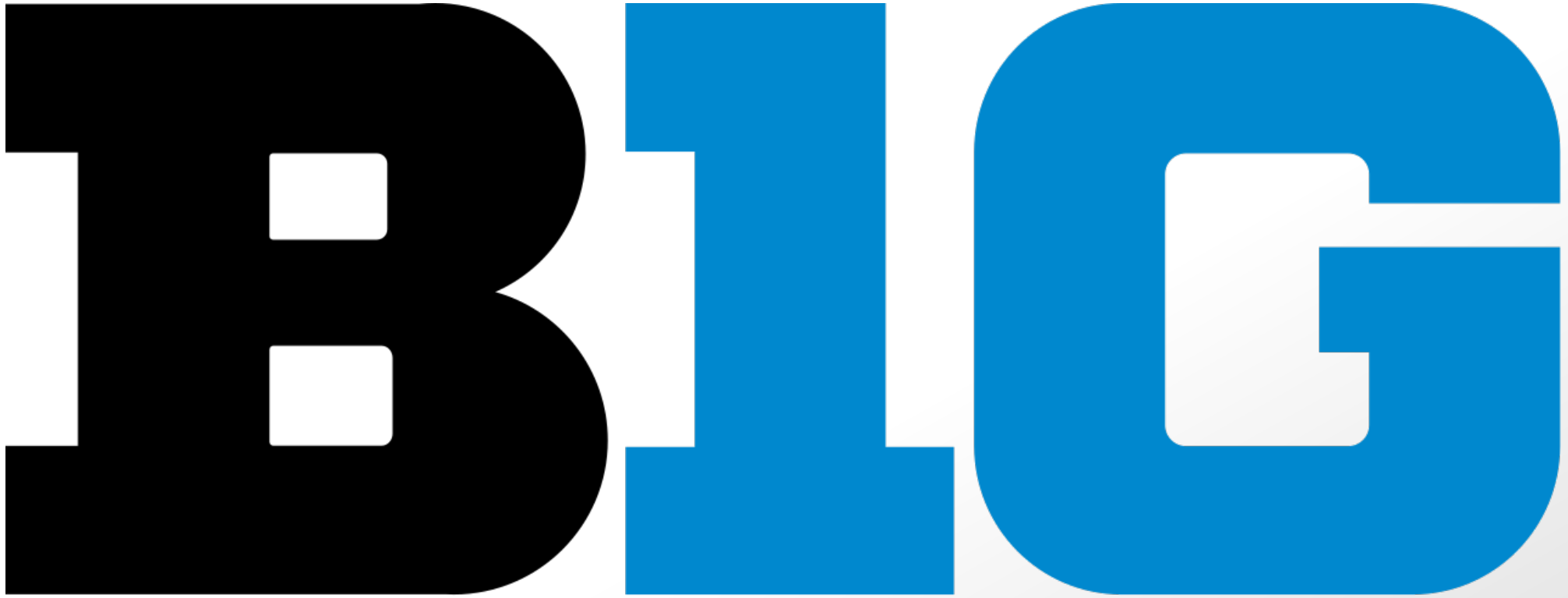
Sungkyunkwan University
16 shared Research Output

Health
MEDICINE & LIFE SCIENCES

Education
MEDICINE & LIFE SCIENCES

Logistic Models
MEDICINE & LIFE SCIENCES

De
ME


B1G

Big Data, Small Math

D. J Patil

Many big data sets are not being analyzed

We need to pay more attention to the complexity of big data
(clear points of reference; credible, high-quality data)

What makes big data important?

Value

Small data has just as much purpose

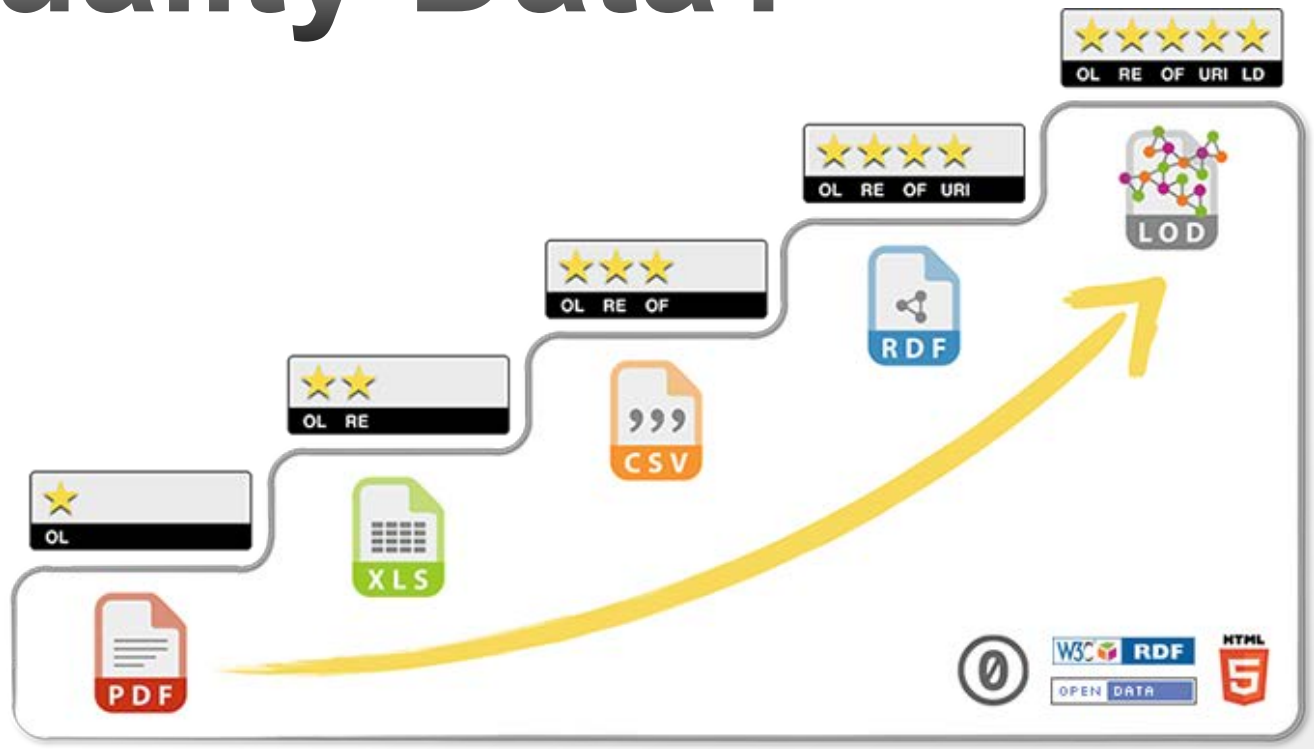
Essential Ingredients for DaaS

- What is the role of the National Assembly Library in DaaS?
- What is the National Assembly Library best equipped to provide with regard to DaaS?

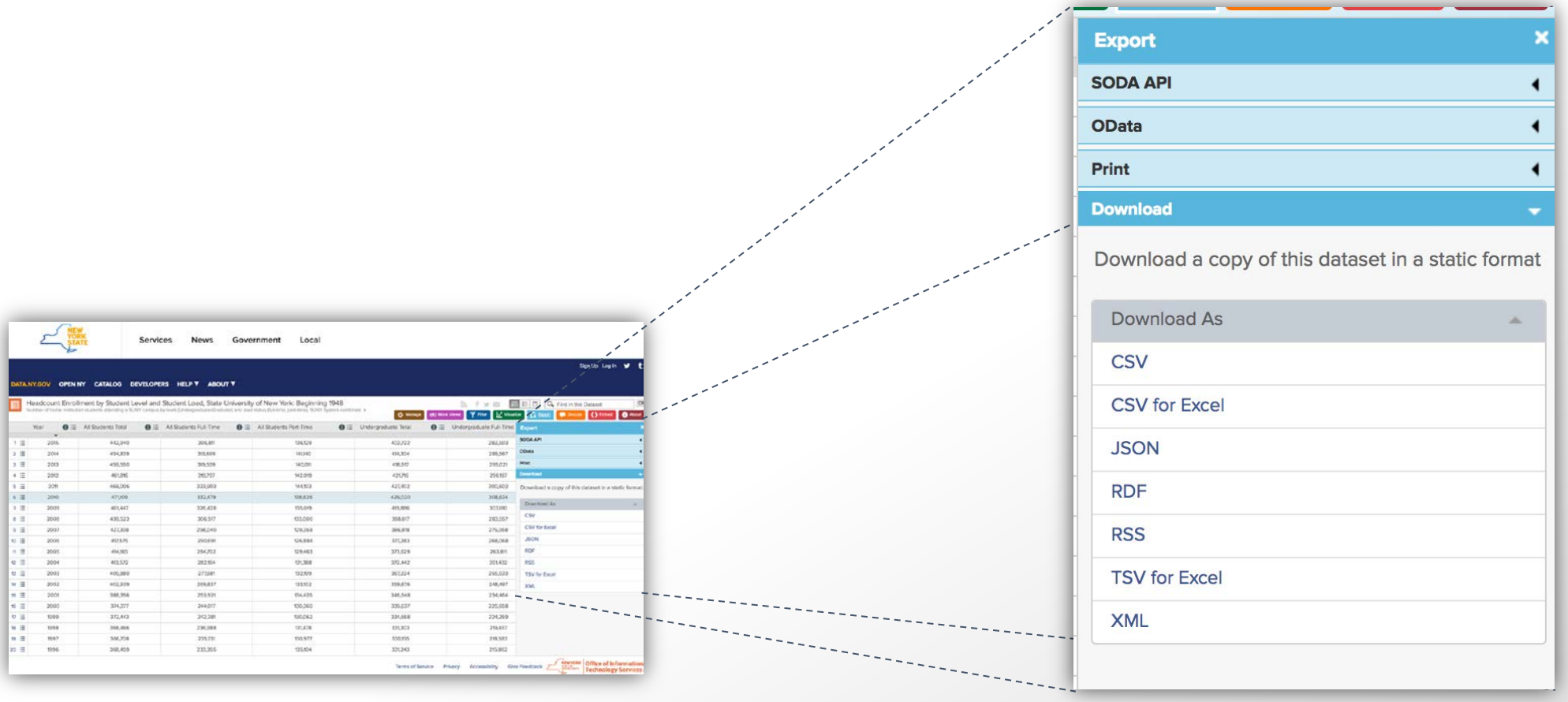
National Assembly Libraries have an obligation to provide **quality data**

Quality data enhances CONVERGENCE and CONNECTIVITY

Quality Data?



Providing necessary data in various forms



Communication Between Involved Parties

NAL should provide
a playground,
platform for diverse
users

Library Data
Social Data
National Assembly Data
Open Government Data
News Data

Data outside the National Assembly Library only seems bigger and more appealing

Users seek out National Assembly Libraries because of an interest in their unique data

The purpose of DaaS is to make
our small data More valuable
and share it across a wider spectrum

Thank you for listening!

Prof. Sam Oh, SKKU iSchool, LIS and Data Science

iSchool Caucus Chair(2018-2019), ISO/IEC JTC1/SC34 Chair, DCMI Board Member

Image Reference

- Slide 3: http://cfile7.uf.tistory.com/image/225C7F4257F3087A2633CB
- Slide 4: http://magazine.hankyung.com/jobnjoy/apps/news?popup=0&nid=05&nkey=2012051000025112945&mode=sub_view
- Slide 7: https://image.freepik.com/free-vector/people-search_23-2147502905.jpg
- Slide 8: http://pruti.tistory.com/entry/%ED%9E%98%EC%9D%B4%EB%93%9C%EB%8B%88
- Slide 11, 22: http://i.dailymail.co.uk/i/pix/2016/03/02/article-urn:publicid:ap.org:305733e99b1a4195a5fde07bb98ac1bf-4TibiYnljpbda813bc04ce986925-212_634x454.jpg
- Slide 12: http://www.instantshift.com/2011/08/25/prepare-for-your-comfortable-death-with-these-online-tools/
- Slide 14: https://blogs.oracle.com/oraclehcm/10-ways-to-move-towards-modern-hr
- Slide 16: http://www.dtelepathy.com/blog/wp-content/uploads/2014/10/Betterment_DataVisualization.jpg
- Slide 22: https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0ahUKEwjIuvTGyLzTAhVROrwKHRsVCqsQjRwIBw&url=http%3A%2F%2Fsoftwaredev.tistory.com%2Fm%2Fentry%2F%25ED%2595%259C%25EA%25B5%25AD%25EC%259D%2598-%25EA%25B0%259C%25EB%25B0%259C%25EC%259E%2590%25EB%258A%2594-%25EC%2593%25B8%25EB%258D%25B0%25EC%2597%2586%25EC%259D%25B4-%25EB%25B0%2594%25EC%2581%2598%25EB%258B%25A4&psig=AFQjCNGrb9bYZt1dW6aFJwczGkjOgCqB6g&ust=1493105223870747
- Slide 25: https://www.ucc.ie/en/media/academic/schoolofpharmacy/images/postgradscroppedbanner.jpg
- Slide 30: http://www.ccts.uic.edu/news/researcher-seeking-collaborator
- Slide 34, 39: ttp://www.v12data.com/wp-content/uploads/drupal-default-files/daas_0.jpg
- Slide 35: hhttp://www.kimdirector.co.kr/bbs/data/webplan/20150626_con_img002.jpg
- Slide 36: https://www.trinitymaxwell.com/wp-content/uploads/2016/01/mobile-connectivity.jpg
- Slide 37: http://5stardata.info/images/5-star-steps.png
- Slide 40, 41: https://pixabay.com/ko/%EA%B3%B5%EC%9B%90-%EC%A0%95%EA%B8%80-%EC%B2%B4%EC%9C%A1%EA%B4%80-%EB%86%80%EC%9D%B4%ED%84%B0-%ED%96%87%EB%B9%9B-%EB%8F%84%EC%8B%9C%EC%9D%98-%EB%8F%99%EB%84%A4-%EC%9D%BC%EC%B6%9C-1608727/