Consuming Norwegian Linked Open Data

Applications in Regional Development and Environmentally Friendly Behavior

Dumitru Roman, SINTEF, Norway

Jens Kilde Mjelva and David Norheim, Computas AS, Norway

dumitru.roman@sintef.no, jens.kilde.mjelva@computas.com, david.norheim@computas.com

@DBKDA2012

Outline

- Open Data Movement
  - Linked Open Data

- Norwegian Linked Open Data – Use Cases
  - Monitoring Regional Development
  - Environmentally Friendly Behavior

- Requirements
- Architecture
- Data Sets
- Applications

- Summary

Open Data Movement

- Open Government - establish a modern cooperation among politicians, public administration, industry and private citizens by enabling more transparency, democracy, participation and collaboration
- The Open Government Partnership (launched on September 20, 2011)

- 8 founding governments
- 43 national governments commitments to OG

http://www.opengovpartnership.org/

- Key enablers: free access to information and the possibility to freely use and re-use this information => Open Government Data (OGD)

Open Data Movement (cont')

- "Open Government Data" - worldwide movement to open up government/public administration data
  - Targeted to both human and machine-readable non-proprietary formats - for re-use

- "Open Data"
  - Data beyond just governmental institutions
  - Includes data from relevant stakeholder groups (e.g. citizens, industry, NGOs, science or education, etc.)

- Examples:
  - DIFI (http://data.norge.no/); UN (http://data.worldbank.org); Open Knowledge Foundation (http://okfn.org); New York Times (http://data.nytimes.com)

Open Government Data Principles

(http://sunlightfoundation.com/policy/documents/ten-open-data-principles/)

1. Data must be complete
2. Data must be primary
3. Data must be timely
4. Data must be accessible
5. Data must be machine-processable
6. Access must be non-discriminatory
7. Data formats must be non-proprietary
8. Data must be license-free
9. Data must be permanently available
10. Usage Costs

From Open Data to Linked Open Data

- Crucial for data to be put into a context - new knowledge and more powerful services and applications
  - Interoperability and standards are key

- 5 Stars Model

★	Information is available on the Web (any format) under an open license
★★	Information is available as structured data (e.g. Excel instead of an image scan of a table)
★★★	Non-proprietary formats are used (e.g. CSV instead of Excel)
★★★★	URI identification is used so that people can point at individual data
★★★★★	Data is linked to other data to provide context

What are the costs and benefits of ★ web data?

As a consumer ...	As a publisher ...
✓ You can see it. ✓ You can print it. ✓ You can store it locally (on your hard drive or on an USB stick). ✓ You can enter the data manually into another system.	✓ It is easy to publish.

What are the costs and benefits of ★★ web data?

As a consumer, you can do everything that you could do with ★ web data, plus:	As a publisher ...
✓ You can directly process it with proprietary software to aggregate it, perform calculations, visualise it, etc. ✓ You can export it into another (structured) format.	✓ It is easy to publish.

What are the costs and benefits of ★★★ web data?

As a consumer, you can do everything that you could do with ★★ web data, plus:	As a publisher ...
✓ You do not have to pay for a format over which a single entity has exclusive control	✓ It is easy to publish.

What are the costs and benefits of ★★★★ web data?

As a consumer, you can do everything that you could do with ★★★ web data, plus:	As a publisher ...
✓ You can link to it from any other place, either on the web or locally. ✓ You can bookmark it. ✓ You can re-use parts of the data.	✓ You will need to invest some time slicing and dicing your data. ✓ You will need to assign URIs to data items and think about how to represent the data. ✓ You have fine-granular control over the data items and can optimise their access (e.g. load balancing, caching, etc.)

What are the costs and benefits of ★★★★★ web data?

As a consumer, you can do everything that you could do with ★★★★ web data, plus:	As a publisher ...
✓ You can discover new data of interest while consuming other information. ✓ You have access to the data schema.	✓ You will need to invest resources to link your data to other data on the web. ✓ You make your data discoverable. ✓ You increase the value of your data.

Taken from *Linked Open Data: The Essentials* (by F. Bauer, M. Kaltenböck), in turn based on http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/

Linked Data Principles

1. Use URIs to identify things

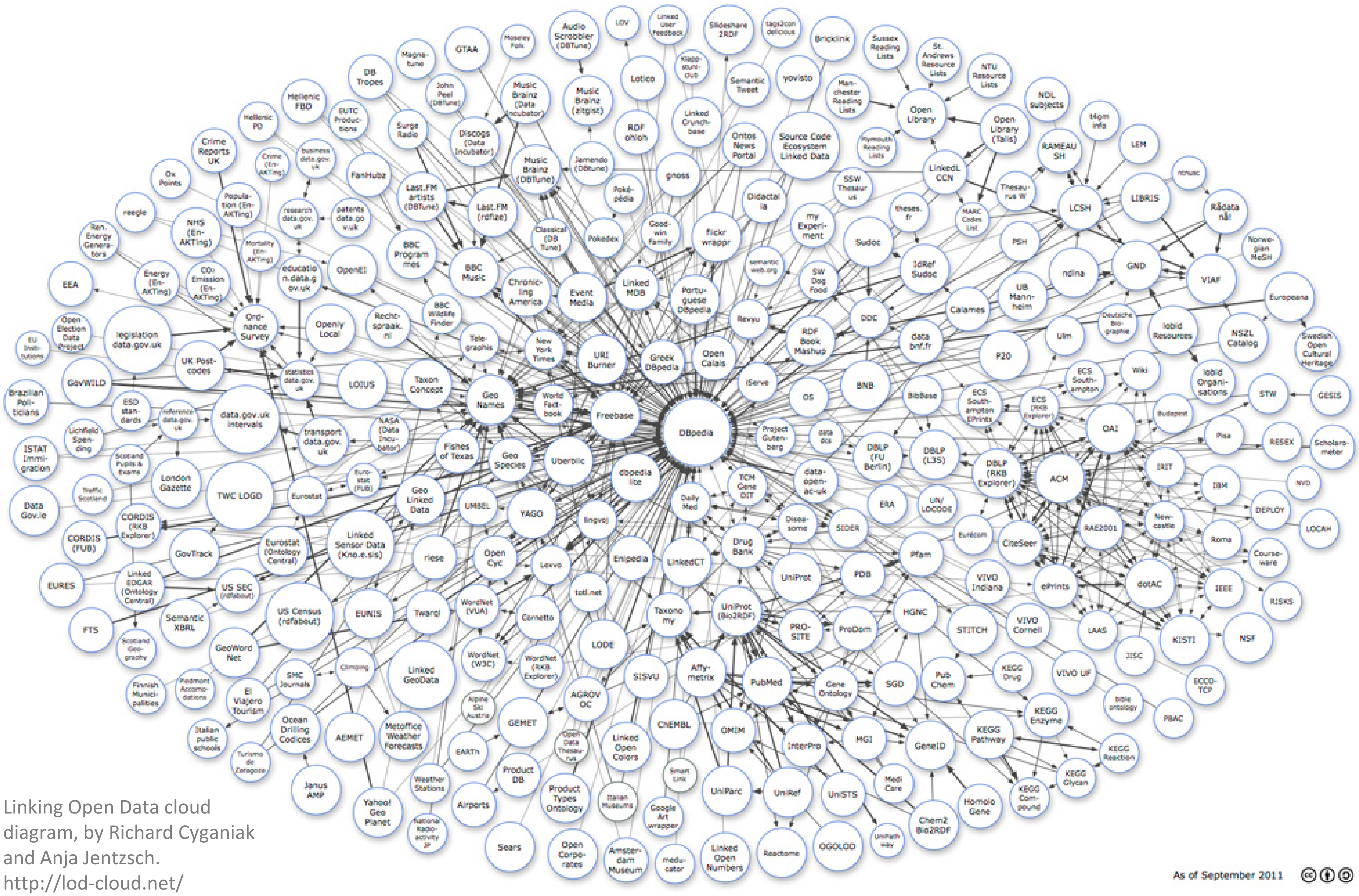
2. Use HTTP URIs so that these things can be referred to and looked up ("dereferenced") by people and user agents

3. Provide useful information about the thing when its URI is dereferenced, using standard formats such as RDF/XML

4. Include links to other, related URIs in the exposed data to improve discovery of other related information on the Web

http://www.w3.org/DesignIssues/LinkedData

LOD Cloud



Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch.
http://lod-cloud.net/

Outline

- Open Data Movement
  - Linked Open Data
- Norwegian Linked Open Data – Use Cases
  - Monitoring Regional Development
  - Environmentally Friendly Behavior
- Summary

The context (1): **PlanetData** Large-scale Data Management

- FP7 Network of Excellence (2010-2014)
- Aim: establish an interdisciplinary, sustainable European community on large-scale data management
  - Publishing and managing new species of interlinked data sets
  - Improving the usefulness of existing data sources
  - Data sets, vocabularies, best practices for publishing self-descriptive data
  - Portal with data provisioning and management tools
  - Training infrastructure, learning resources, summer schools, standards
- PlanetData Programs:
  - Call 1: “Consuming Linked Data”

http://planet-data.eu/

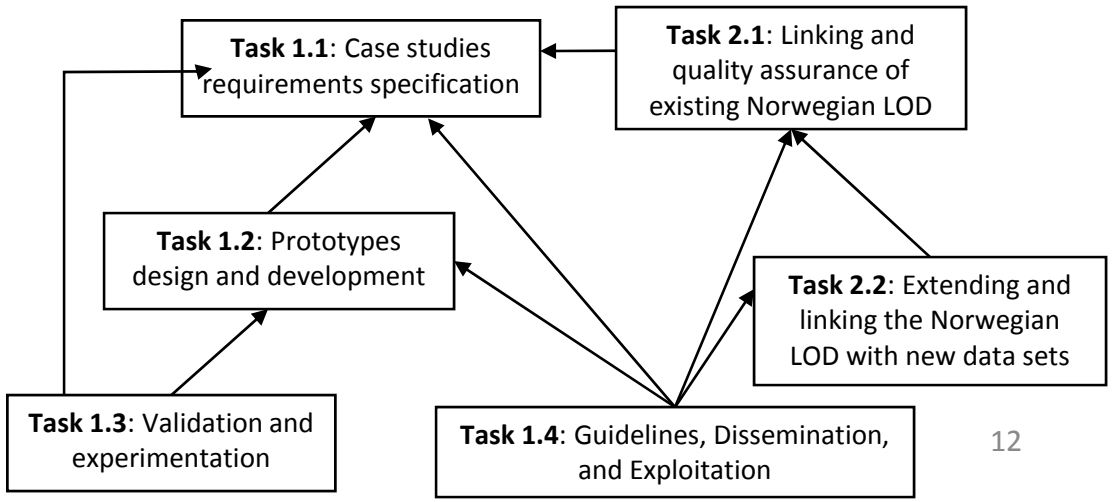
PlanetData-NorthPole

Consuming and Improving Norwegian Linked Open Data for Regional Development and Environmental Friendly Behavior

The objectives of PlanetData-NorthPole are:

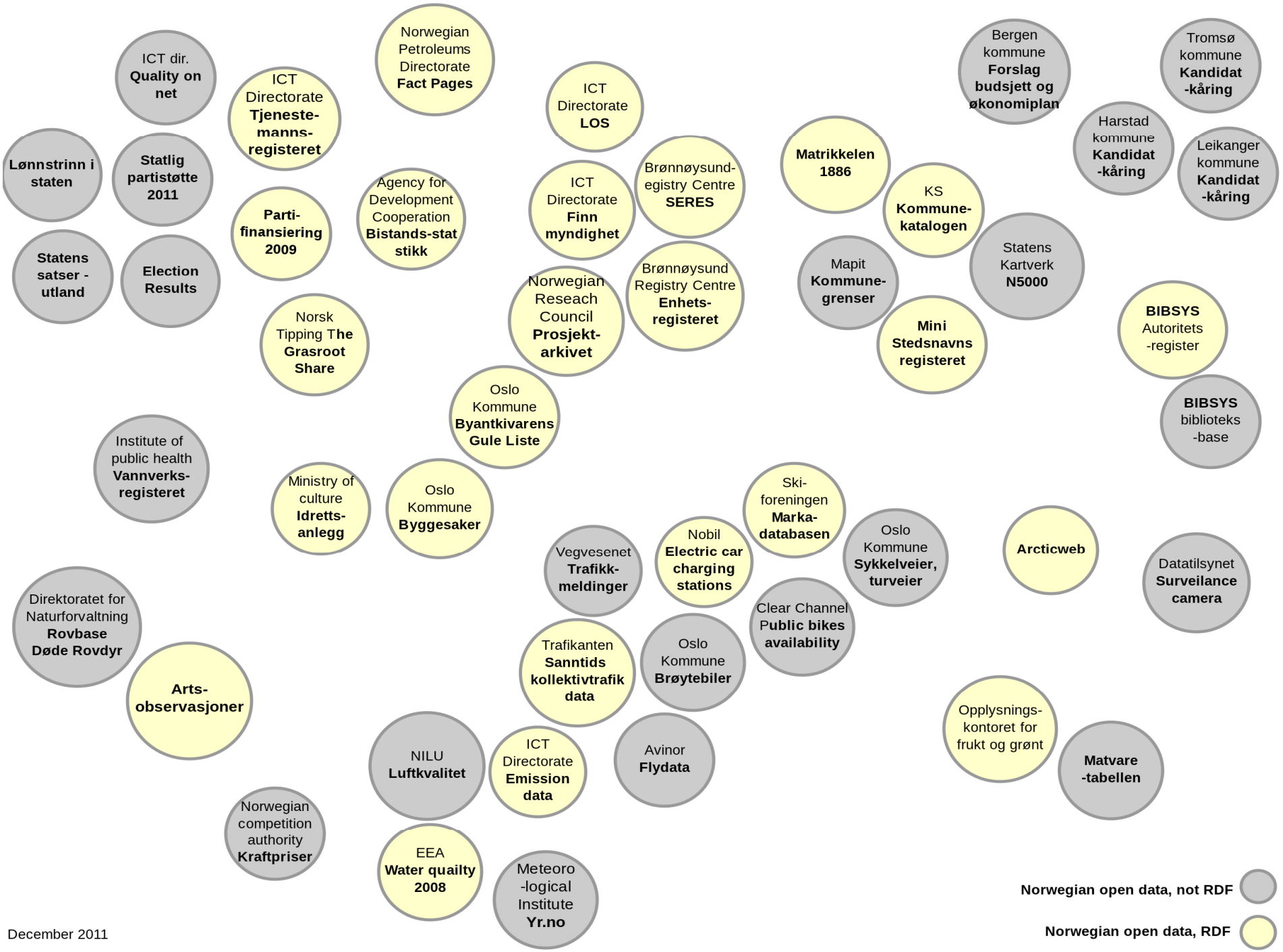
1. To **specify and implement two case studies** for demonstrating the use and benefits of LOD in *regional development* and *environmental friendly behaviour*, with a particular localization on Norway;
2. To **improve the existing Norwegian LOD and extend it** with new data sets to support the proposed case studies;
3. To **provide guidelines** for other countries in the use of LOD for regional development and environmental friendly behaviour applications.

- Participants:
  - Computas AS
  - SINTEF
- Duration: 6 months
- Budget: approx 120K Euro



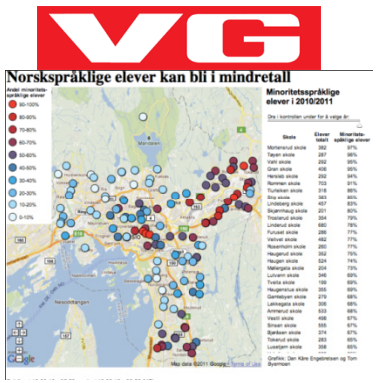
The context (2):
Semicolon II – Interoperability in Public Sector

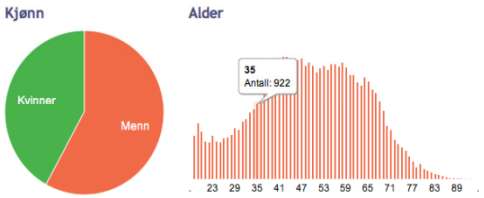
- Aim: test and establish methods, tools and performance indicators to be used as the basis for recommendations and standards for enhancing collaboration across the public sector in Norway
- A multidisciplinary project
  - Policy development, Semantic interoperability, Interoperability platforms, Legislation, Cost/benefit analyses, Organisational interoperability
- http://semicolon.no

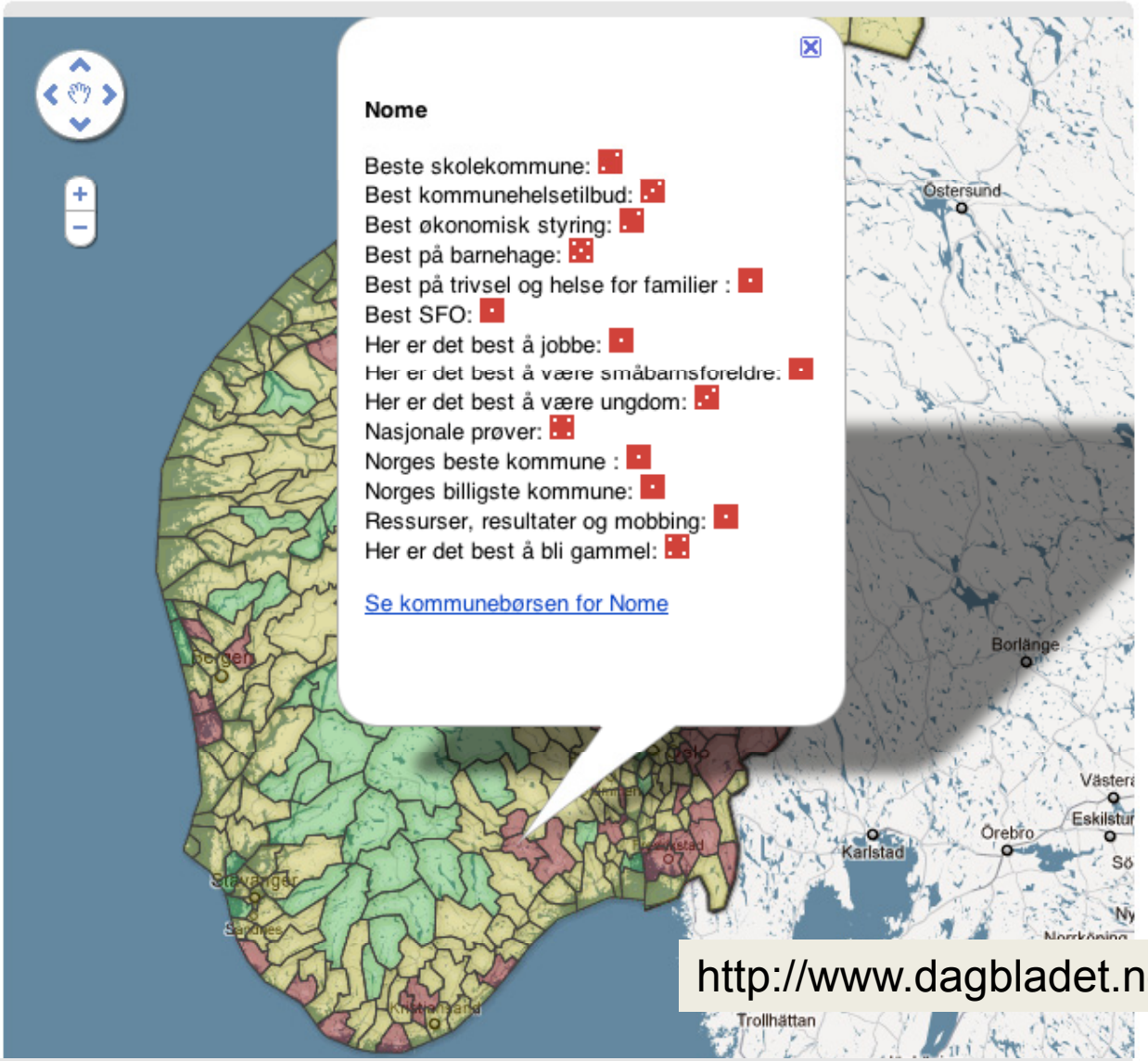

ICT dir.
Quality on net
ICT Directorate
Tjeneste-manns-registeret
Norwegian Petroleums Directorate
Fact Pages
ICT Directorate
LOS
Bergen kommune
Forslag budsjett og økonomiplan
Tromsø kommune
Kandidat-kåring
Harstad kommune
Kandidat-kåring
Leikanger kommune
Kandidat-kåring
Lønnstrinn i staten
Statlig partistøtte 2011
Parti-finansiering 2009
Agency for Development Cooperation
Bistands-statistikk
ICT Directorate
Finn myndighet
Brønnøysund-egistry Centre
SERES
Matrikkelen 1886
KS
Kommune-katalogen
Statens satser - utland
Election Results
Mapit
Kommune-grenser
Statens Kartverk
N5000
Norsk Tipping The Grasroot Share
Norwegian Reseach Council
Prosjekt-arkivet
Brønnøysund Registry Centre
Enhets-registeret
Mini Stedsnavns registeret
BIBSYS
Autoritets-register
BIBSYS
biblioteks-base
Oslo Kommune
Byantkivarens Gule Liste
Institute of public health
Vannverks-registeret
Ministry of culture
Idretts-anlegg
Oslo Kommune
Byggesaker
Ski-foreningen
Marka-databasen
Oslo Kommune
Sykkelveier, turveier
Arcticweb
Datatilsynet
Surveilance camera
Vegvesenet
Trafikk-meldinger
Nobil
Electric car charging stations
Clear Channel
Public bikes availability
Direktoratet for Naturforvaltning
Rovbase
Døde Rovdyr
Trafikanten
Sanntids kollektivtrafik data
Oslo Kommune
Brøytebiler
Arts-observasjoner
Opplysnings-kontoret for frukt og grønt
Matvare-tabellen
NILU
Luftkvalitet
ICT Directorate
Emission data
Avinor
Flydata
Norwegian competition authority
Kraftpriser
EEA
Water quailty 2008
Meteoro-logical Institute
Yr.no
December 2011
Norwegian open data, not RDF
Norwegian open data, RDF

Regional development (Use Case 1)

- Data journalism for monitoring regional developments in municipalities in Norway
  - Journalists spend significant time collecting and aggregating data for analyzing trends and monitoring regional developments
  - E.g.
    - How many job openings are there in the municipalities?
    - What's the distribution of gender and age in the political landscape?
    - Where are accidents happening?
- **Problem**: How can we speed up and improve the process of collecting and aggregating data for monitoring regional developments?
- **Added value proposition**: Enable smarter/faster monitoring of regional developments







http://www.dagbladet.no/kommuneborsen/

Klikk på kartet for å sjekke din kommune.

Kategoriene: Alle NORGES BESTE KOMMUNE Best økonomisk styring
Beste skolekommune Nasjonale prøver Ressurser, resultater og mobbing
Best kommunehelsetilbud Best å jobbe Best på barnehage Best SFO
Best å være småbarnsforeldre Best å være ungdom Norges billigste kommune

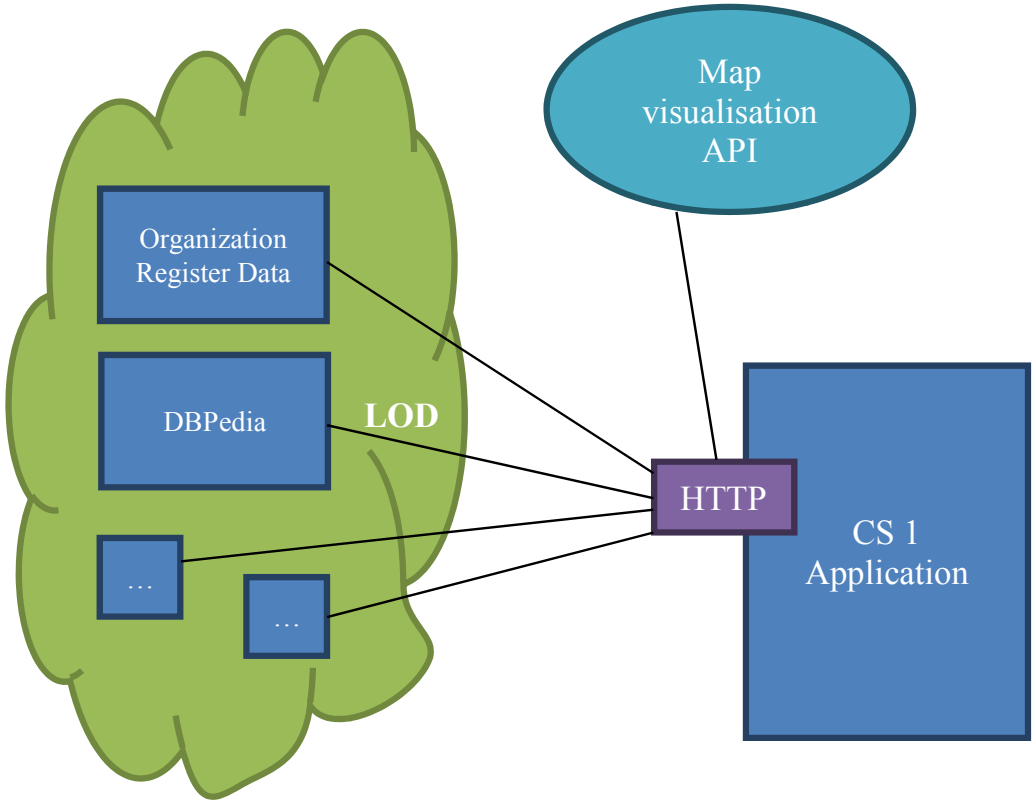
Use Case 1 – Requirements

1. The system shall register data set selections
2. The system shall register data variable selections
3. The system shall register selections of parts of the visualized data
4. The system shall be connected to a chart visualization system
5. Various Linked Open Data shall be available to the system either through RESTful Web service request(s) or SPARQL-endpoints
6. The system should be able to add more LOD sets to visualize data from

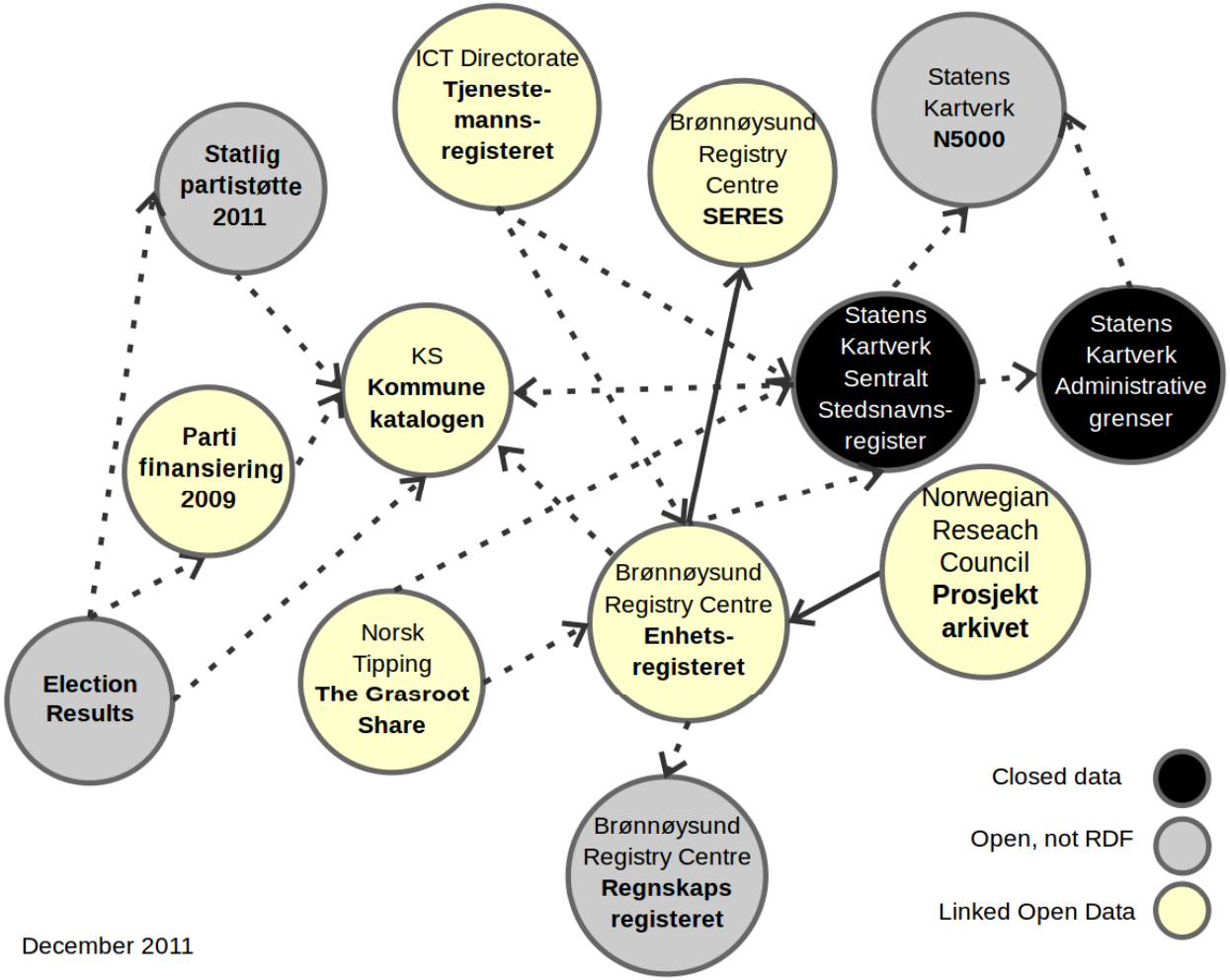
Use Case 1 – Requirements (cont’)

7. The system shall present the user with visualizations of the selected data sets
8. The system shall present the user with datasets and variables to choose to base the visualization on
9. The system shall visualize data connected to a geographical location on a geo chart
10. The application should be available through Web browsers without plugins
11. The application shall be user friendly

Use Case 1 – Application architecture outline



Use Case 1 – (Linked) (Open) Data Sources



Use Case 1 – LOD example

http://opendata.computas.no/lod/id/enhet/986429360

KARDE AS, en Enhet

Organisasjonsnr: 986429360

Organisasjonsform: Aksjeselskap (AS)

Nacekode:
70.220 - Bedriftsrådgivning og annen administrativ rådgivning

Antall ansatte: 8 (per 15.01.2011)

Adresse:
Skafjellveien 20
3070 SANDE I VESTFOLD
Norge

Kontaktperson:
Terje Johan Grimstad

- andre roller:
  - Daglig leder/ adm.direktør i KARDE AS
  - Styremedlem i KARDE AS
  - Kontaktperson i TGR INVEST AS
  - Styrets leder i TGR INVEST AS
  - Styremedlem i TELLU AS
  - Styremedlem i PALPACOM AS

Styremedlemmer:

- Styrets leder
  - Arthur Berg Reinertsen
    - andre roller:
      - Daglig leder/ adm.direktør i KARDE AS
      - Styrets leder i KARDE AS
      - Kontaktperson i ABR INVEST AS
      - Styrets leder i ABR INVEST AS
      - Styrets leder i TELLU AS
      - Styrets leder i PALPACOM AS

```
<http://opendata.computas.no/lod/id/enhet/986429360>
      a       <http://opendata.computas.no/dict/guid/Brønnøysundregistrene/Begrep/Enhet/4974> ;
      vocab:adresse
              [ a
<http://opendata.computas.no/dict/guid/Brønnøysundregistrene/Begrep/Adresse/4967> , vcard:Address
;
                vocab:gateadresse "Skafjellveien 20" ;
                vocab:kommunenavn "SANDE" ;
                vocab:kommunenr "0713" ;
                vocab:land
                        [ a       geoinfo:self_governing ,
<http://opendata.computas.no/dict/guid/Brønnøysundregistrene/Begrep/Land/4908> ;
                          rdfs:seeAlso <http://www.geonames.org/countries/NOR/> ;
                          vocab:landNavn "Norge" ;
                          vocab:landkode "NOR" ;
                          geoinfo:codeISO3 "NOR"
                        ] ;
                vocab:postnr "3070" ;
                vocab:poststed "SANDE I VESTFOLD" ;
                hvor:kommunenr "0713" ;
                hvor:postnr "3070" ;
                hvor:poststed "SANDE I VESTFOLD" ;
                vcard:street-address
                        "Skafjellveien 20"
              ] ;
      vocab:ansvarskapital
              [ a       vocab:Ansvarskapital ;
                vocab:kapitalbeløp "100000.0"^^<http://www.w3.org/2001/XMLSchema#double> ;
                vocab:kapitaltype "Aksjekapital"@no ;
                vocab:kapitalvaluta "NOK"
              ] ;
      vocab:antAnsattePåDato
              [ a       vocab:AntAnsattePåDato ;
                vocab:antAnsatte "8"^^<http://www.w3.org/2001/XMLSchema#long> ;
                vocab:gyldigFraDato "2011-01-
14T23:00:00Z"^^<http://www.w3.org/2001/XMLSchema#dateTime>
              ] ;
      vocab:enhetsnavn "KARDE AS" ;
      vocab:erEnhetIRolle _:b22 , _:b1 , _:b13 , _:b14 , _:b24 , _:b32 , _:b23 , _:b28 , _:b6 ,
_:b33 , _:b8 , _:b30 ;
      vocab:nacekode
              [ a       nace:Activity ,
<http://opendata.computas.no/dict/guid/Brønnøysundregistrene/Begrep/Næringskode/4960> ;
                rdfs:seeAlso "http://www4.ssb.no/stabas/ItemsFrames.asp?ID=5552001&Language=nb" ;
                nace:broader <http://ec.europa.eu/eurostat/ramon/rdfdata/nace_r2/70.22> ;
                nace:code "70.220" ;
                nace:name "Bedriftsrådgivning og annen administrativ rådgivning"@no ;
                vocab:nacekodekode "70.220" ;
                vocab:nacekodetekst "Bedriftsrådgivning og annen administrativ rådgivning"@no
              ] ;
```

Select question(s) for visualization

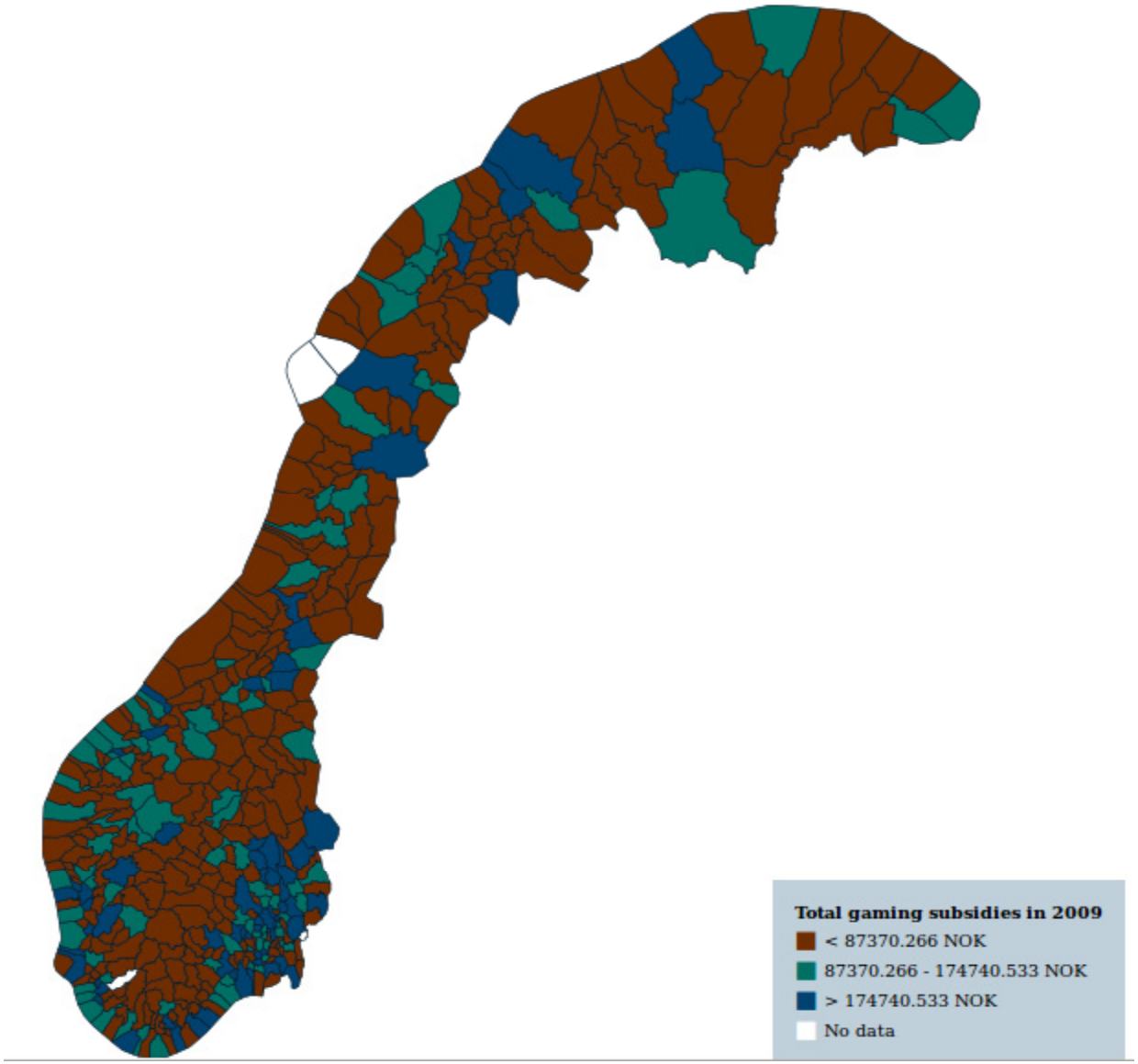
- [ ] What is the total amount of gaming subsidies given to the organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to sports organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to religious organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to art- and culture organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to nature- and environment organizations in each municipality? (2009)

Select question(s) for visualization

- [x] What is the total amount of gaming subsidies given to the organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to sports organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to religious organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to art- and culture organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to nature- and environment organizations in each municipality? (2009)

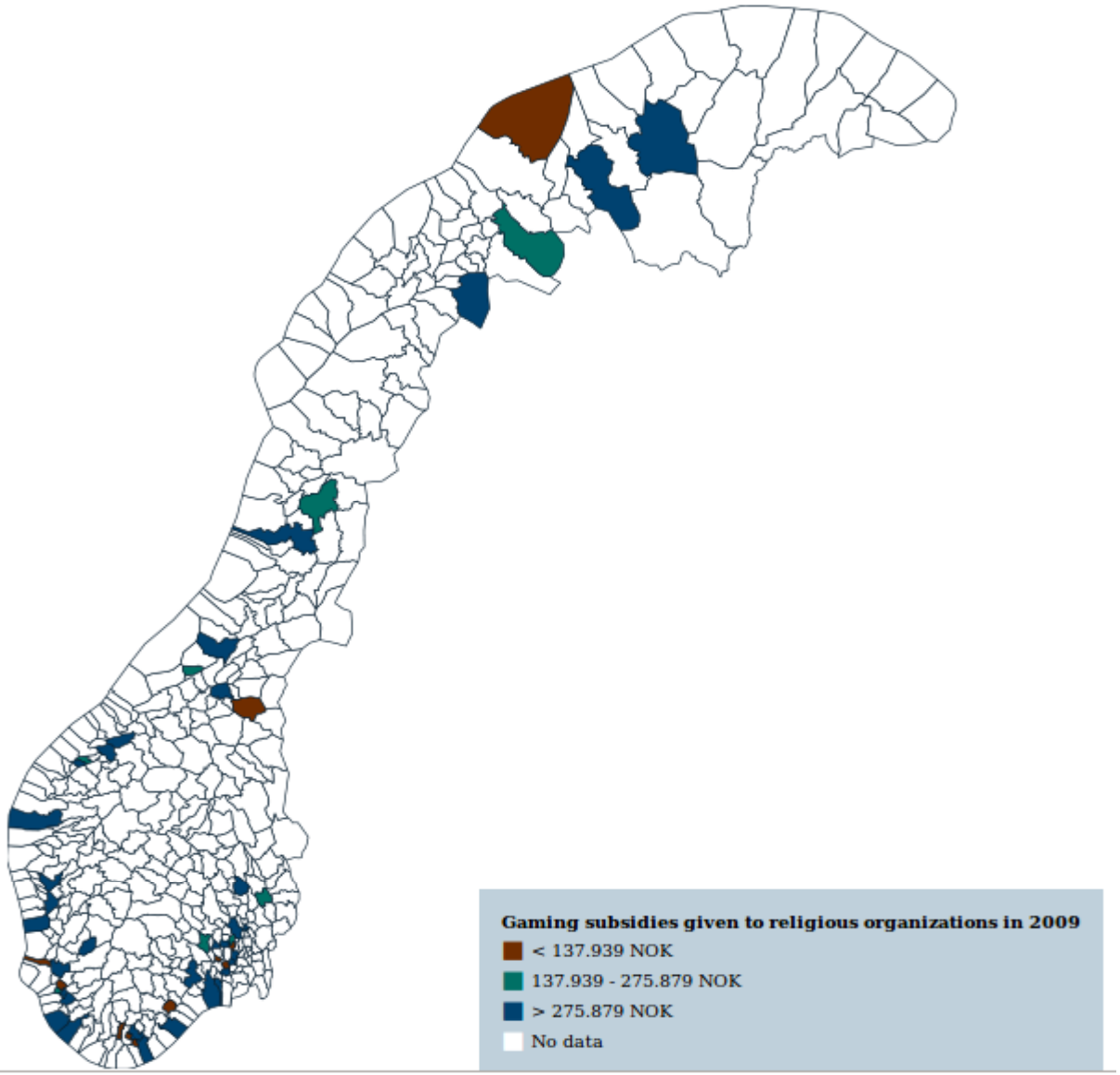
Select question(s) for visualization

- [x] What is the total amount of gaming subsidies given to the organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to sports organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to religious organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to art- and culture organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to nature- and environment organizations in each municipality? (2009)



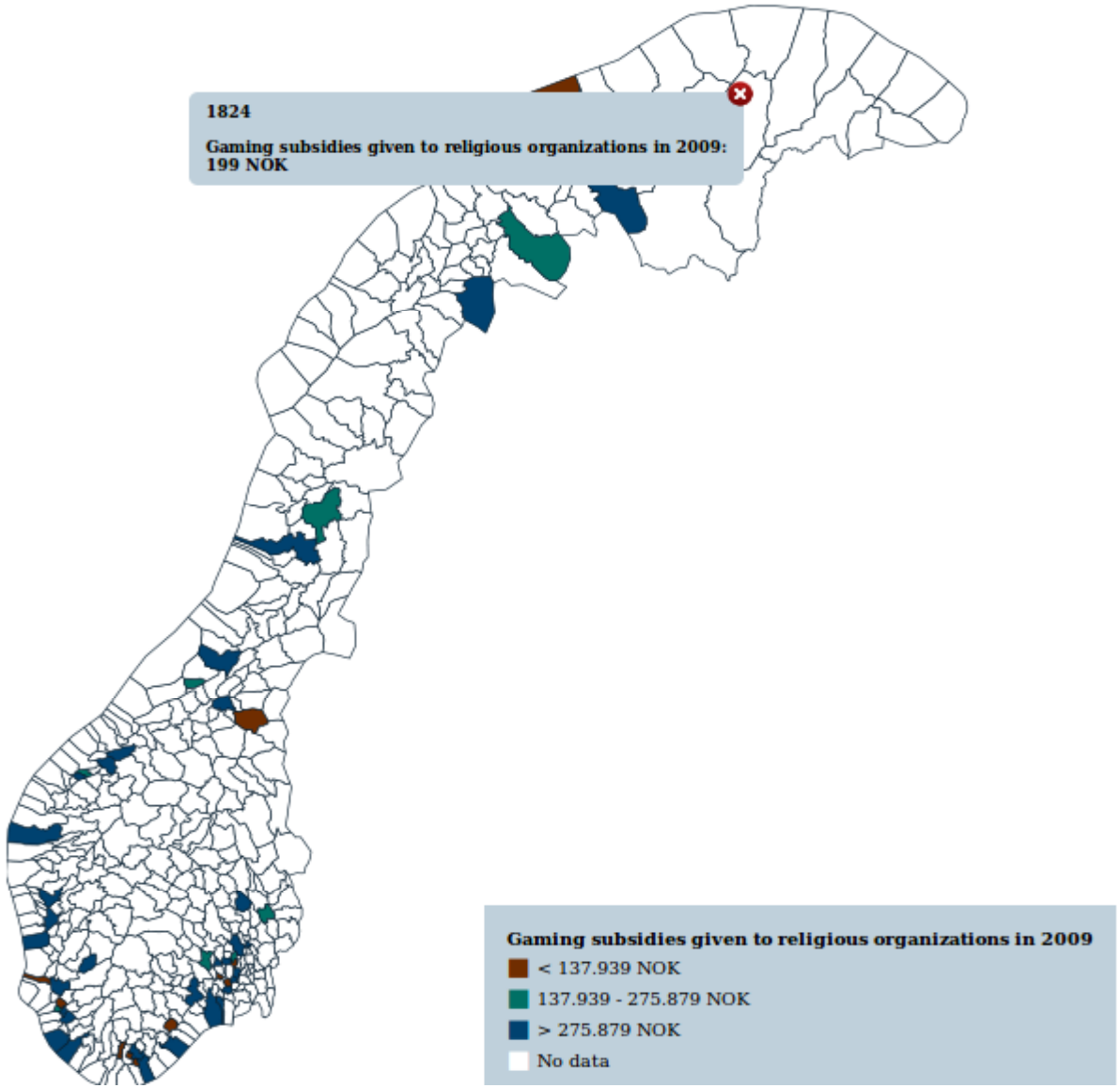
Select question(s) for visualization

- [ ] What is the total amount of gaming subsidies given to the organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to sports organizations in each municipality? (2009)
- [x] What is the amount of gaming subsidies given to religious organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to art- and culture organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to nature- and environment organizations in each municipality? (2009)



Select question(s) for visualization

- [ ] What is the total amount of gaming subsidies given to the organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to sports organizations in each municipality? (2009)
- [x] What is the amount of gaming subsidies given to religious organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to art- and culture organizations in each municipality? (2009)
- [ ] What is the amount of gaming subsidies given to nature- and environment organizations in each municipality? (2009)

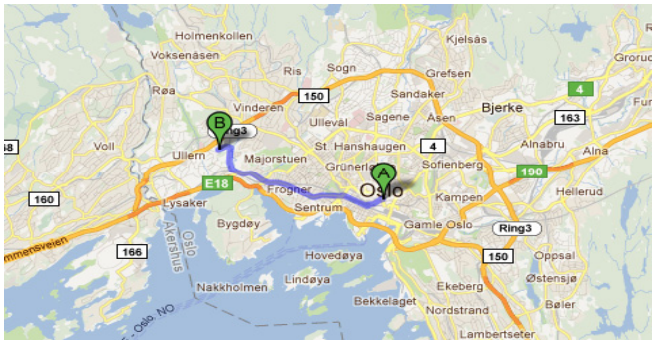


UC1 – Demo

Available @
http://opendata.computas.no/RegionalDevelopment/

Environmental friendly behaviour (Use Case 2)

- **Problem statement**: Faced with different transportation options for a short trip, which are the most environmental-friendly options given constraints like time, weather, traffic and private preferences.



- Typically different options
  - Public transportation (bus/tram/metro/train)
  - Private car (electric/gas/diesel) car, taxi
  - Cycling, walking
- Constraints: Time, avoid bad weather, polluted zones, traffic, private preferences
- Environmental parameters: CO2 emissions, energy efficiency
- **Added value proposition**: Enable smarter/faster environmental friendly decision making for local trips when options are available
  - Assist the user's decision making wrt travelling from his current position to the position of the next event in the user's calendar

Use Case 2 – Requirements

1. The user shall not have to enter any information when starting the app
2. Various Linked Open Data shall be available to the system either through RESTful Web service request(s) or SPARQL-endpoints
3. Private user data – calendar events and the geographic position of the user – shall be made available to the app
4. Communication, forming requests - The system shall generate requests to the required LOD systems
5. An initial request will be formed with arguments fetched from the closed (user) data from req. #3, e.g.: “Find the 15 closest public transportation stops to the user’s position”

Use Case 2 – Requirements (cont')

6. Communication, handling replies - The system shall parse the data it receives and add emission information to a transport alternative by combining emission data with the transport data

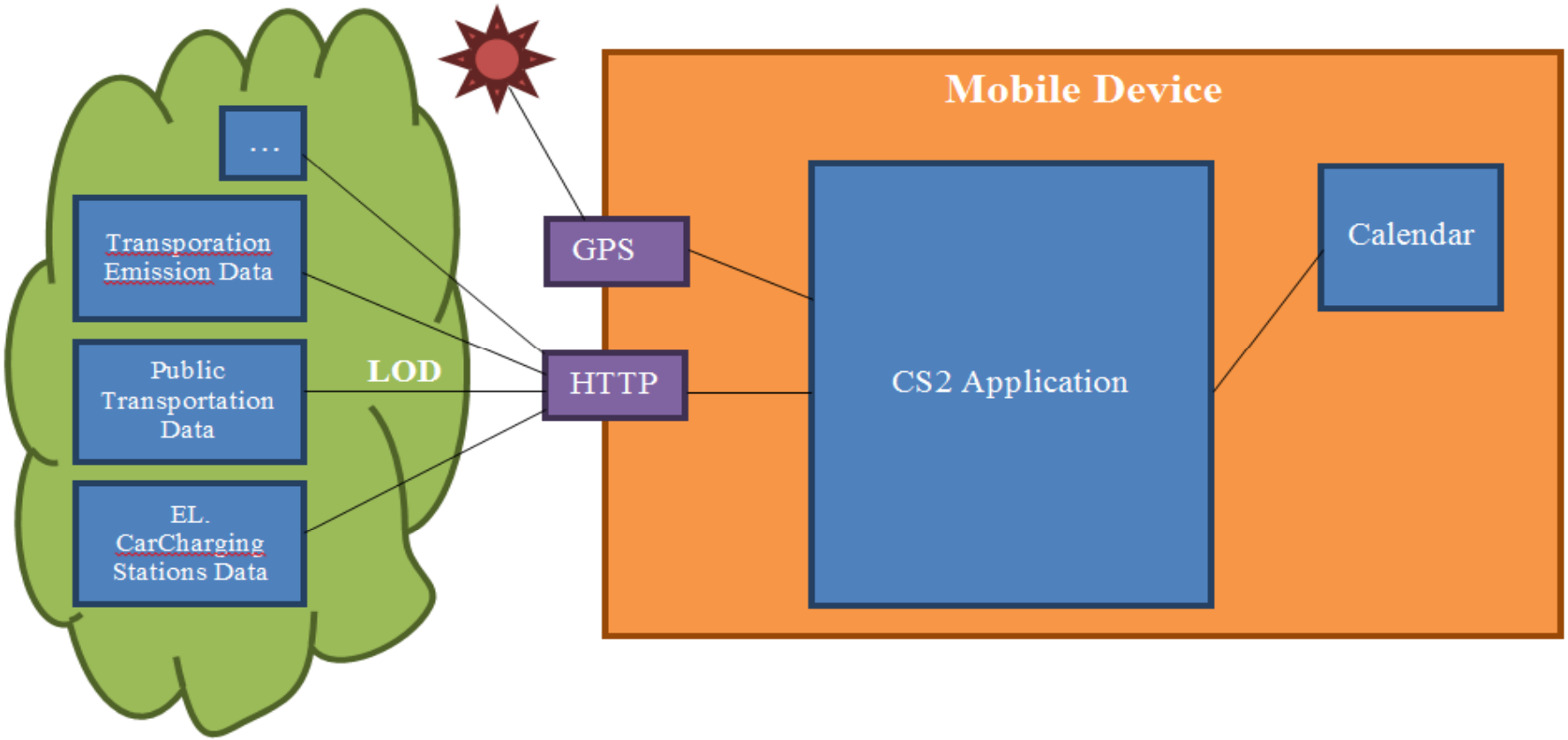
7. The system shall add entries (events) to the user's calendar corresponding to the transportation alternatives found. The added calendar event(s), i.e. travel alternatives shall at least contain the following information: Start time, end time, transport type, route description and environmental friendliness (e.g. CO2 emission)

8. Application result quality - The added calendar entries should to be of quality wrt taking the user from his current position to the actual address of the calendar event the entries where based on.

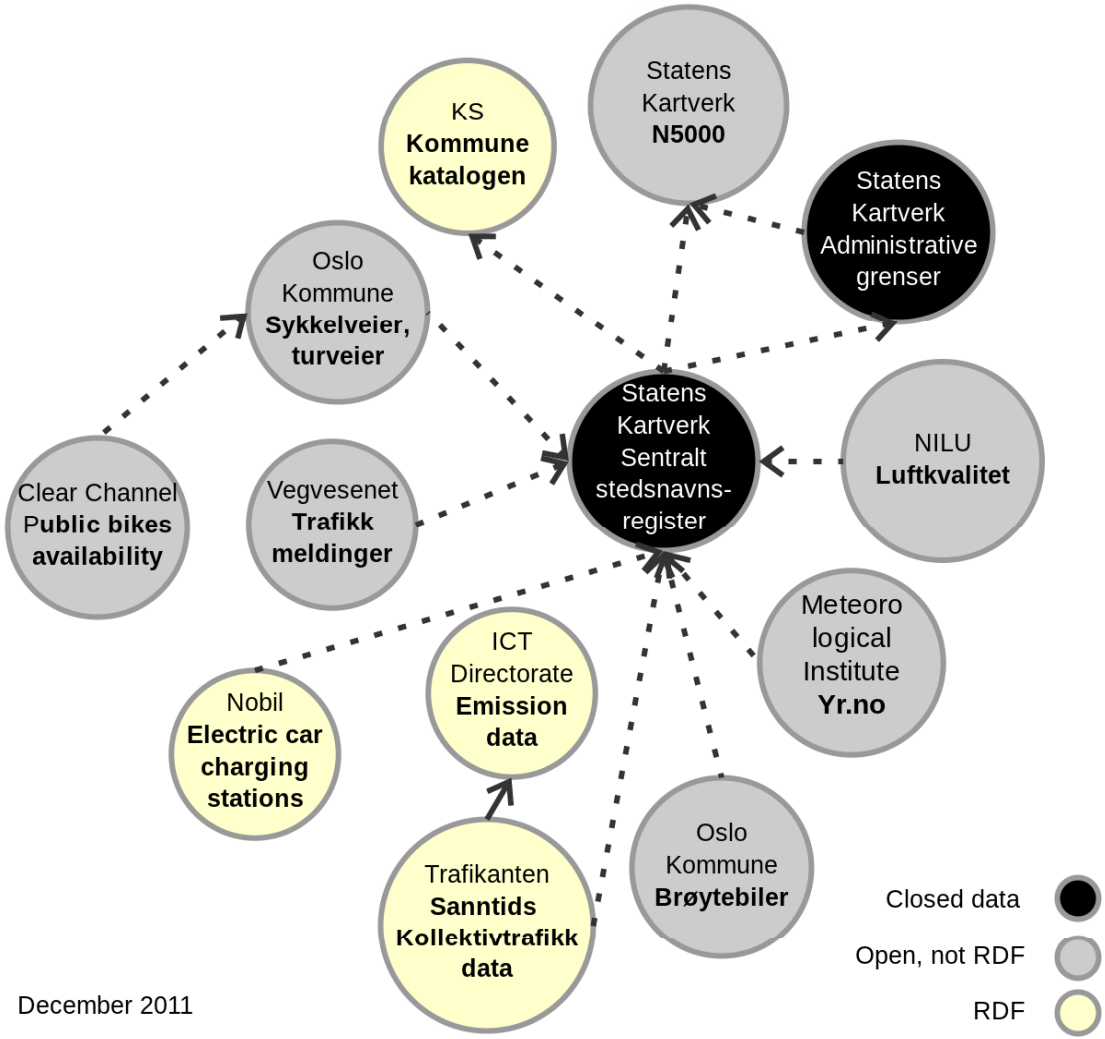
9. The app could display a notification telling whether transportation alternatives where found or not

10. The data lifting and application shall try to utilise common linked data vocabularies for transportation and environment

Use Case 2 –
Application structure outline



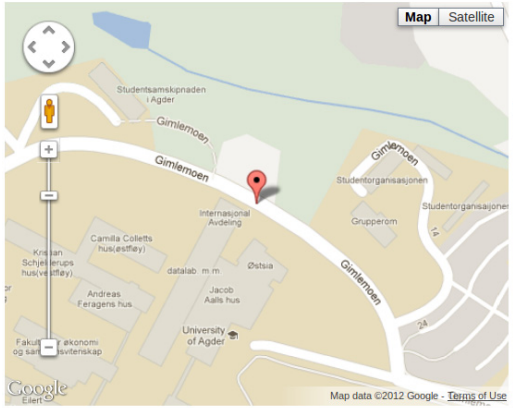
Use Case 2 – (Linked) (Open) Data Sources

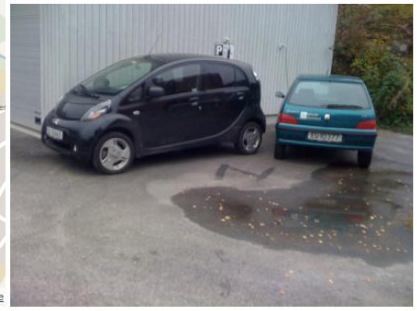


Use Case 2 – LOD example

http://opendata.computas.no/nobil/id/chargingStation/1121

Universitetet i Agder, Kristiansand

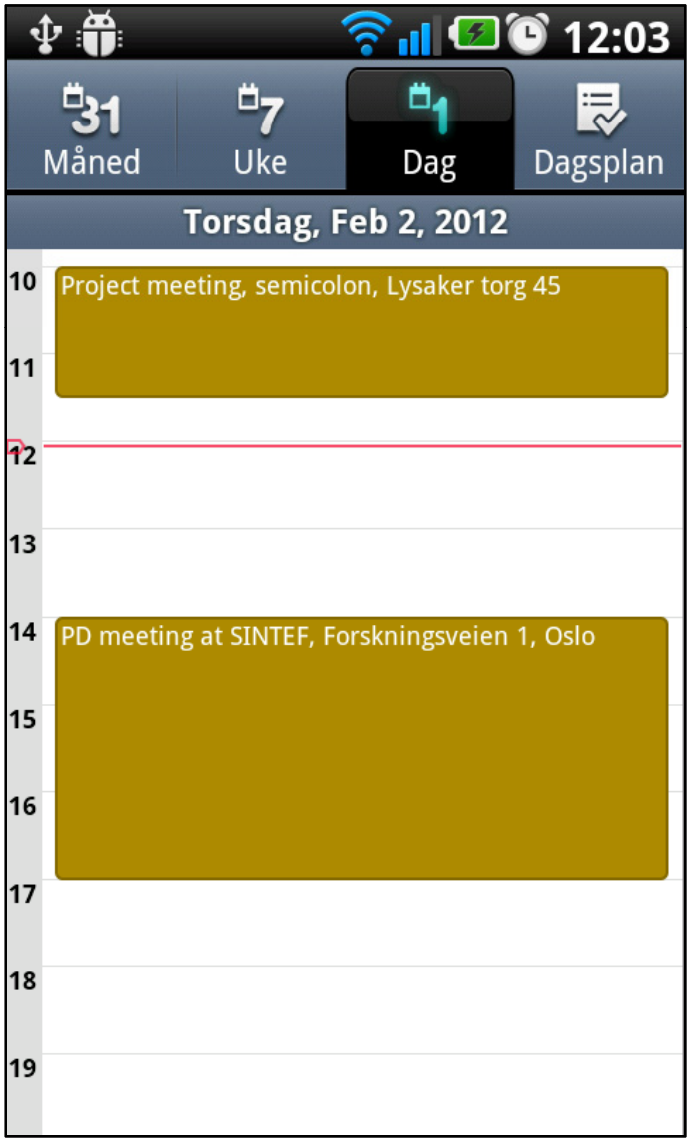




Picture:	http://www.nobil.no/img/ladestasjonbilder/1121.jpg
Thumbnail:	http://www.nobil.no/img/ladestasjonbilder/tn_1121.jpg
Owner:	Universitetet i Agder
Available slots:	2
Place type:	Gateplan
Charging station ID:	1121
Availability:	Besøkende
Contact info:	drift-krs@uia.no
Place description:	Til høyre for hovedbygning. Ingen ordentlige parkeringsplasser, men et skilt og to stikkontakter i en vegg.
Time limit (hours):	0
Street address:	Gimlemoen 25
Postal code:	4604
Municipal number:	1001
Postal area:	KRISTIANSAND S
Latitude:	58.1644
Longitude:	8.0038
Geometry:	POINT(58.1644 8.0038)
Point:	8.0038 58.1644
Charging speed:	16A
Access key:	Åpen
Charging outlets:	2

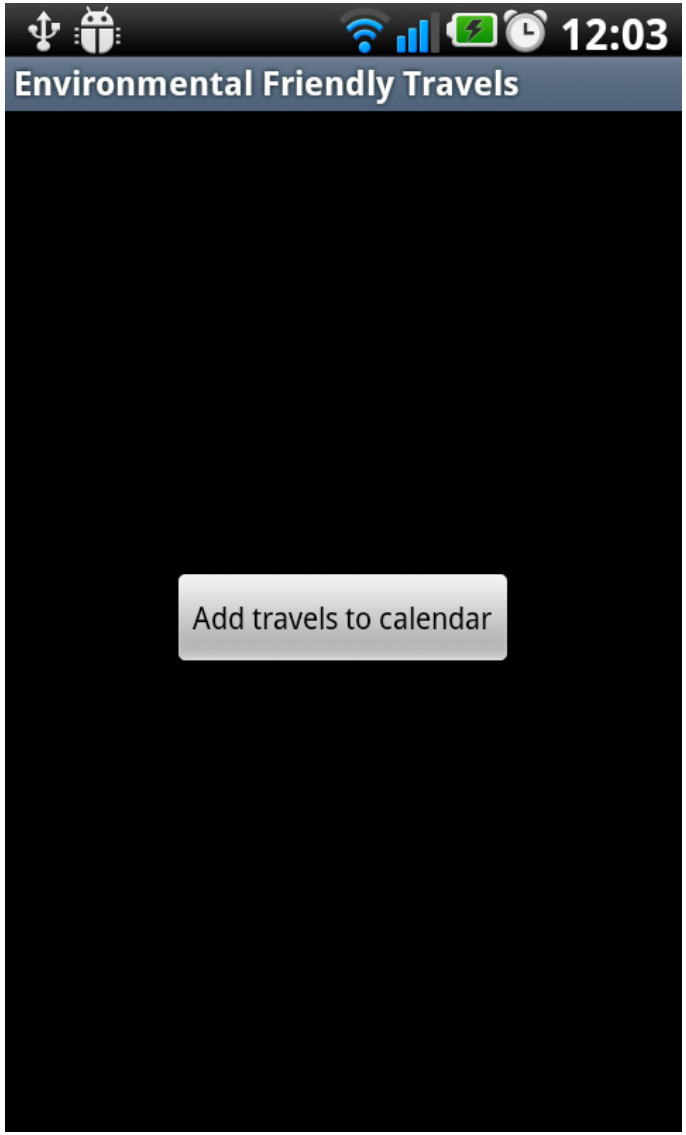
UC2 – Demo Screenshots

UC2 – Demo Screenshots

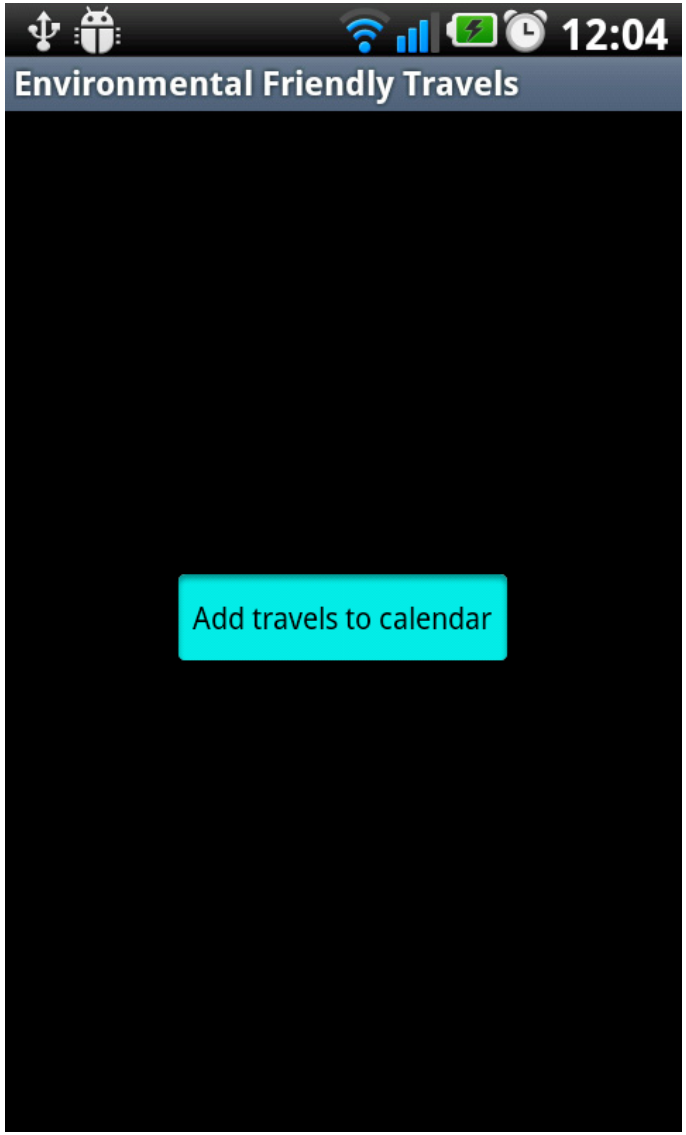


UC2 – Demo Screenshots

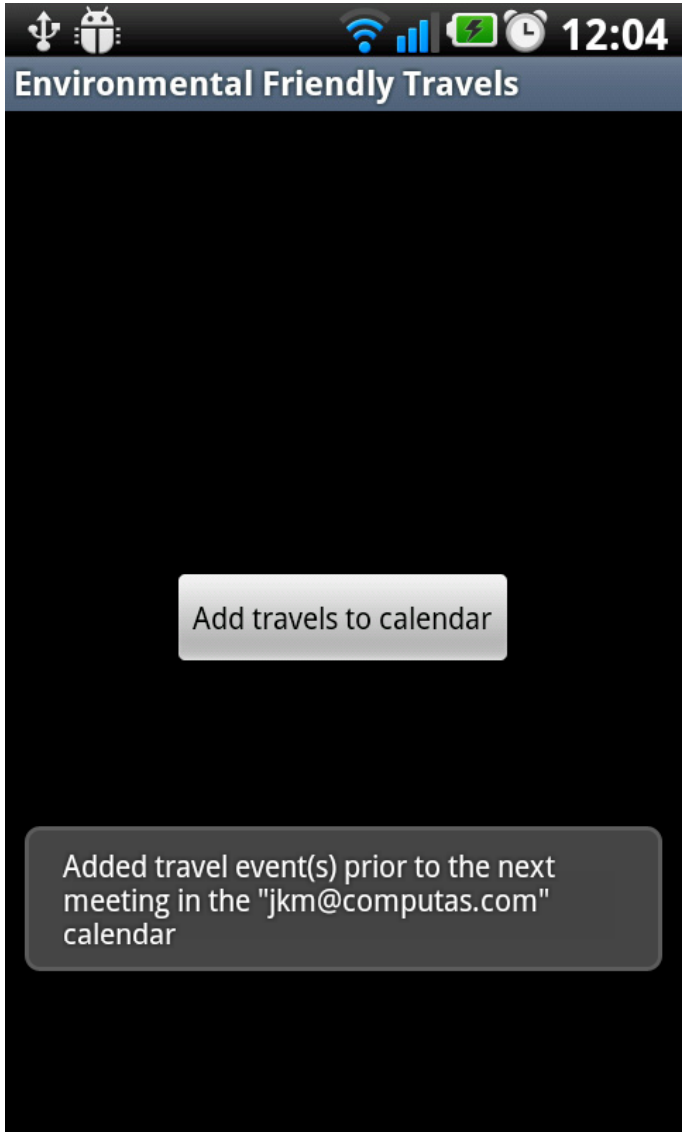
UC2 – Demo Screenshots



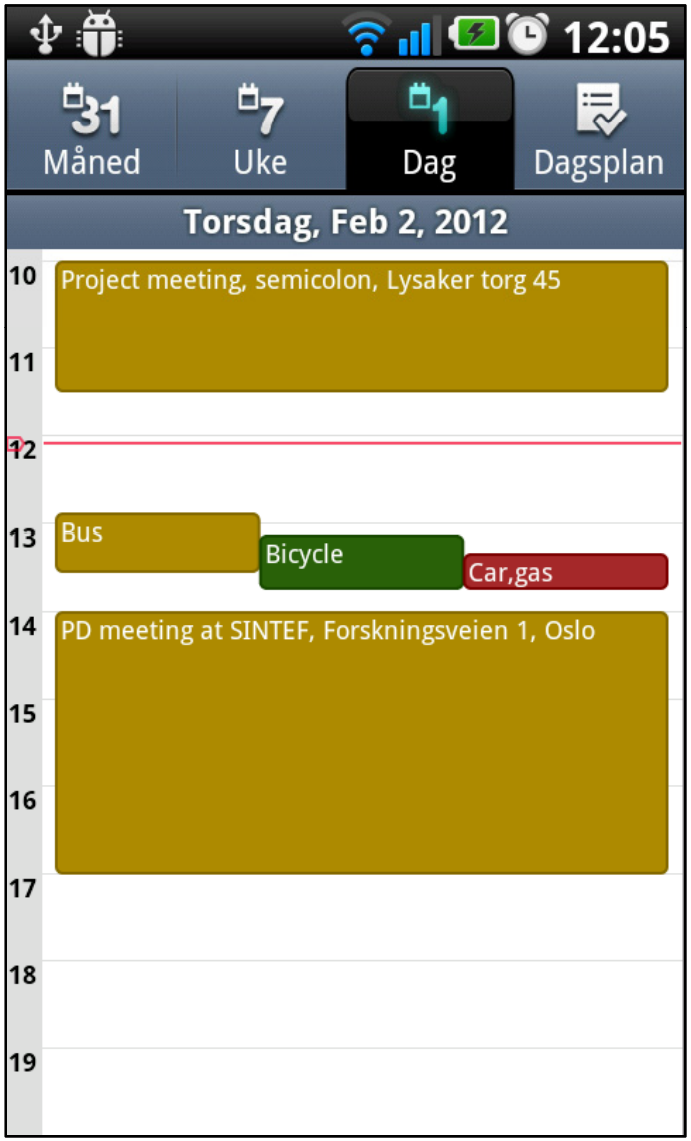
UC2 – Demo Screenshots



UC2 – Demo Screenshots

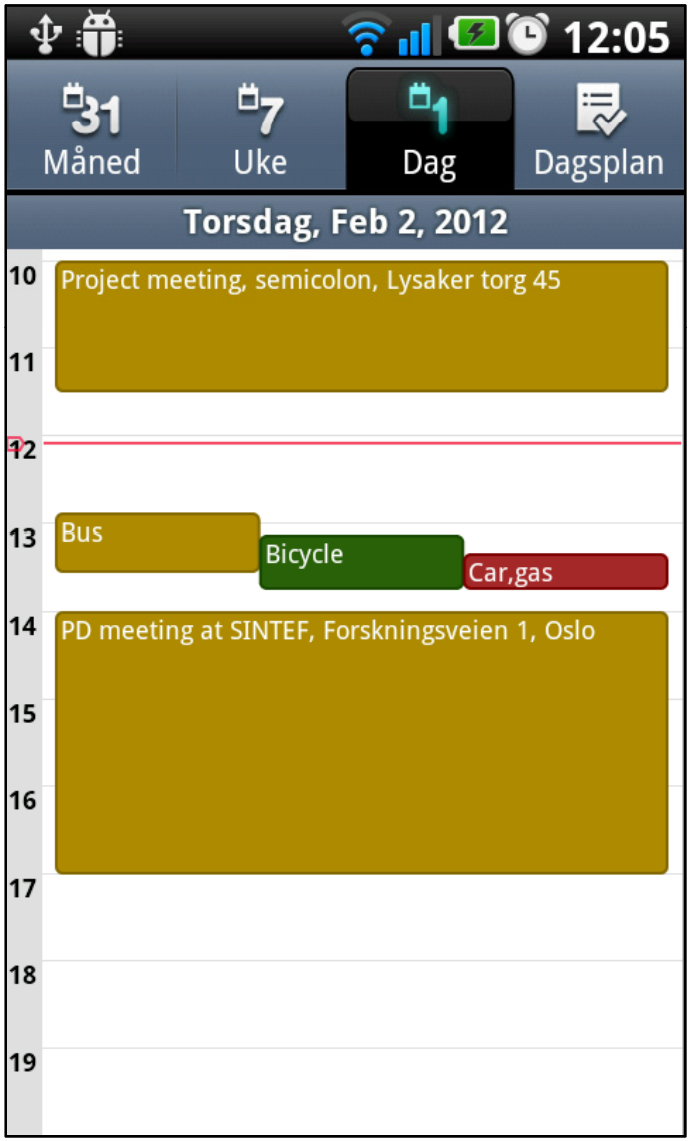


UC2 – Demo Screenshots

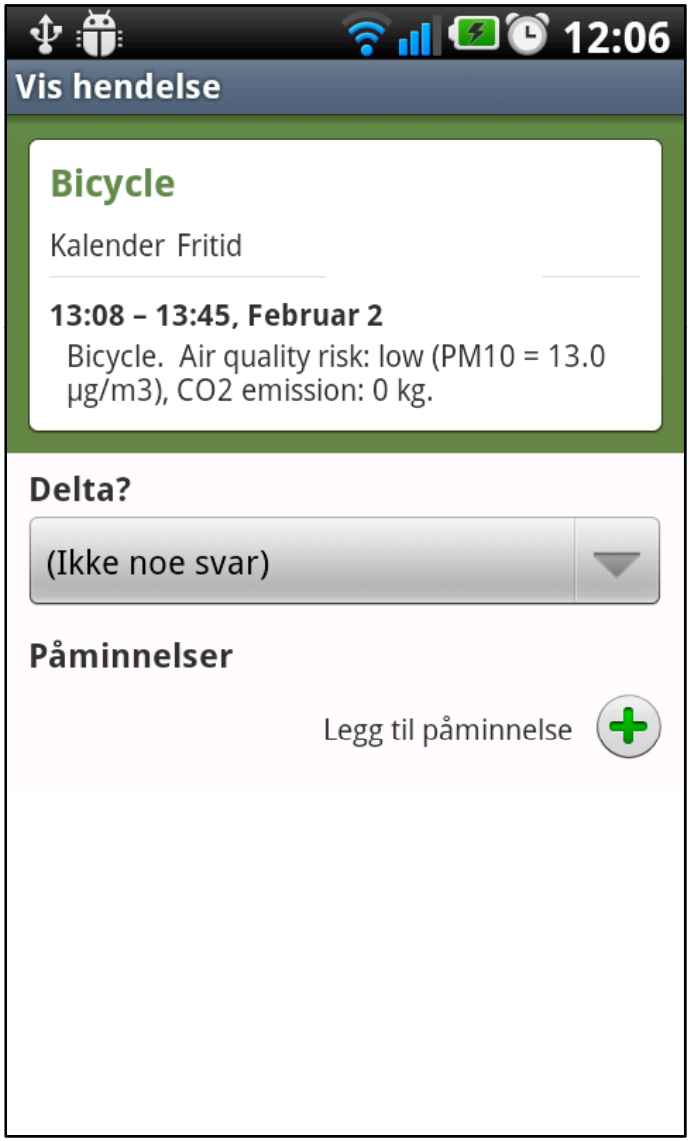


UC2 – Demo Screenshots

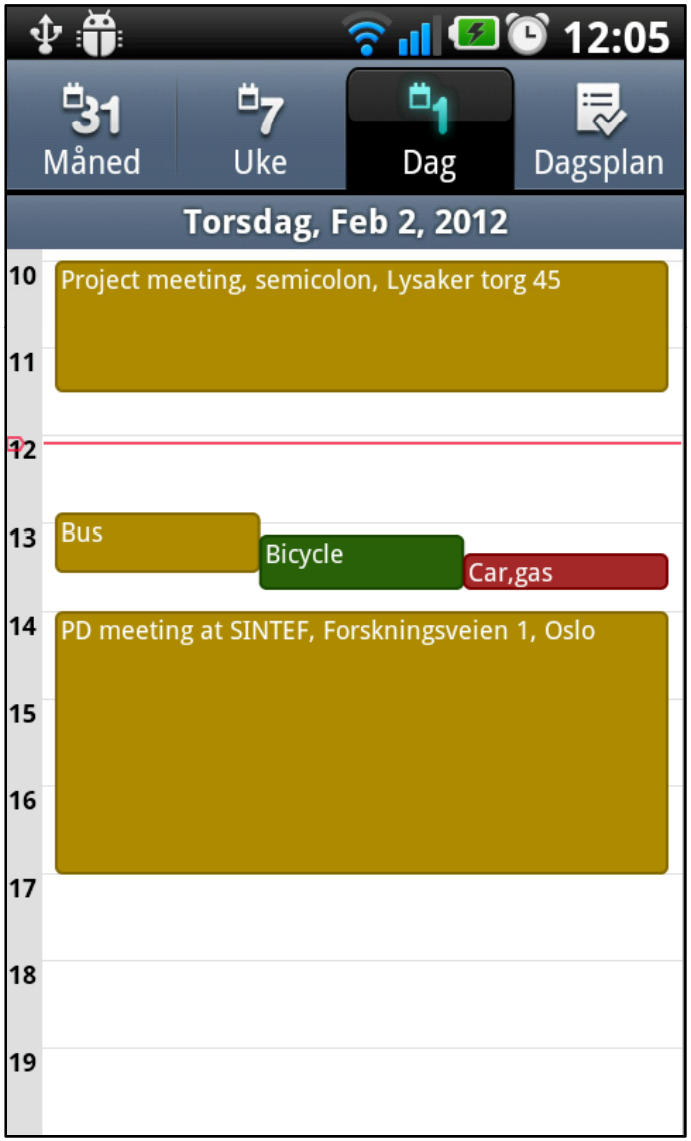
UC2 – Demo Screenshots



UC2 – Demo Screenshots

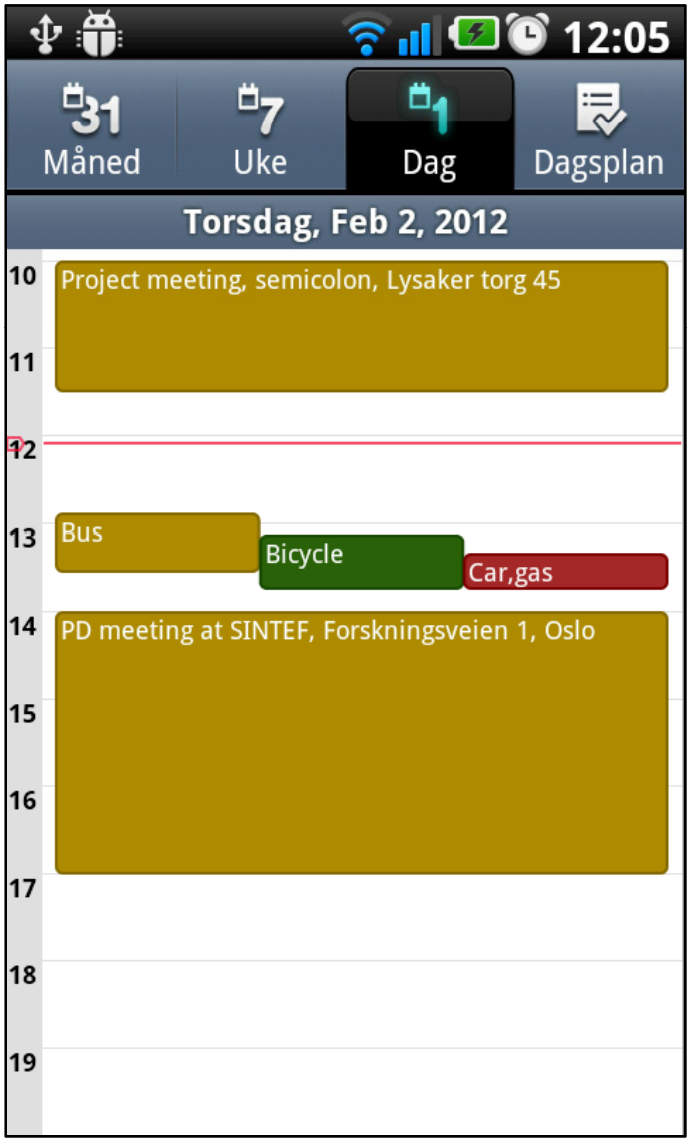


UC2 – Demo Screenshots



UC2 – Demo Screenshots

UC2 – Demo Screenshots



Outline

- Open Data Movement
  - Linked Open Data
- Norwegian Linked Open Data – Use Cases
  - Monitoring Regional Development
  - Environmentally Friendly Behavior
- Summary

UC1 – Further improvements

- Plug-in more data sets to the application
  - More questions could be answered
- Improve the graphical experience
- Validation with data journalists (e.g. through workshop/questionnaires)
  - MediArena

    http://www.mediarena.no/

    "*The MediArena Business Cluster is a unique collaboration project between global technology corporations, national broadcasters and small, forward-leaning meditech entrepeneurs - all situated in Bergen, Norway.*"
  - Dagbladet Kommunebørsen

    http://www.dagbladet.no/kommuneborsen/

UC2 – Further improvements

- Support for user preferences, so that the user can configure
  - Bike/car availability
  - Car type
  - Which calendar(s) to generate events from
- Integrate the app with data from the Veimeldinger data set
- Validation
  - Distribute application for use to Computas/SINTEF employees and collect feedback (e.g. through questionnaires)

Summary

- Open Data Movement
  - More transparency, participation and collaboration
  - Free access to information and the possibility to freely use and re-use this information
- Linked Open Data
  - Interoperability, standards, tools for Open Data
- Two use cases for the Norwegian Linked Open Data
  - Monitoring Regional Development
  - Environmentally Friendly Behavior

Thank you!

Questions?

Title	*Author*	*URL to data set homepage*	*Triples*	**Star scheme score (Lee, 2006)**	**Case study applicability**	**Comments**
Enhetregisteret	Brønnnøysund Registry Centre	Computas http://opendata.computas.no/about/#BR-Enhetsregisteret	> 4.500.000	5	Case study #1	RESTful RDF WS Wrapper
Kommunekatalogen	KS	Univ. of Oslo http://sws.ifi.uio.no/sparqler/	Ca. 2.000	3	Case study #1 Case study #2	SPARQL
NFR prosjektarkiv	Norwegian Research Council	Computas http://opendata.computas.no/about/#NFR-Prosjektkatalog	Ca. 200.000	5	Case study #1	SPARQL
Partifinansiering2009	The Ministry of Government Administration, Reform and Church Affairs	Univ. of Oslo http://heim.ifi.uio.no/audus/rdf/partifinansiering.rdf	Ca. 100.000	4	Case study #1	RDF file dump
Grasrotandelen	NorskTipping	Computas http://opendata.computas.no/joseki/	Ca. 70.000	4	Case study #1	SPARQL
Trafikkantensanntid	Oslo Kommune	Computas http://opendata.computas.no/trafikanten/id/stop/3010030	Ca. 50.000	3	Case study #2	RESTful RDF WS Wrapper
Electric car charging stations	Nobil	Computas http://opendata.computas.no/nobil/id/chargingStation/902	Ca. 1.100	3	Case study #2	RESTful WS Wrapper (RDF returning)
Idrettsanlegg	Ministry of Culture	Computas http://opendata.computas.no/joseki/	Ca. 1.000.000	3	Case study #2	SPARQL
Emission data	Norwegian ICT Directorate	Computas http://opendata.computas.no/resource/emission/El-bil	Ca. 150	4	Case study #2	SPARQL + RESTful WS (RDF returning)
Tjenestemanns-registeret	Norwegian ICT Directorate	Computas http://opendata.computas.no/about/#Difi-Tjenestemannsregisteret	Ca. 1.000.000	5	Case study #1	SPARQL + RESTful WS (RDF returning)
SERES	Brønnøysund Registry Centre	Computas http://opendata.computas.no/about/#BR-Seres	Ca. 50.000	5	Case study #1	RESTful WS Wrapper (RDF returning)