

Helsinki Region Infoshare

Technical Description

HiQ Finland Oy
20.1.2015

1 HELSINKI REGION INFOSHARE SERVICE

The **Helsinki Region Infoshare (HRI) service** aims to make regional information quickly and easily accessible to all. Essentially, HRI is a web service for fast and easy access to open data sources between the cities of Helsinki, Espoo, Vantaa and Kauniainen. The data published is mainly statistical, giving a comprehensive and diverse outlook on different urban phenomena, such as living conditions, economics and well-being, employment and transport. A good proportion of the data material offered by the service is GIS based.

The data can be used in research and development activities, decision-making, visualization, data journalism and in the development of apps. The data may be used by citizens, businesses, universities, academies, research facilities or municipal administration. The data on offer is ready to be used freely at no cost. There are no limitations on users; anyone interested in open data can participate.

For example, existing open data has already been utilized in different ways. Take a look at the [HRI app gallery](#) (mainly in Finnish) for an overview of apps and services created with the use of open data.

There are **four operational areas** in which the HRI service mainly operates in:

1. Producing data
2. Opening data
3. Sharing data
4. Utilising data

The main operational activity is to support the producers of information in opening their data and to increase its utilization by multi-channel communication. More information about the operational areas can be found [here](#).

Our code is available from city of Helsinki Urban Facts GitHub: <https://github.com/Helsingin-kaupungin-tietokeskus> under an open license.

2 TECHNICAL DETAILS

Helsinki Region Infoshare runs on a mix of [WordPress](#) and [CKAN](#). CKAN is used for the data catalogue and WordPress for the blogs, discussion, news articles and comments. WordPress and CKAN are seamlessly integrated with each other such as WordPress provides commenting, user authentication and authorization services for the CKAN. CKAN synchronizes dataset information in real time to WordPress for the purpose of keeping track of the dataset commenting, google analytics statistics and to present datasets on the front page.

All the services run on cloud infrastructure in Ireland provided by Amazon Web Services (AWS) and are developed and maintained by [HiQ Finland Oy](#). Amazon Virtual Private Cloud (Amazon VPC) is used to logically isolate the services from other networks in the AWS cloud and to keep complete control over the Helsinki Region Infoshare virtual network. AWS cloud architecture is illustrated in Figure 1.

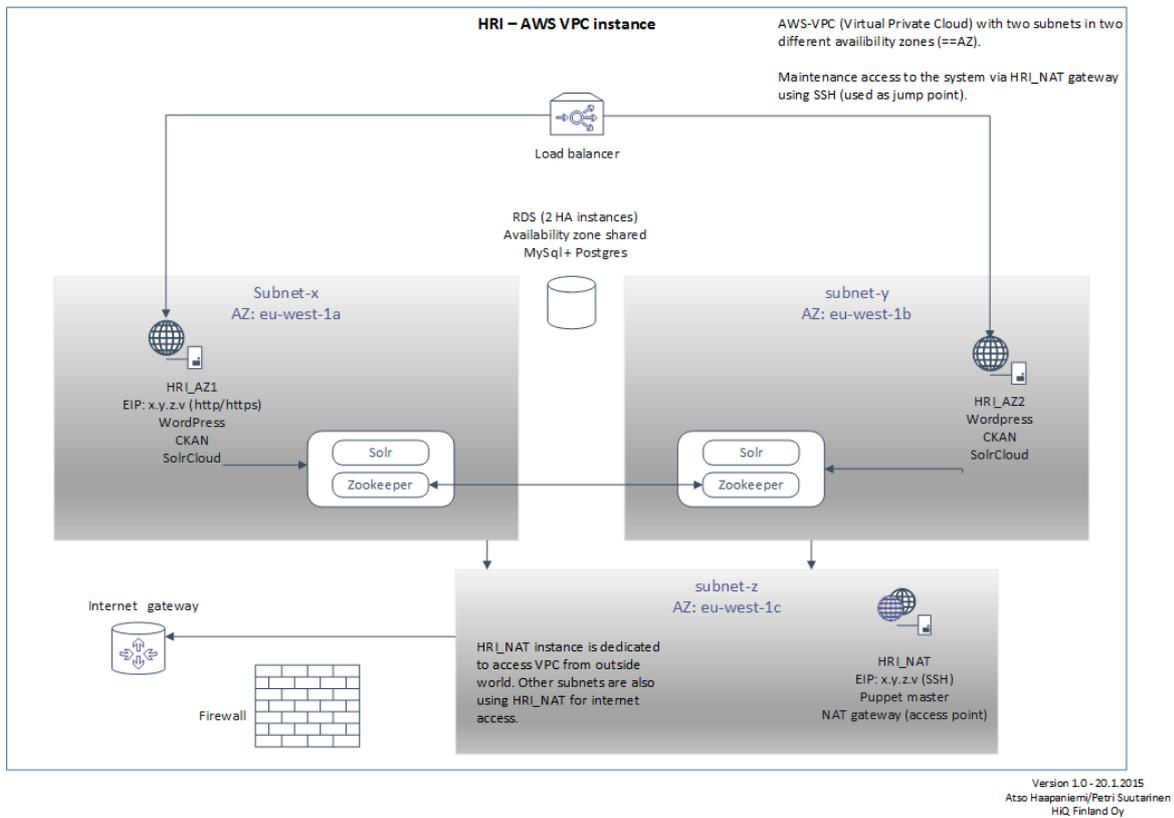


Figure 1: Helsinki Region Infoshare AWS cloud architecture.

WordPress and CKAN with other supporting services are installed on Linux operating system in EC2 instances. Ubuntu (12.04) is currently used as the Linux distribution. In production environment the EC2 instances are distributed to multiple Availability Zones for high level of fault tolerance. Elastic Load Balancing is used to automatically distribute the incoming traffic across the multiple EC2 instances.

Wordpress uses MySQL as the database and CKAN PostgreSQL database. Amazon Relational Database Service (Amazon RDS) is used for both of these databases. Databases are deployed as Multi-Availability Zone (Multi-AZ) services for high level of fault tolerance.

CKAN uses Solr for providing fast search capabilities by indexing the stored metadata. In our distributed environment ZooKeeper is used to manage distributed indexing nodes with centralized management.

2.1 ADDITIONAL SERVICES

Helsinki Region Infoshare uses AWS Simple Storage Service (S3) bucket as a data storage for some 3rd party data providers who do not have other place to store the data. This data storage is accessible with URL <http://datastore.hri.fi/>.

In addition AWS Route 53 service is used as the Domain Name System (DNS) for the Helsinki Region Infoshare.

2.2 CKAN API

Developers access to the data catalogue with an API is made possible through the CKAN RESTful and functional interfaces, all in JSON format, making it suitable for a wide range of clients.

The base URL for the Helsinki Region Infoshare API is: <http://ckan.hri.fi/api/3>.

The API is provided by the CKAN software and full instructions are provided here: <http://docs.ckan.org/en/ckan-2.1.1/api.html>.

2.3 CKAN METADATA DAILY DUMP

Helsinki Region Infoshare provides daily dump of the whole metadata database as JSON file. This file can be loaded from the services metadata description page: <http://www.hri.fi/fi/dataset/hri-tietoaineistojen-metatiedot>.