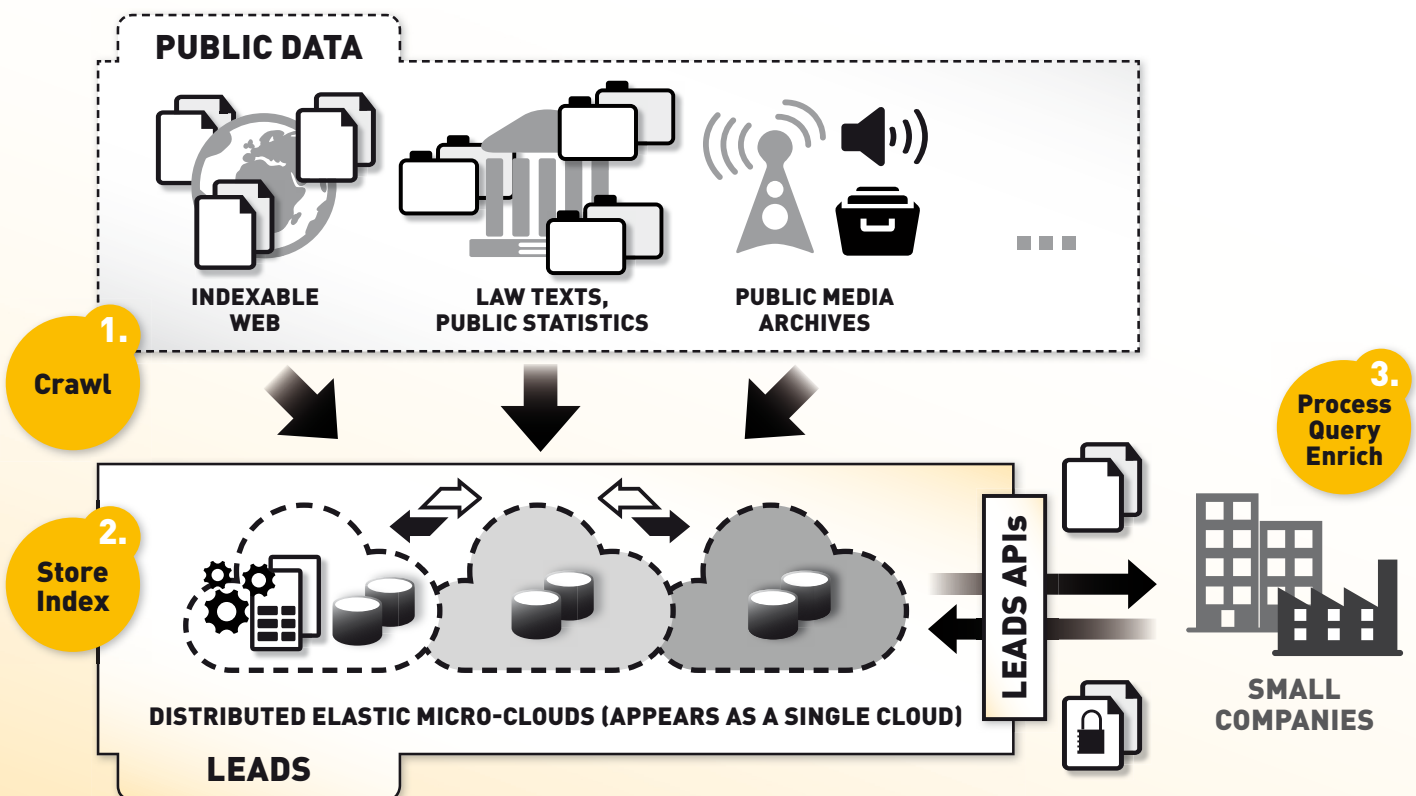# LEADS
LARGE-SCALE ELASTIC ARCHITECTURE FOR DATA AS A SERVICE

SINCE ITS CREATION, THE WEB HAS GROWN AT EXPONENTIAL RATE:

· Every day, 8 times the total content of all the U.S. libraries is added to the web;

· Every minute 35 hours of video are added to YouTube; and

· Every second 4000 tweets are sent.

Processing these vast quantities of data is a task that only the major Internet companies can afford. In particular, the cost of crawling, storing and processing public data is orders of magnitude beyond the capabilities of small and medium scale companies.

The **LEADS** (Large-scale Elastic Architecture for Data-as-a-Service) project is an academic-industrial collaboration funded by the European Commission that involves three universities and four companies. **LEADS** aims at bringing web mining and web analysis capabilities to a broad spectrum of companies, academic institutions and non-profit organizations. The project is based on a novel DaaS (Data-as-a-Service) model that leverages a unique infrastructure of local and energy-efficient micro-clouds.

PUBLIC DATA

INDEXABLE WEB

LAW TEXTS, PUBLIC STATISTICS

PUBLIC MEDIA ARCHIVES

1. Crawl

2. Store Index

3. Process Query Enrich

LEADS APIs

SMALL COMPANIES

DISTRIBUTED ELASTIC MICRO-CLOUDS (APPEARS AS A SINGLE CLOUD)

LEADS

LEADS

LARGE-SCALE ELASTIC ARCHITECTURE FOR DATA AS A SERVICE

INNOVATIONS IN LEADS INCLUDE:

▶ **An elastic network of widely distributed micro-clouds.**

This environmentally friendly network of servers offers a high degree of reliability, scalability, availability and energy efficiency.

▶ **Complex geo-distributed load-balancing algorithms.**

These algorithms ensure that data is stored in close proximity to where it is used, reducing network traffic and latency and increasing speed and efficiency.

▶ **Confidentiality of user data.**

The project will use new cryptographic techniques (some developed within the project itself) for comparing, sorting and filtering user-encrypted data in the context of large public data sets.

▶ **Historical data.**

The storage system supports versioning of real time public and private data to allow historical data analytics and time-based comparisons.

The **LEADS** project began in October 2012, and will run for a total of three years.
It is funded by the Seventh Framework Program of the European Commission.

LEADS Project:

🌐 http://www.leads-project.eu/

f www.facebook.com/LeadsProject

🐦 twitter.com/leads_project

Seventh Framework Programme:

🌐 http://cordis.europa.eu/fp7/home_en.html

Project participants are:

adidas
(Germany)

AoTerra GmbH
(Germany)

Technische Universität Dresden
(Germany)

Barcelona Media/Yahoo!
(Spain)

Université de Neuchâtel
(Project Coordinator, Switzerland)

Telecommunication Systems Institute
(Greece)

Red Hat
(Ireland)

SEVENTH FRAMEWORK PROGRAMME