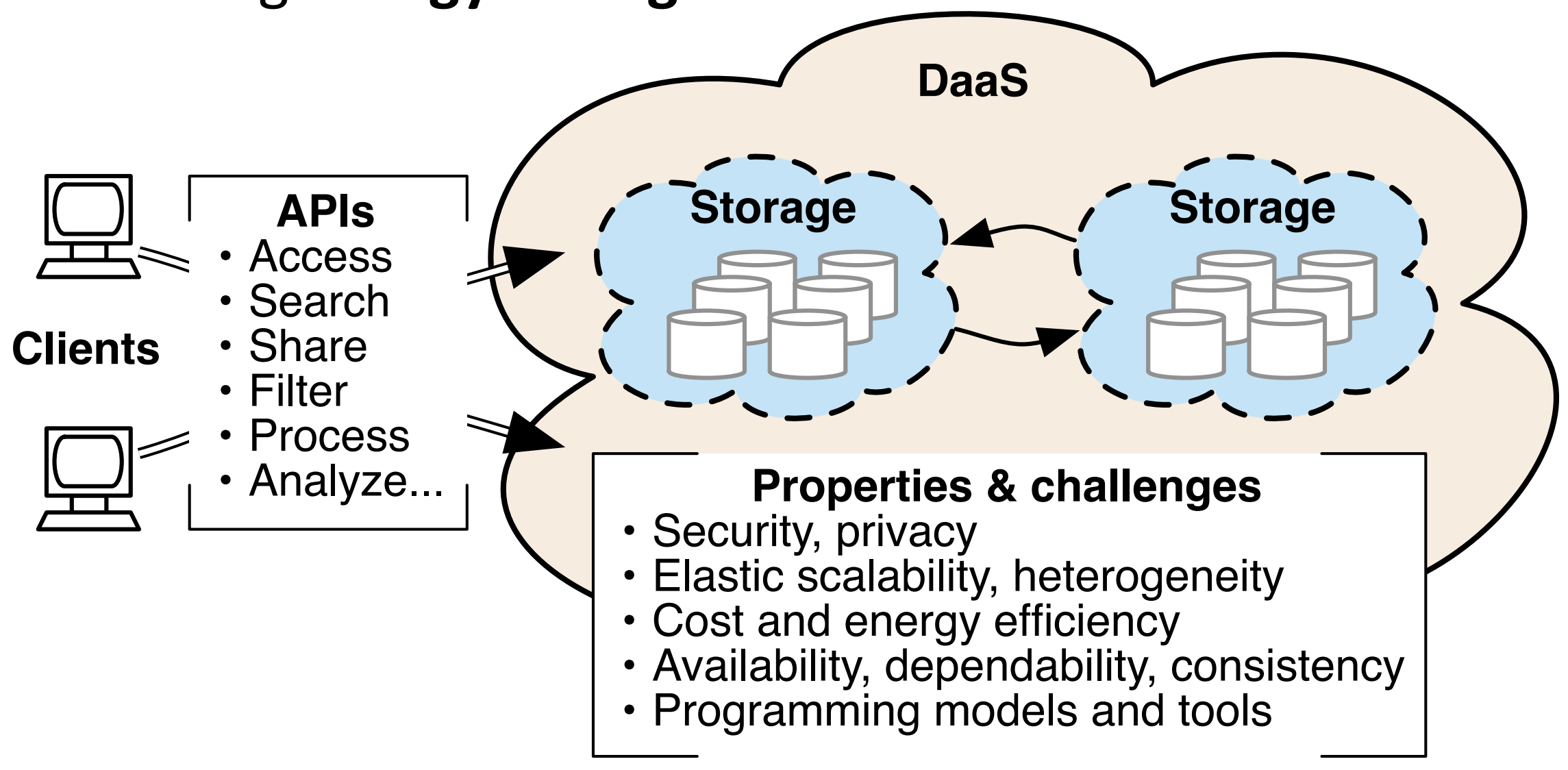
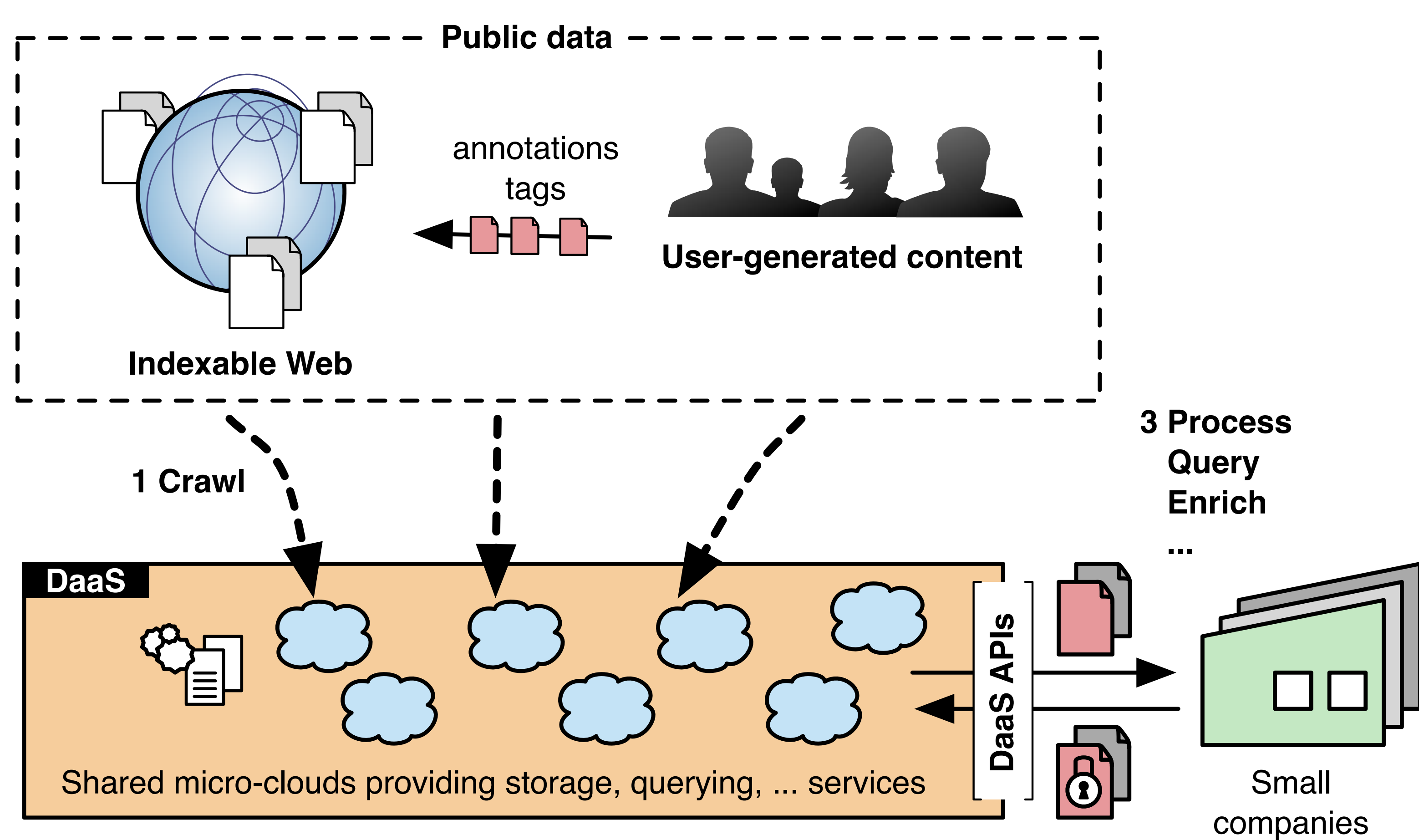


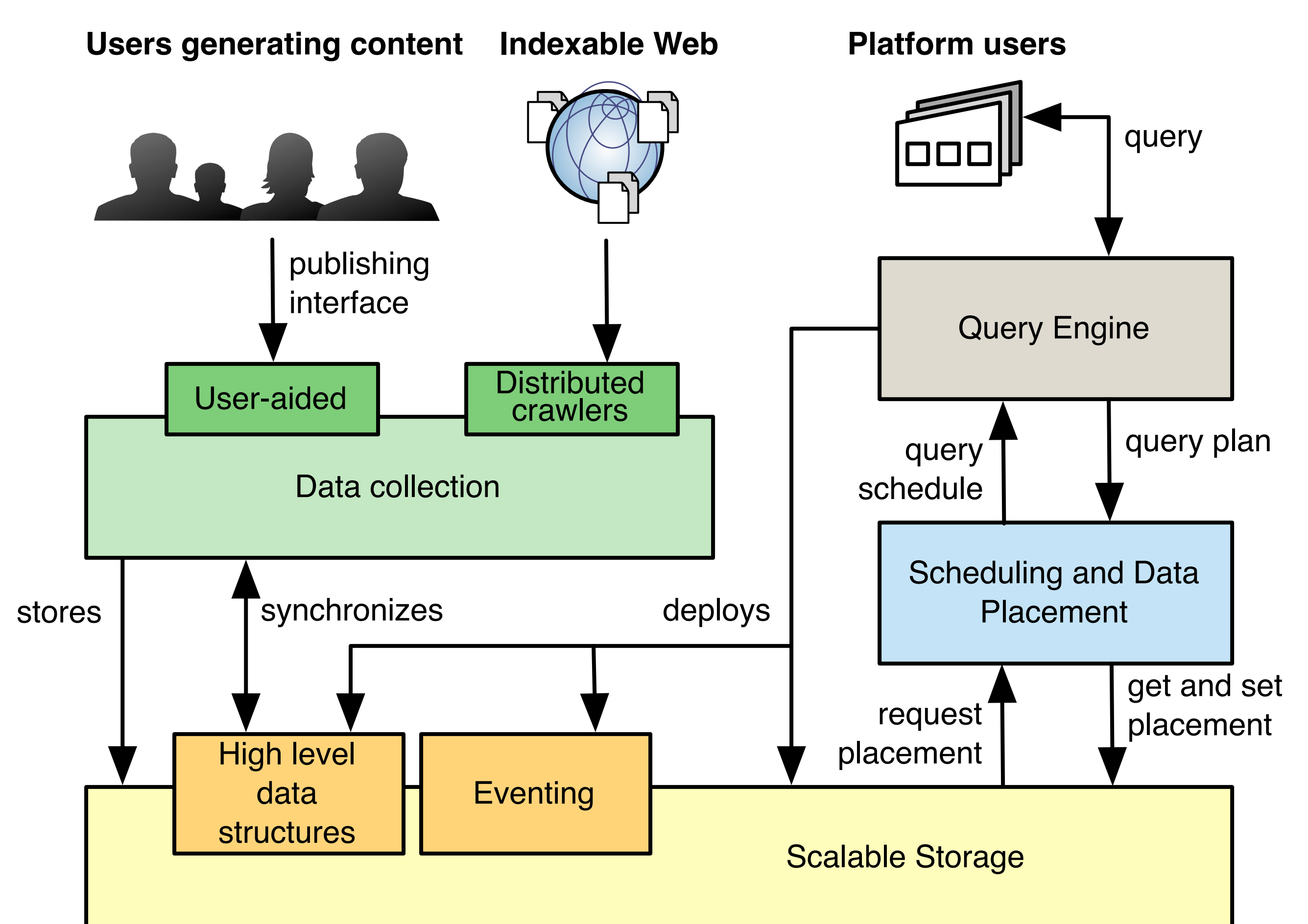


- 1 Motivation**
- Only the big players can crawl and process web-scale data!
 - SMEs need access to public data to offer innovative services
 - Find the trends for company's products
 - Detect correlations between events and sentiments
 - Find the most influential advertisement venues
 - Pay-as-you-go pricing
 - Exploit local energy-efficient micro-clouds

- 2 Goals**
- A novel **Data-as-a-Service (DaaS)** model to enable **SMEs** to process, analyze, and query **web-scale** data
 - A platform to enable DaaS
 - With a **pay-as-you-go** approach
 - **Sharing the cost** of indexing and processing of public data
 - Promoting **energy savings**



- 3 Unique features**
- **Green computing** using of excess heat to heat-up buildings
 - **Geo-distribution** of algorithms to reduce overall cost
 - Future cloud infrastructures: **elastic set of micro-clouds**
 - Processing of a combination of **streaming** and **historic** data
 - Support for **public** and **private** data
 - Combines **real-time queries** and **big data analytics**



4 Research challenges	Key Results
<p>Data collection</p> <ul style="list-style-type: none"> • Distributed crawling over the micro-clouds • User-aided data collection 	<ul style="list-style-type: none"> • Scalable and location-aware crawling • User-aided collection to reveal hidden web
<p>Scalable storage</p> <ul style="list-style-type: none"> • Data storage across micro-clouds • Robust and efficient • Secure for private data 	<ul style="list-style-type: none"> • Key Value Store over a cloud federation • Support of atomic data structures and events
<p>Query Engine</p> <ul style="list-style-type: none"> • Processing across micro-clouds • Streaming and historic data • Private and public data • Web data analytics 	<ul style="list-style-type: none"> • Access through well-known SQL syntax <ul style="list-style-type: none"> • Distributed execution over the cloud • Scalable • Selected queries over encrypted data • Link analysis across micro-clouds
<p>Scheduling and data placement</p> <ul style="list-style-type: none"> • Geo-distribution and scheduling • Both storage and processing • Move processing where heating is needed 	<ul style="list-style-type: none"> • Support for micro-cloud federations <ul style="list-style-type: none"> • Minimize cost • Maximize energy efficiency
<p>User support</p> <ul style="list-style-type: none"> • Graphical User Interfaces • Query models • Application programming interface 	<ul style="list-style-type: none"> • Graphical access to the web graph <ul style="list-style-type: none"> • Ability to enrich with user-generated data • API support <ul style="list-style-type: none"> • Query execution and data management

- 5 Technologies**
- Open-source projects with a large user base
-



CLOUD COMPUTING
LEADS – Large-scale Elastic Architecture for Data As a Service
<http://www.leads-project.eu/>

PROJECT DATA
Start date: 01-Oct-2012
Duration: 36 months
Funding: 2.89 M€

Contact:
Coordinator
Etienne Rivière
Université de Neuchâtel
Switzerland
etienne.riviere@unine.ch

Presenter
Evangelos Vazaios
Technical University of Crete
Greece
vagvaz@softnet.tuc.gr