Pythia: A Privacy-enhanced Personalized Contextual Suggestion System for Tourism

George Drosatos

Pavlos S. Efraimidis, Avi Arampatzis, Giorgos Stamatelatos and Ioannis N. Athanasiadis

Institute for Language and Speech Processing
Athena Research and Innovation Center

Department of Informatics
Aristotle University of Thessaloniki

Department of Electrical and Computer Engineering
Democritus University of Thrace

This work was supported by the project ATLAS (No. 11SYN-10-1730), funded in part by the European Commission and Greek national funds.

Summary

Summary of this work

- We present Pythia, a privacy-enhanced non-invasive contextual suggestion system for tourists
  - high quality personalized recommendations
  - non-invasive operation
  - protection of user privacy
- Make contextual suggestions (POIs) based on automatically built user profiles
- The user profile is POI-based using as personal data from user's Smartphone:
  - Location traces
  - Browsing history
  - Web searches
- Strong privacy guarantees are achieved by placing both mechanisms at the user-side
- Proof of concept: a prototype implementation of Pythia
  - as an Android application, as well as a web application

- Background
  - *Privacy*
  - *Tourism recommendations*

Privacy & Legislation

- **What is privacy?**
  - "The right to be let alone" [Warren and Brandeis, 1890]
  - "The right of the individual to decide what information about himself should be communicated to others and under what circumstances" [Westin, 1970]
  - The right to informational self-determination [1983]

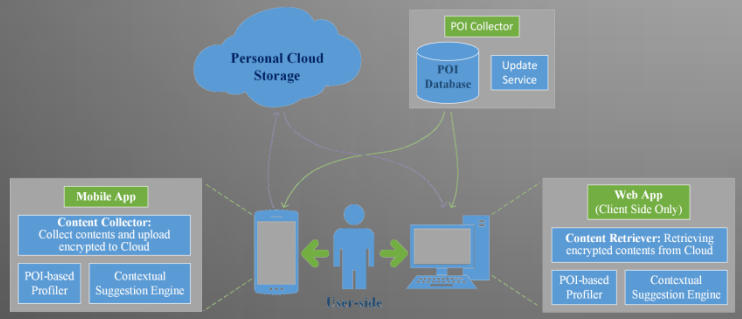
- **Personal Data**: Any information that refers to a person

- **Related Legislation**: e.g. EU Data Protection Directive 95/46/EC
  - Indicative principles:
    - Reported and transparent processing
    - Finality & Purpose Limitation
    - Personal data quality
    - Security
    - Personal data traffic outside EU

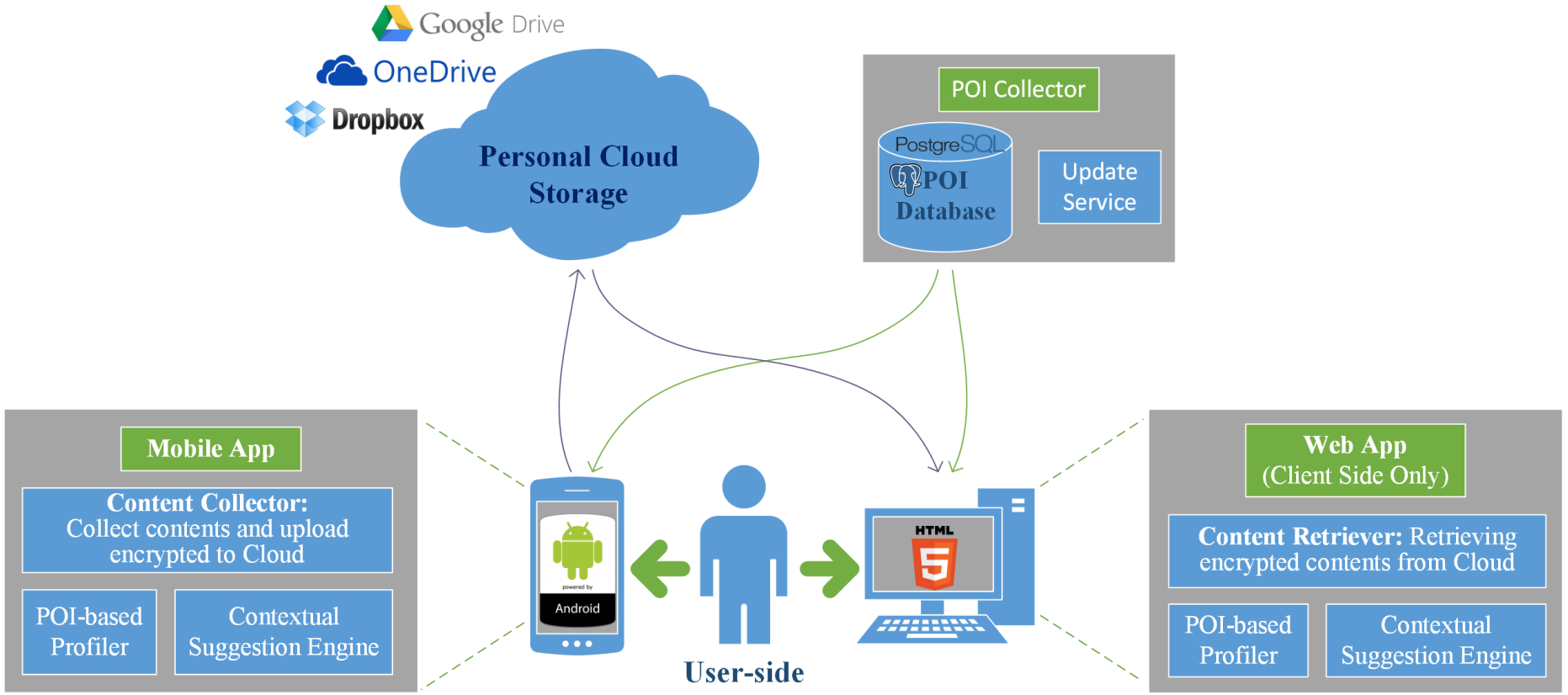
Recommender Systems in Tourism

- A recommender system requires some input data in order to make appropriate suggestions
  - Simple preferences: needs, interests and constraints
  - Evaluations and ratings of other tourists (collaborative filtering)
  - Travelers who are in close spatial and temporal proximity often share common travel interests
- Common practice of such systems
  - Users have to enter their personal information into some system
  - User profiles have to be stored and managed by the recommender service
  - The recommendations are performed in server-side
- The usage of personal data concerns privacy issues
- Pythia system combines a set of unique features in the field of mobile recommendation systems
  - User-centric architecture
  - Privacy-preserving
  - Rich user profiles
  - Non-invasive creation of user's profile

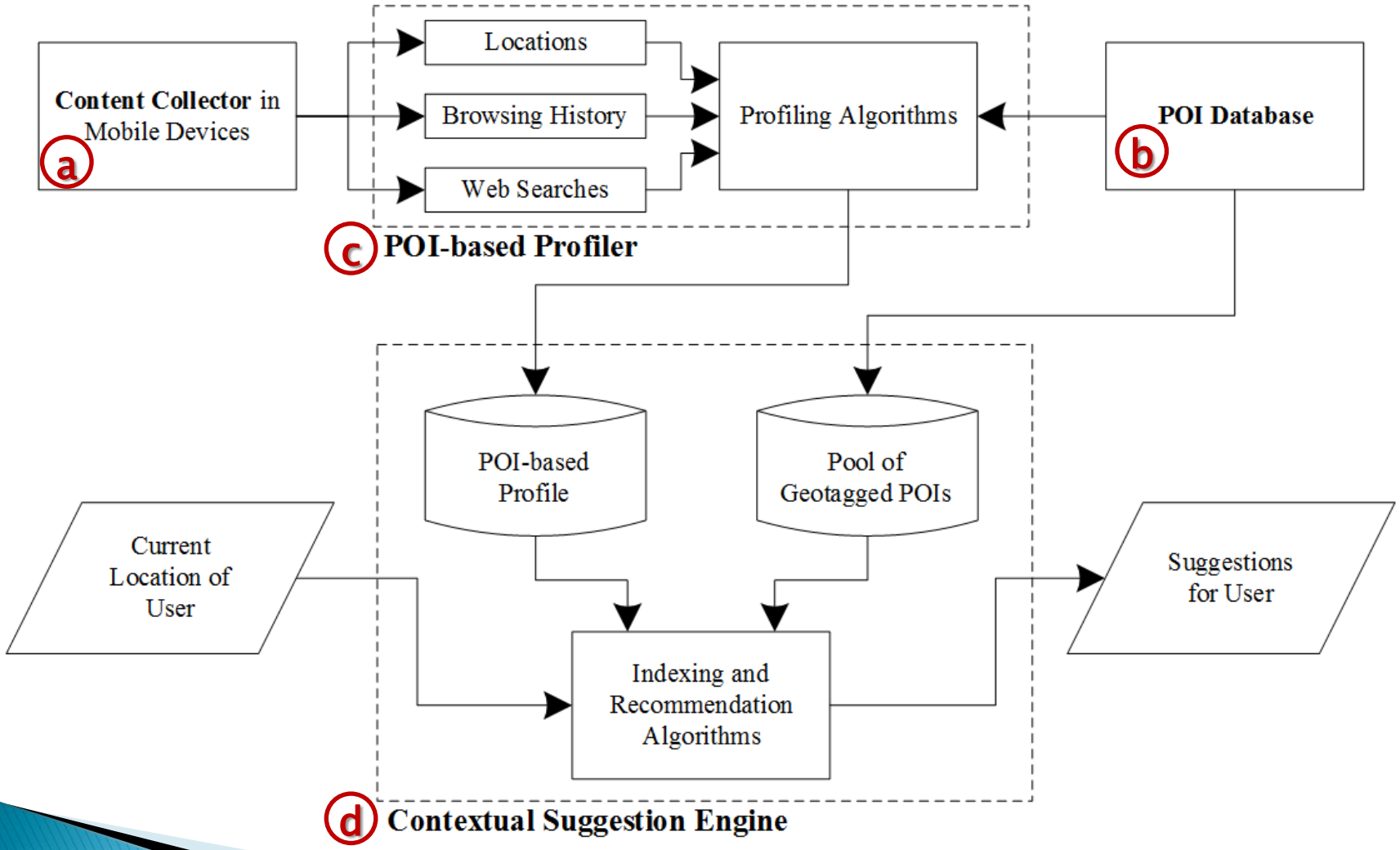
The Pythia System



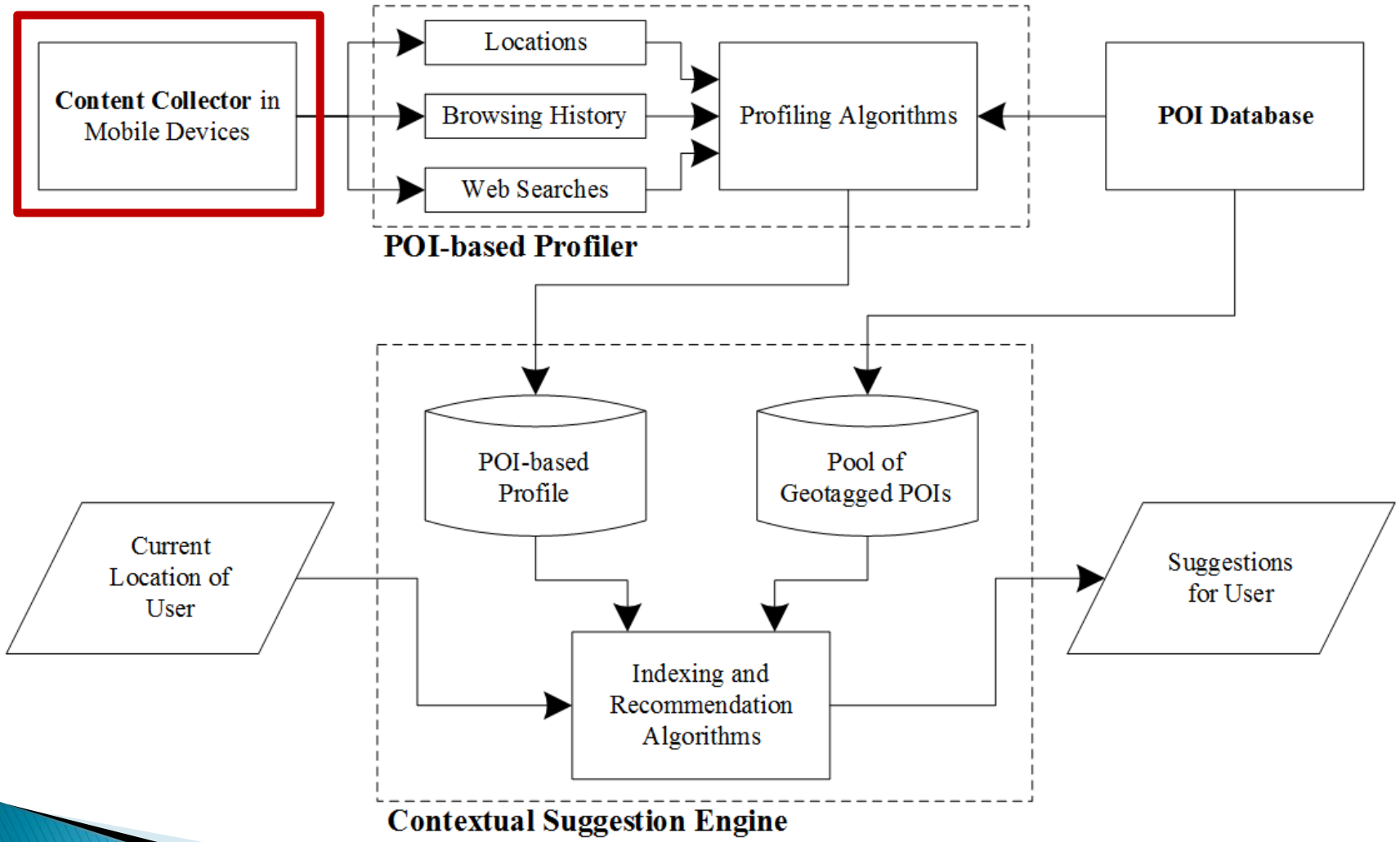
The architecture of the Pythia system



Interaction of Pythia components



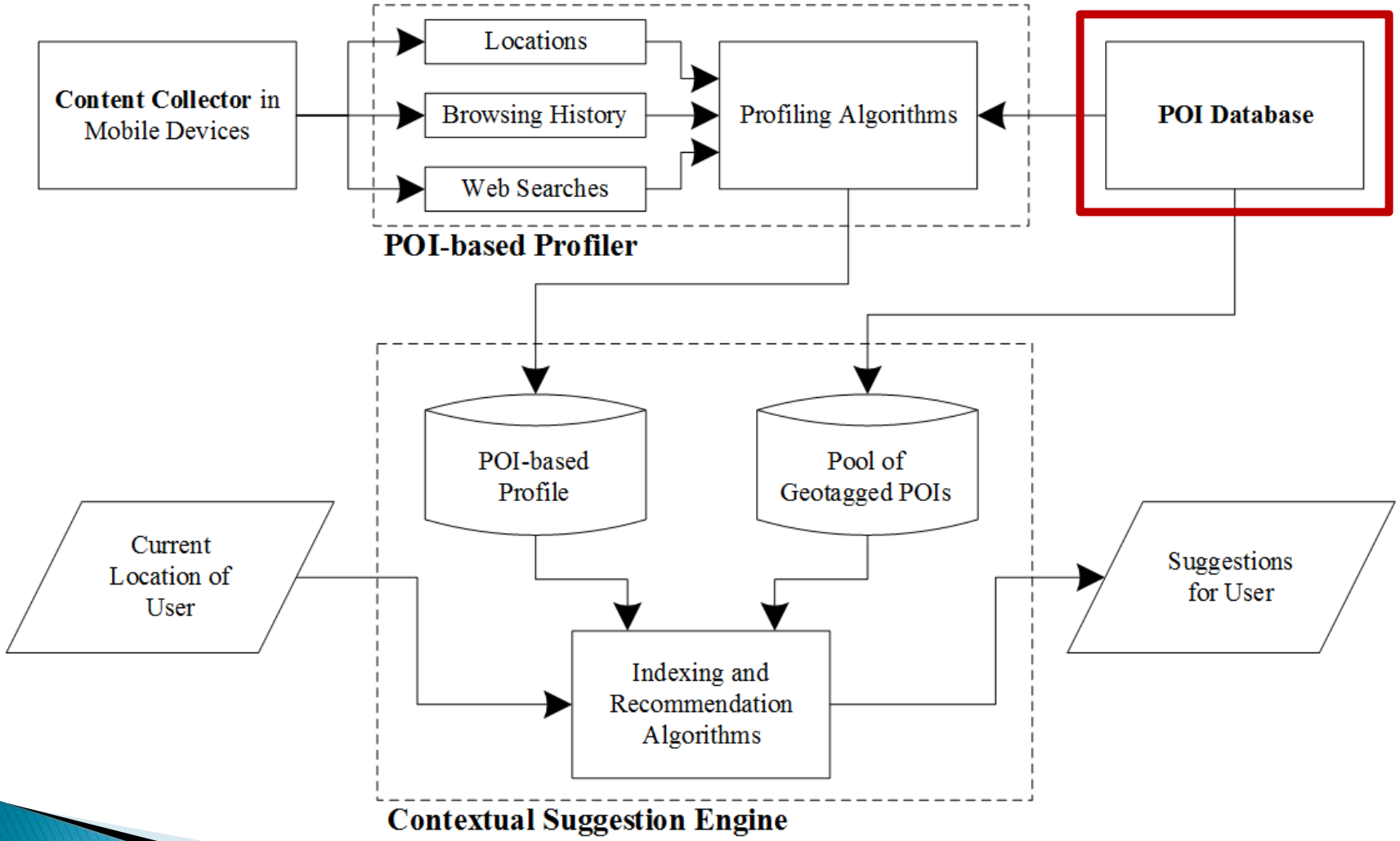
ⓐ Personal Data Collection



ⓐ Personal Data Collection

- Content Collector collects personal data of the digital trace generated by the everyday usage of the smart devices
- This data is stored locally on each device
- Focus on the following types of personal data:
  - **Location traces:** Records of locations in a particular time frequency
    - [Latitude, Longitude]: GPS or Network provider
    - Speed, Altitude, Bearing
    - Accuracy
    - Time
  - **Browsing history:** Records of browsing history
    - Title
    - URL
    - # of visits
    - Time of last visit
  - **Web searches:** Query logs
    - Query
    - Submission time

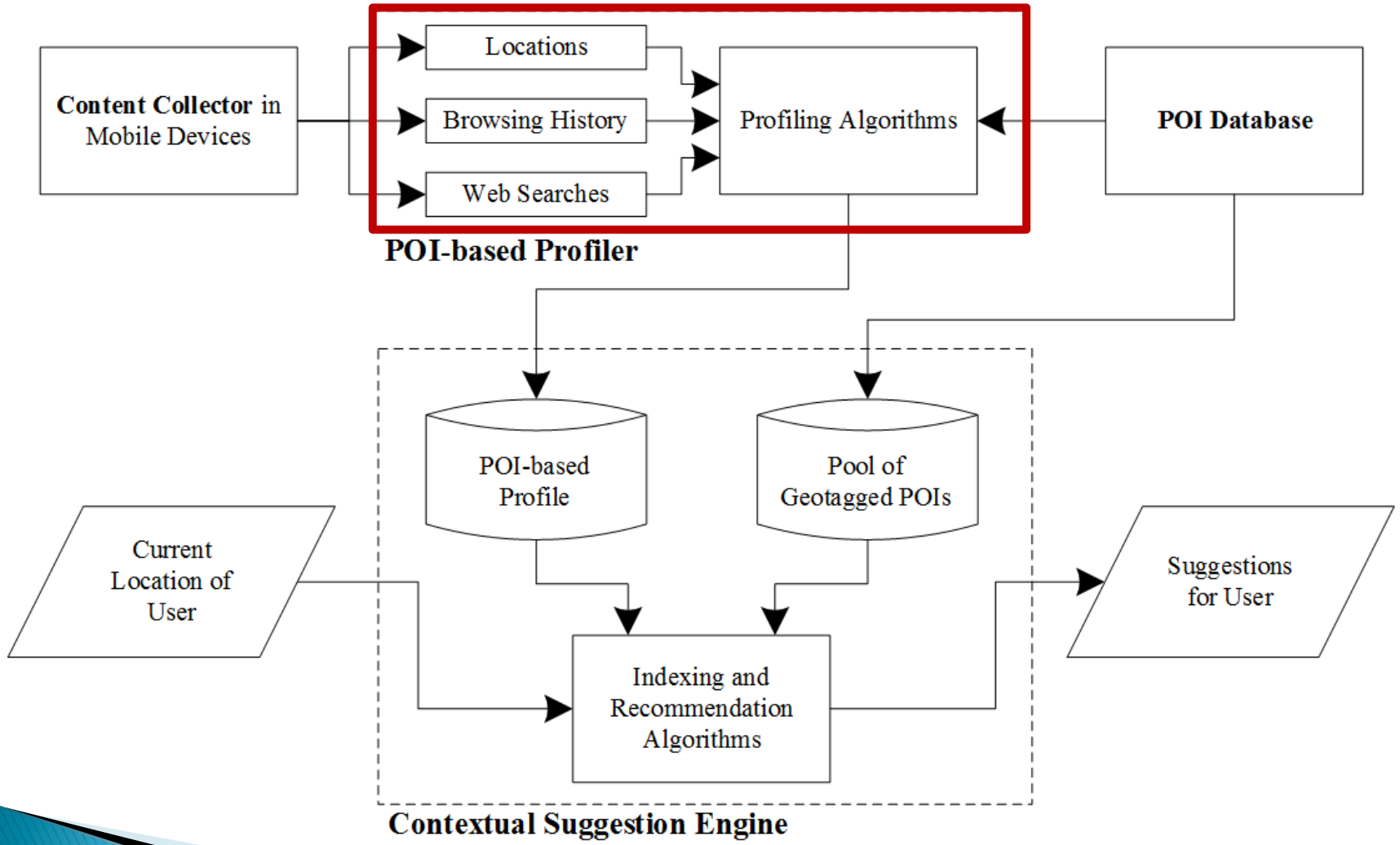
ⓑ POI Collection Framework



ⓑ POI Collection Framework

- This framework implements a Web API which is used by other Pythia components to access the POI database
- The POIs can be retrieved via API:
  - one by one
  - as chunks corresponding to a specific county or area
- The data available for each POI are:
  - title
  - geo-location
  - address
  - phone number
  - categories
  - several URLs
  - rating
  - total unique visitors and total visits
  - a collection of terms describing the POI
- An update service continuously updates and extends the database with new information about POIs from Foursquare and Google Places

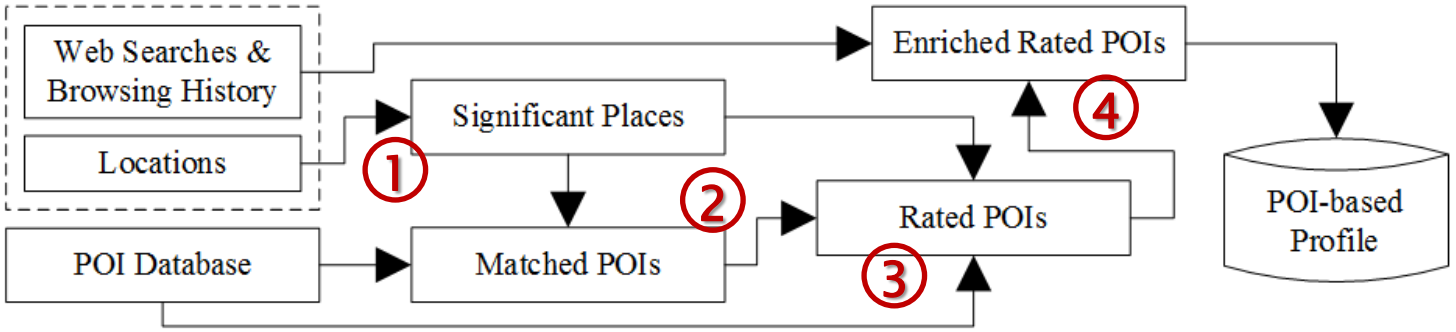
© POI-based Profiling Process



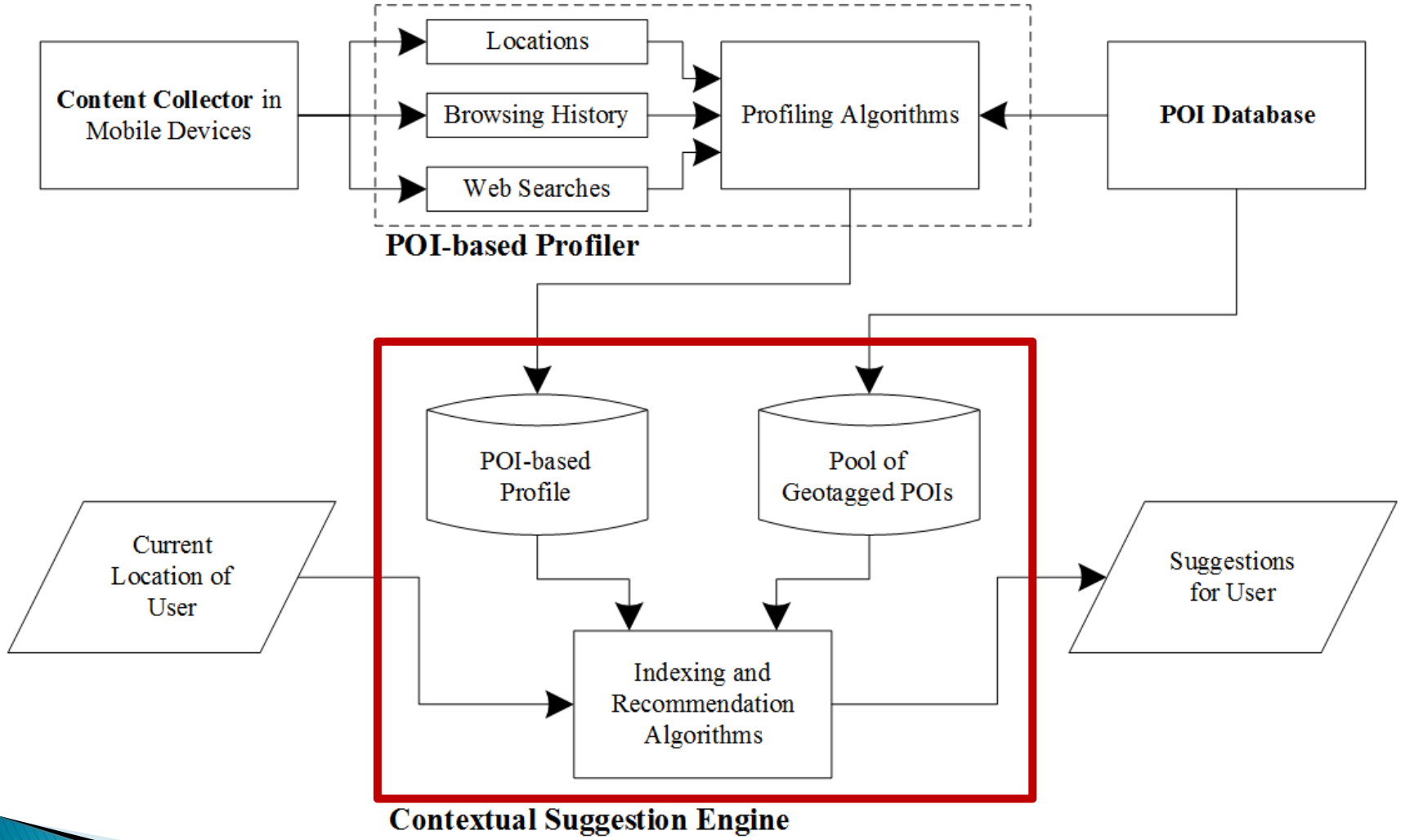
ⓒ POI-based Profiling Process

- Data processing steps:

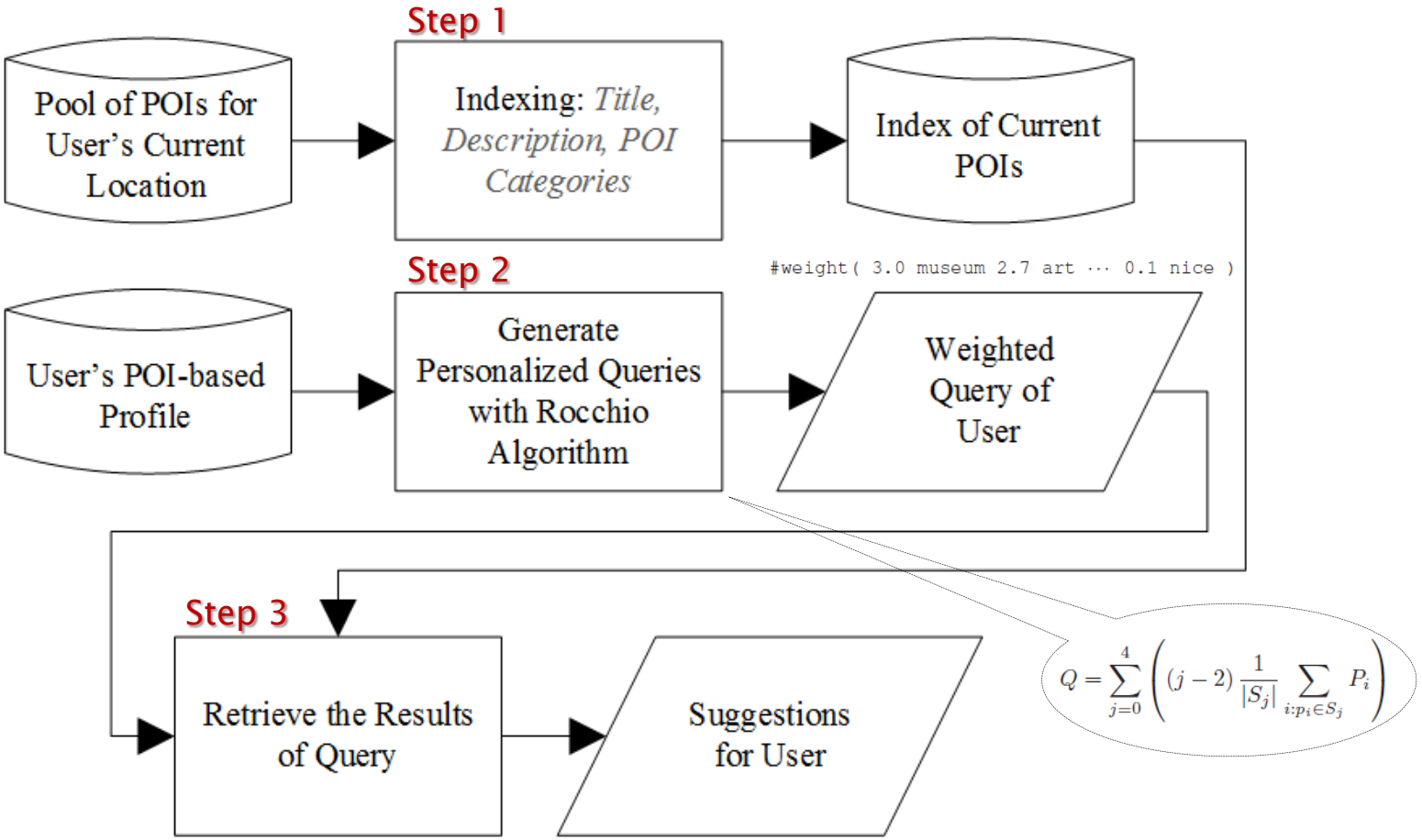
1. **Significant places extraction**: Apply time-based clustering on location traces
2. **POI matching**: Match significant places with POIs from the POI database
3. **POI rating**: Estimate the user's interest for the POI based on the user's number of visits and the average number of visits per user to this POI
4. **Enrichment of Rated POI**: Adjust the POI ratings based on the web searches and titles of browsing history by applying an information retrieval approach



The Rocchio-like Contextual Suggestion model*



ⓓThe Rocchio-like Contextual Suggestion model*



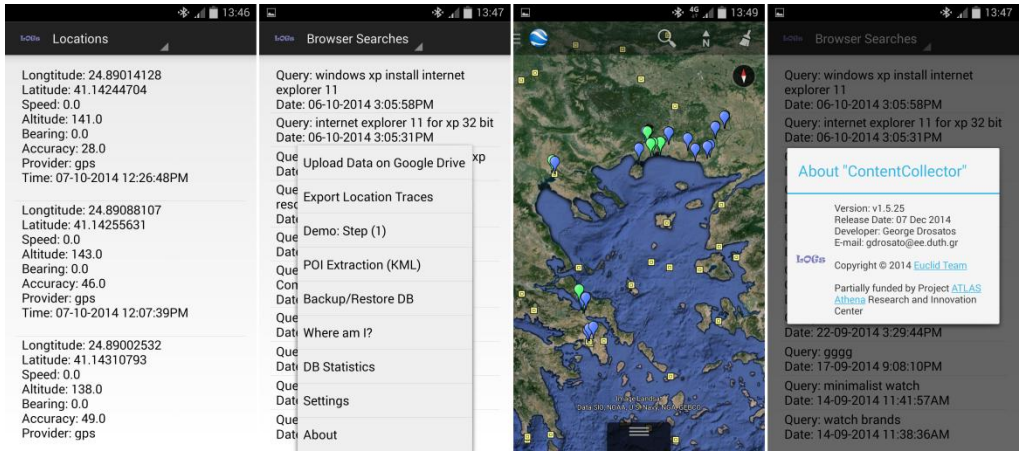
** TREC 2013, Contextual Suggestion Track: We ranked with this algorithm at 2nd place among 15 research groups*

The Pythia Prototype

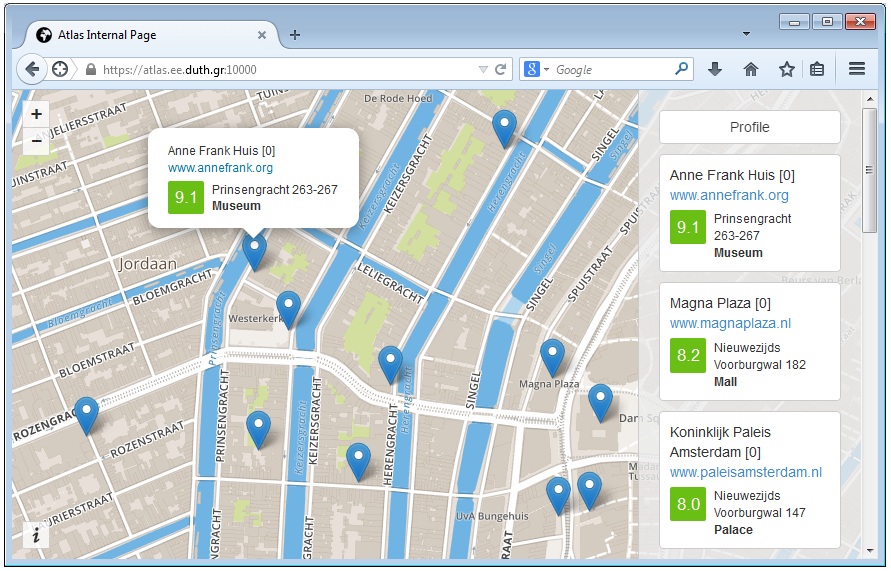
Current status of implementation

- Fully implemented
  - POI Collector (PHP, PostgreSQL)
  - Content Collector (Android platform)
- Only implemented as web application
  - POI-based Profiler (JavaScript)
  - Contextual Suggestion Engine
- The current prototype are implemented as a collection of standalone applications
  - Communicate through the personal Cloud storage of the user

Content Collector



POI-based Profiler & Contextual Suggestion Engine



Conclusions

Conclusions

- Proposed the Pythia system: a privacy-enhanced non-invasive contextual suggestion system for tourists
- The main system components operate in the background without user interaction
- The personal data of the user are kept on his own device
- Utilize sensitive personal data for personalized suggestions without violating the individuals' privacy
- The protection of user privacy is achieved by placing the profiling and recommendation procedures at the user-side
  - Mobile device OR desktop
- Developed prototype implementation

- Future work
  - Improve the automatic POI matching and POI rating algorithms
  - Upgrade the implementation to a more complete and stable system
  - Perform a more extensive evaluation of the complete system

Thank you,

» any questions ?

Acknowledgement

ATLAS: Advanced Tourism Planning System.
http://atlas.web.auth.gr/

This research was supported by the **European Union** (European Social Fund - ESF) and **Greek national funds** through the Operation Program "Competitiveness-Cooperation 2011" - Research Funding Program: **11SYN-10-1730-ATLAS.**