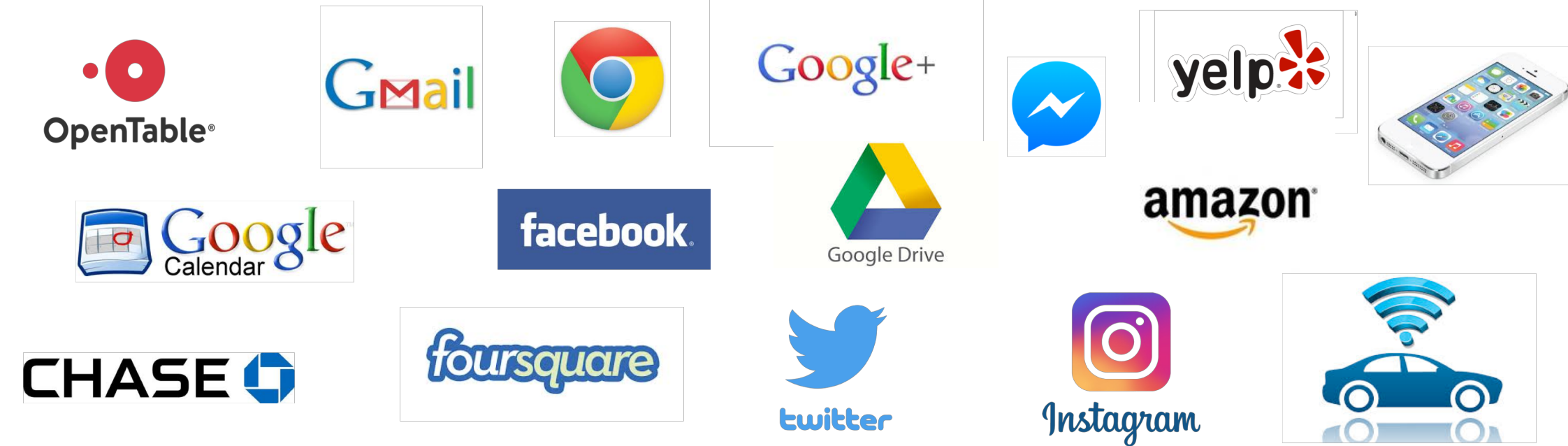
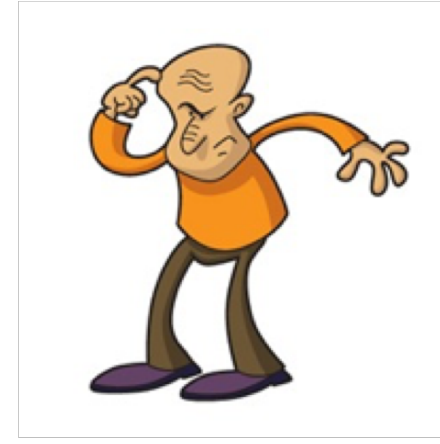


1. Motivation

- What was the name of that restaurant where I was
 - with Mary?
 - last Month?
- Some sources of helpful data
 - "with Mary": calendar, emails, text
 - "restaurant": check-ins, cell phone GPS logs, credit card statements, reservations
- Problem: Personal data is fragmented, heterogeneous**



2. Background

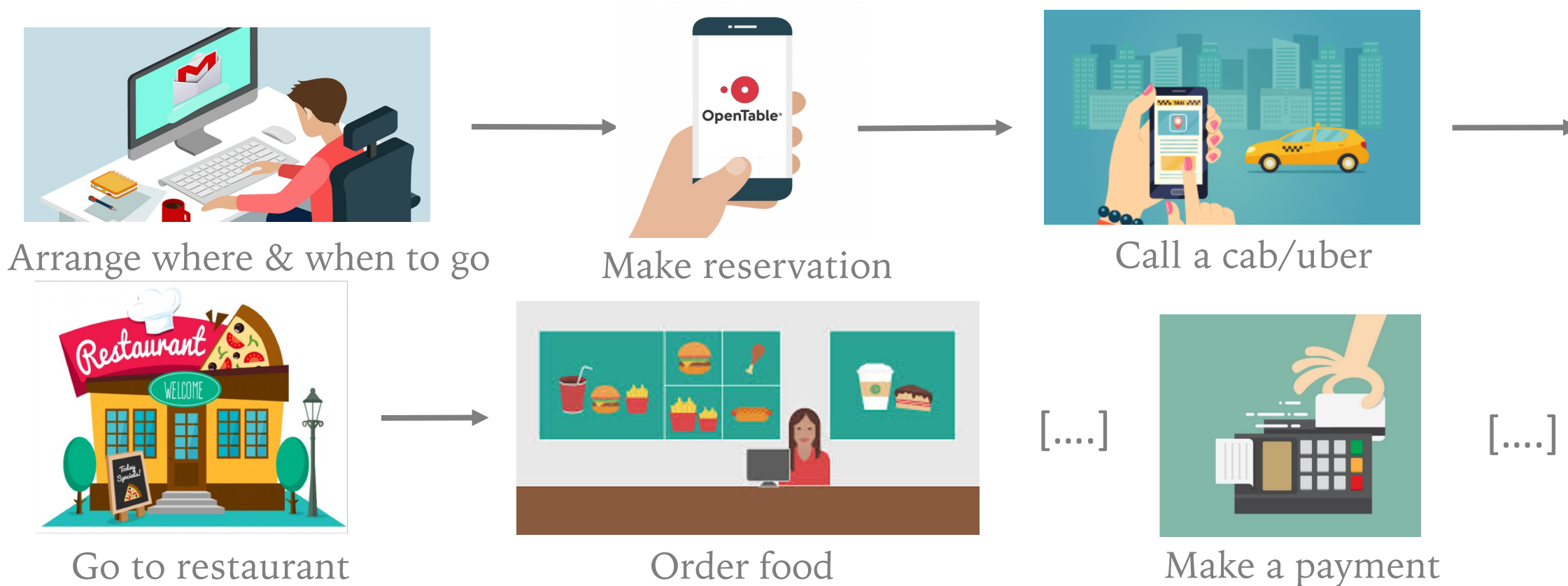
- Research in psychology: **Episodic memory** – memory of autobiographical events.
- Scripts** – stereotypical plans whose instances form part of the episodic memory.
- Natural way to remember past events is by pertinent contextual information
- Answers to: **What, When, Where, Who, Why, How (w5h)**

3. Contribution

- Goal:** develop tools and techniques to organize, summarize, make inferences and personalize **personal digital traces (PDTs)**
- Current work's contributions:
 - Episodic script model** to link and represent PDTs and their episodes. ODBASE '17
 - Algorithm** to find and combine evidence from PDTs for creating episodic memory instances and filling properties. ExploreDB '17@SIGMOD/PODS
 - Case study & experimental evaluation** : "Going out to eat at a restaurant"
 - Interactive tool** with narrative views of users' digital memories. CIKM '18

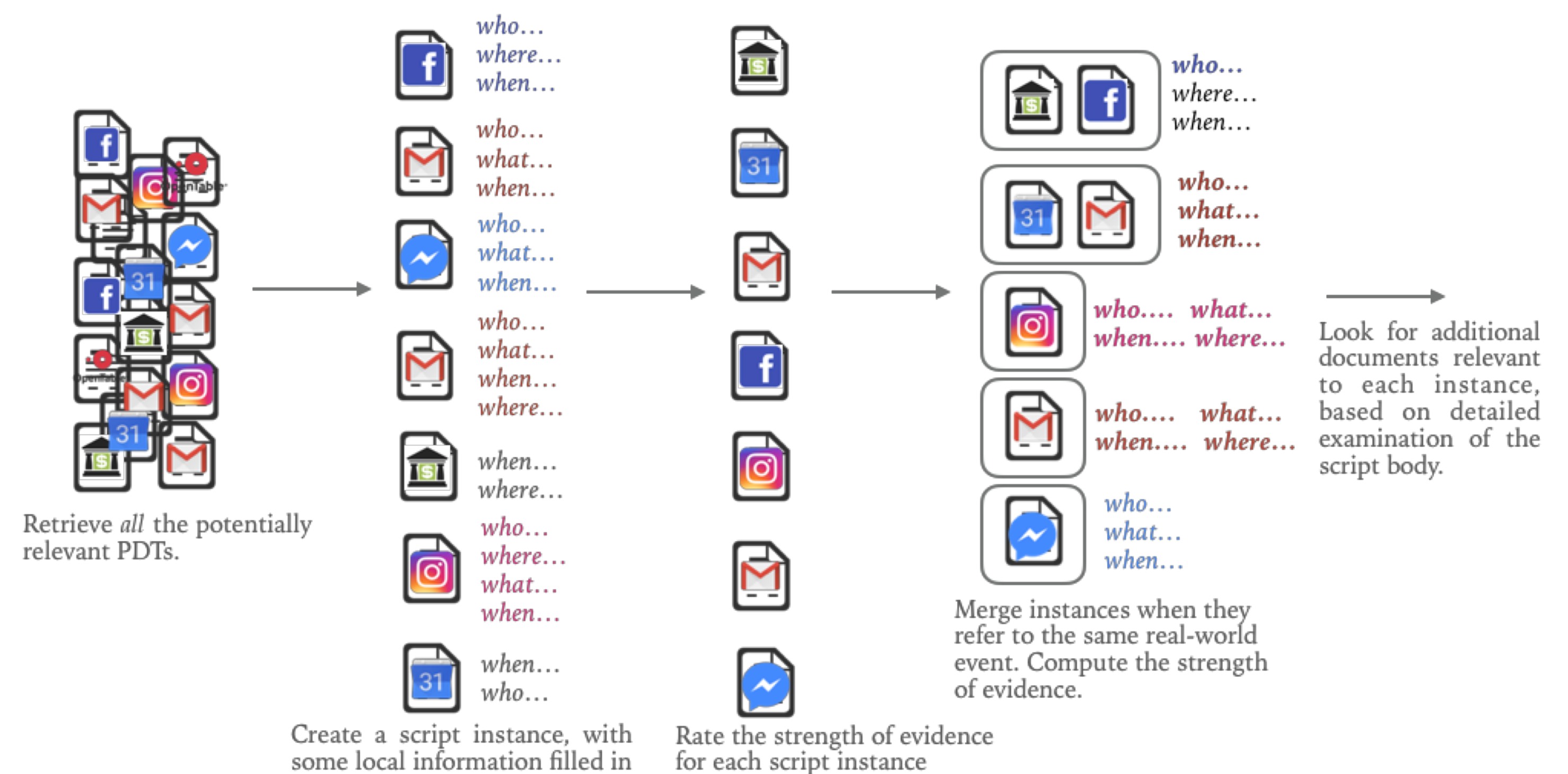
4. Ontology for Scripts

- Atomic event:** described by w5h, extractable from docs
- Script:** describes possible event/sub-script flows & aggregates w5h information from them

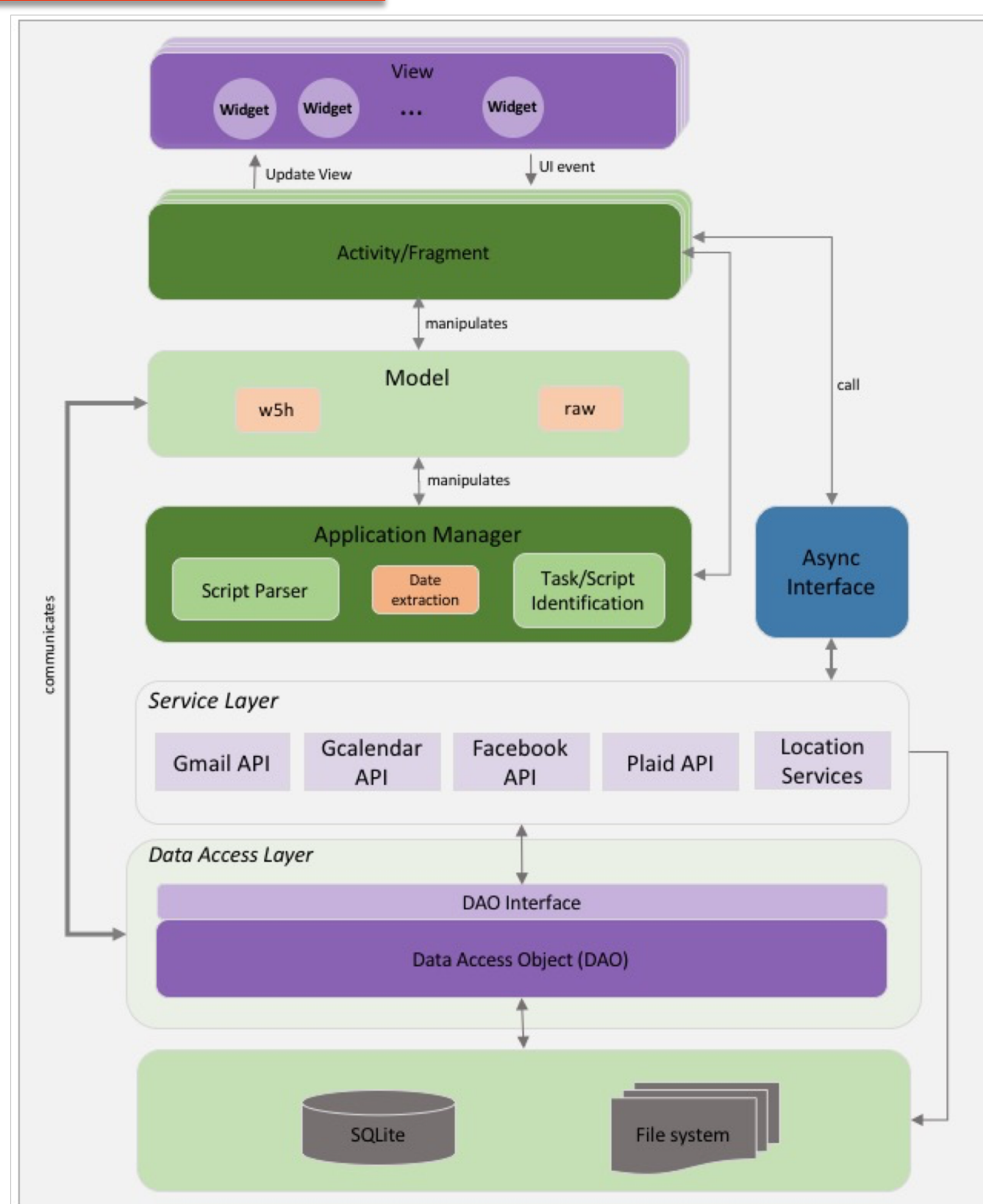


- Candidate script instances** created from actual docs and scored for likelihood

5. Algorithm

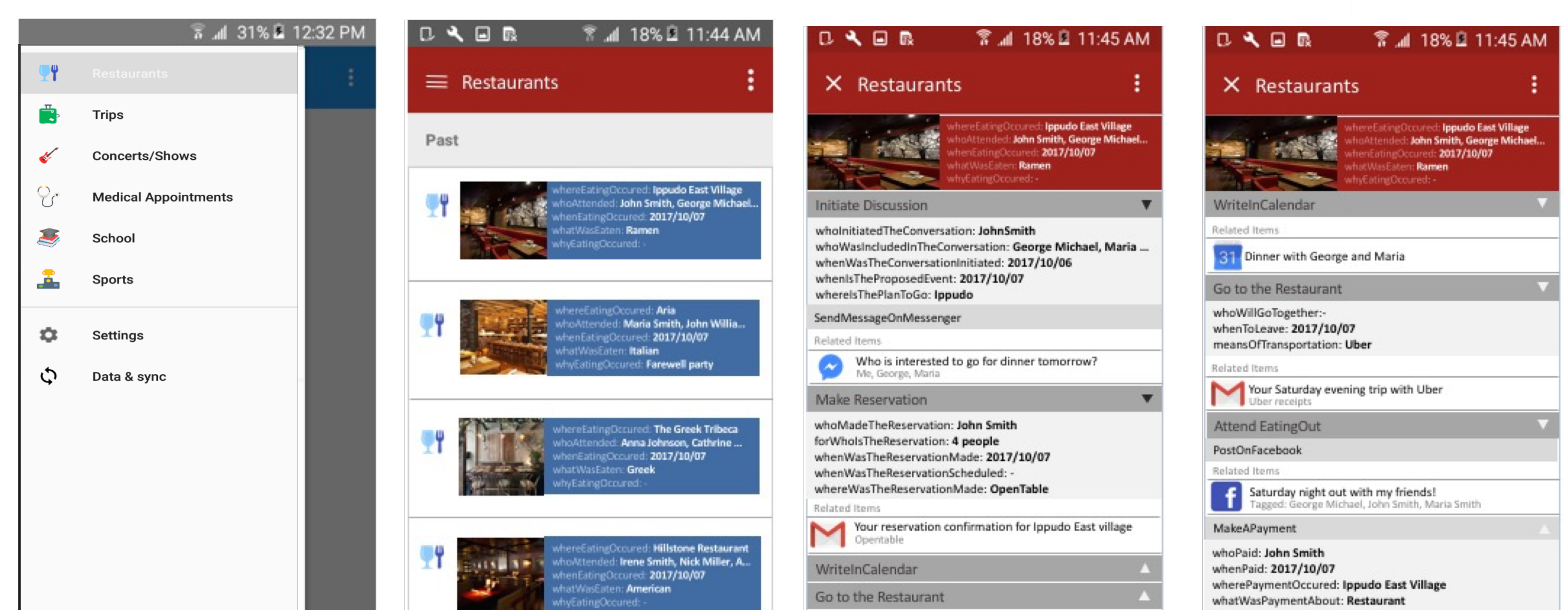


6. Architecture



7. YourDigitalSelf

- A Personal Digital Trace Integration tool - Android app



8. Discussion

- Any approach to retrieve user memories of events must consider several sources to adapt to the wide variety of user behaviors.
- The quality of the information given by different sources vary.
- Problems:**
 - Misclassified "restaurants" in bank statements -Use of Google Maps -Partial success
 - no NLP, so we miss "cannot make it for dinner"
 - Personalization issues: each person uses PDTs consistently but very differently (e.g. shared bank accounts)
 - Constantly changing APIs.

