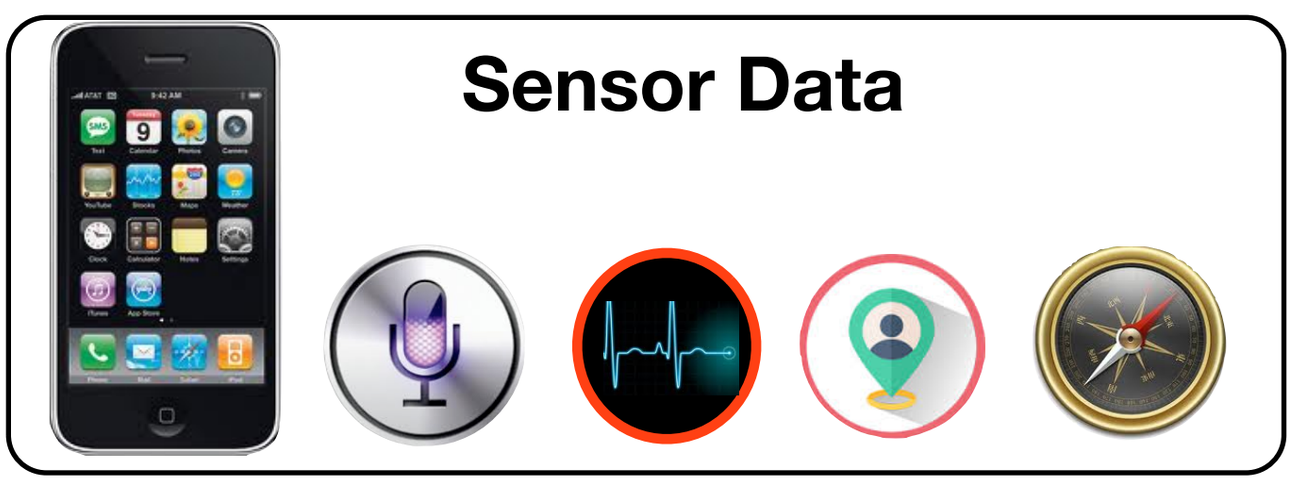
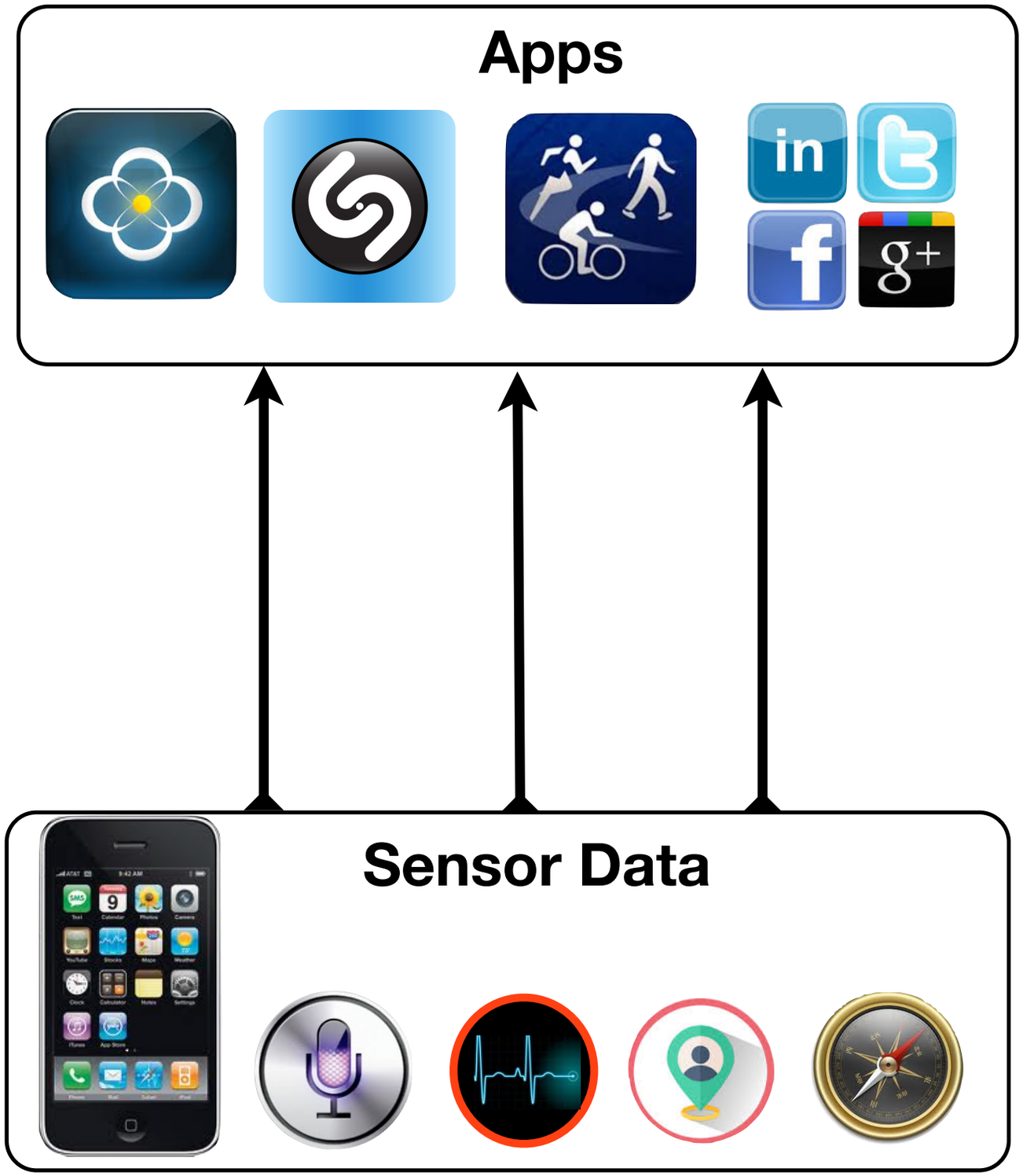
# ipShield: A Framework For Enforcing Context-Aware Privacy

**Supriyo Chakraborty**, Chenguang Shen, Kasturi Rangan Raghavan, Yasser Shoukry, Matt Millar, Mani Srivastava

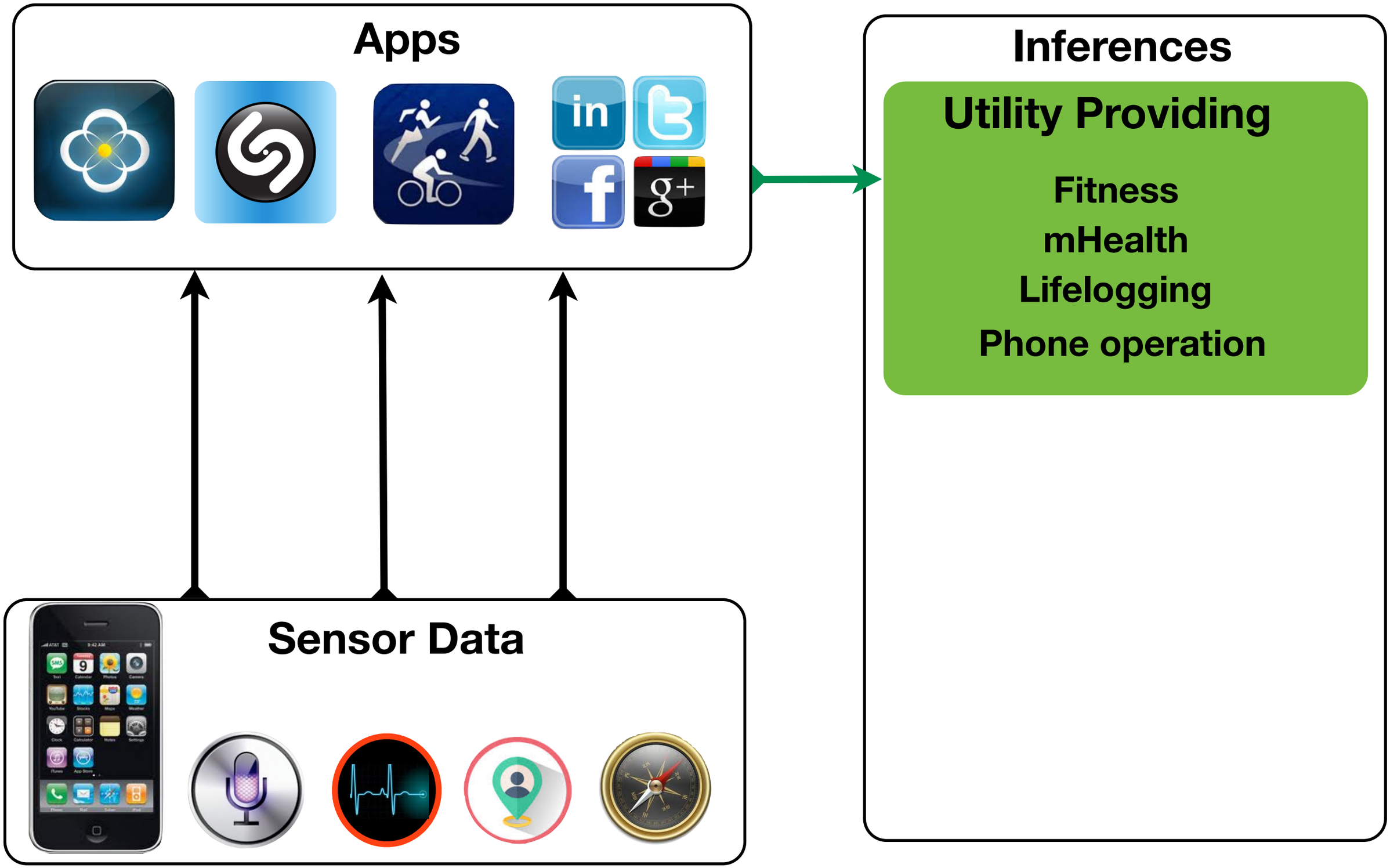
# From sensor data to inferences



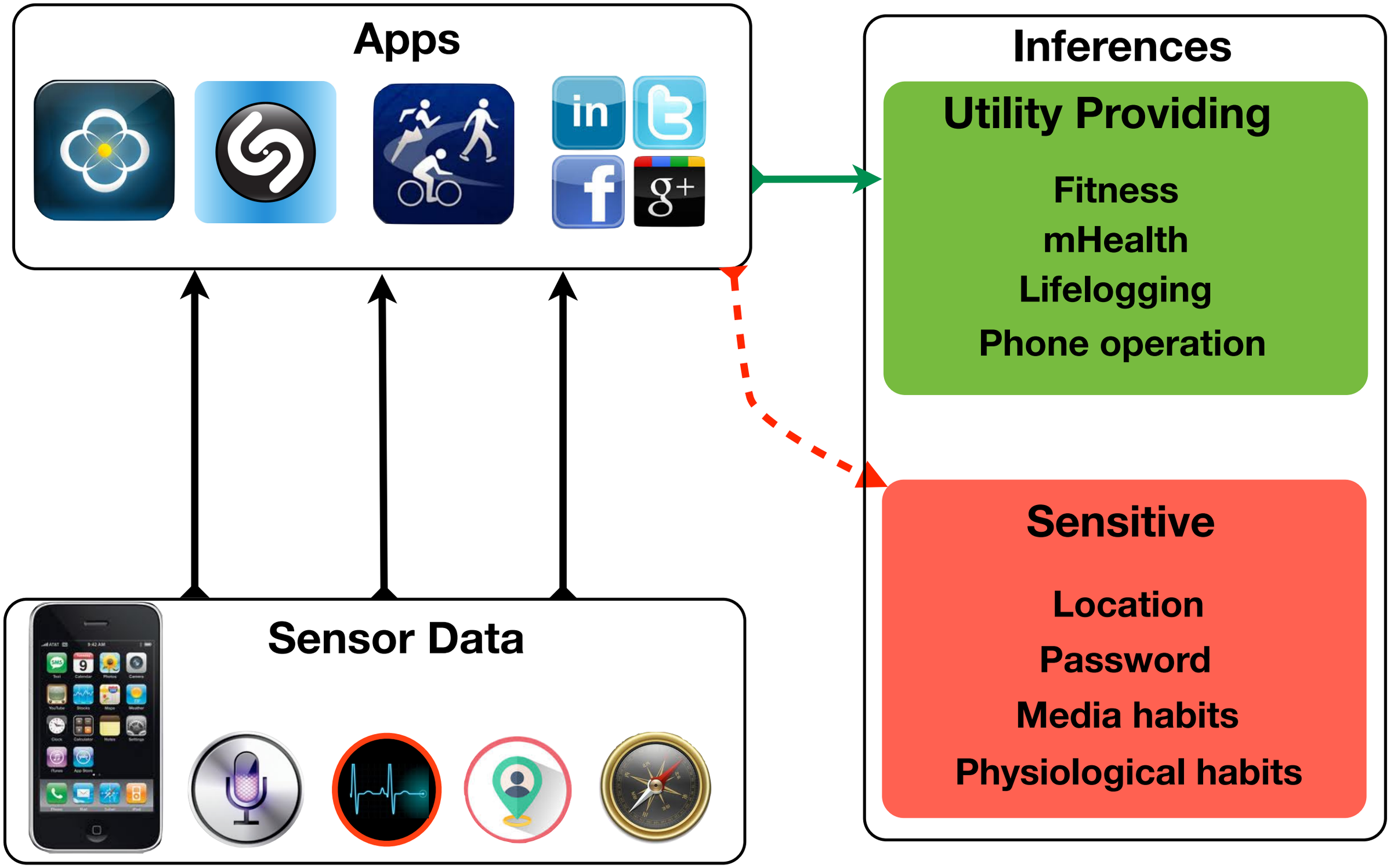
# From sensor data to inferences



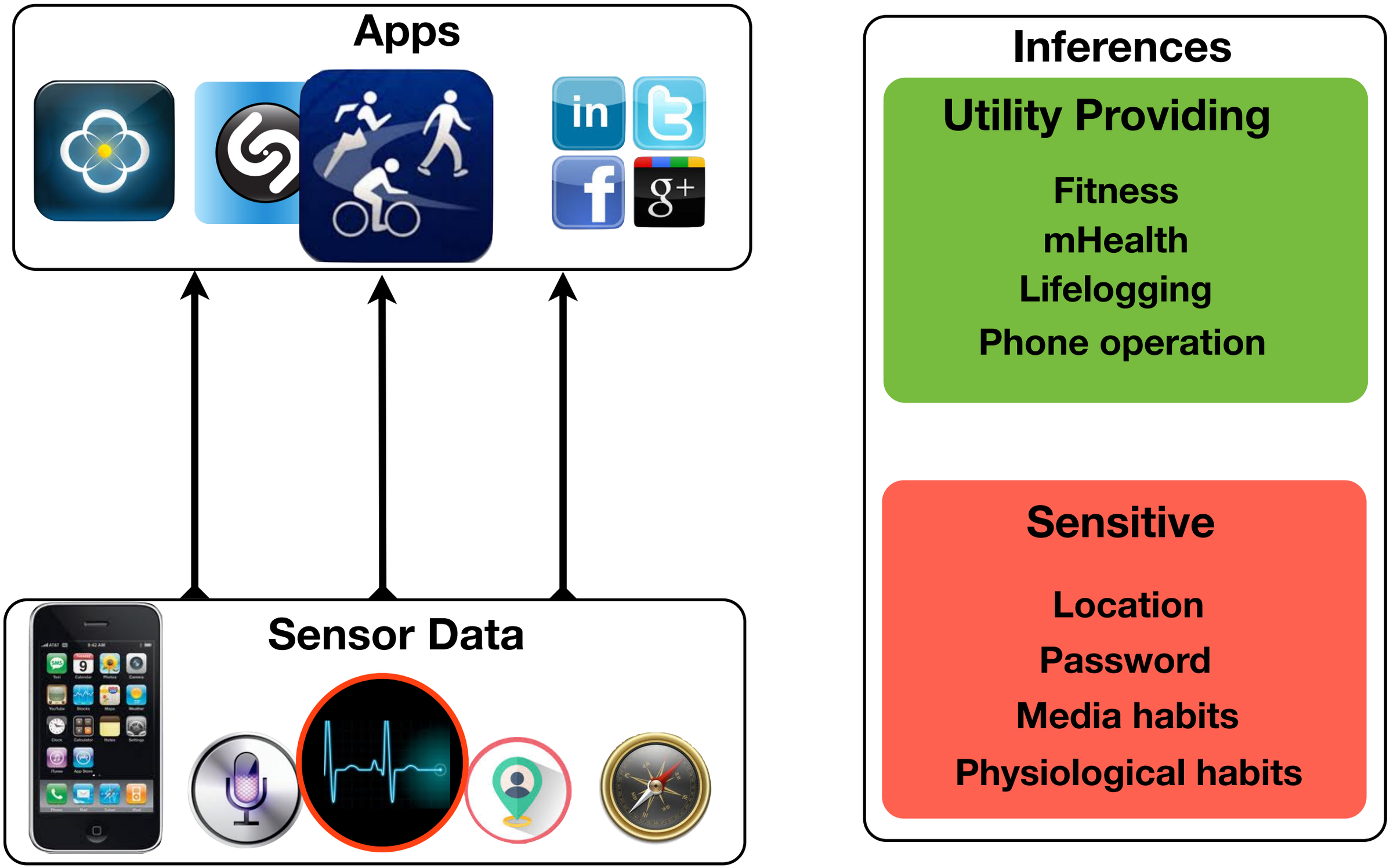
# From sensor data to inferences



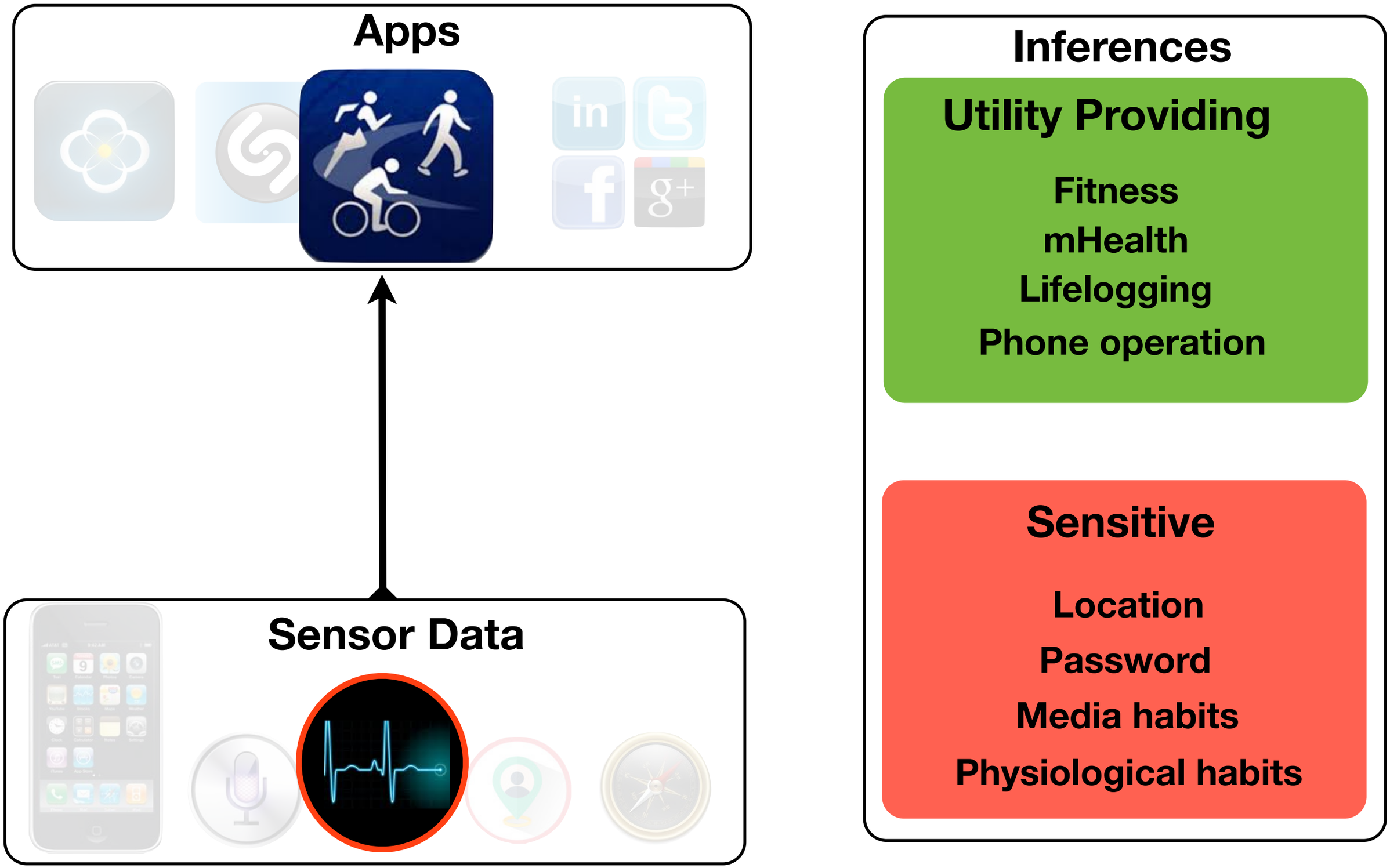
# From sensor data to inferences



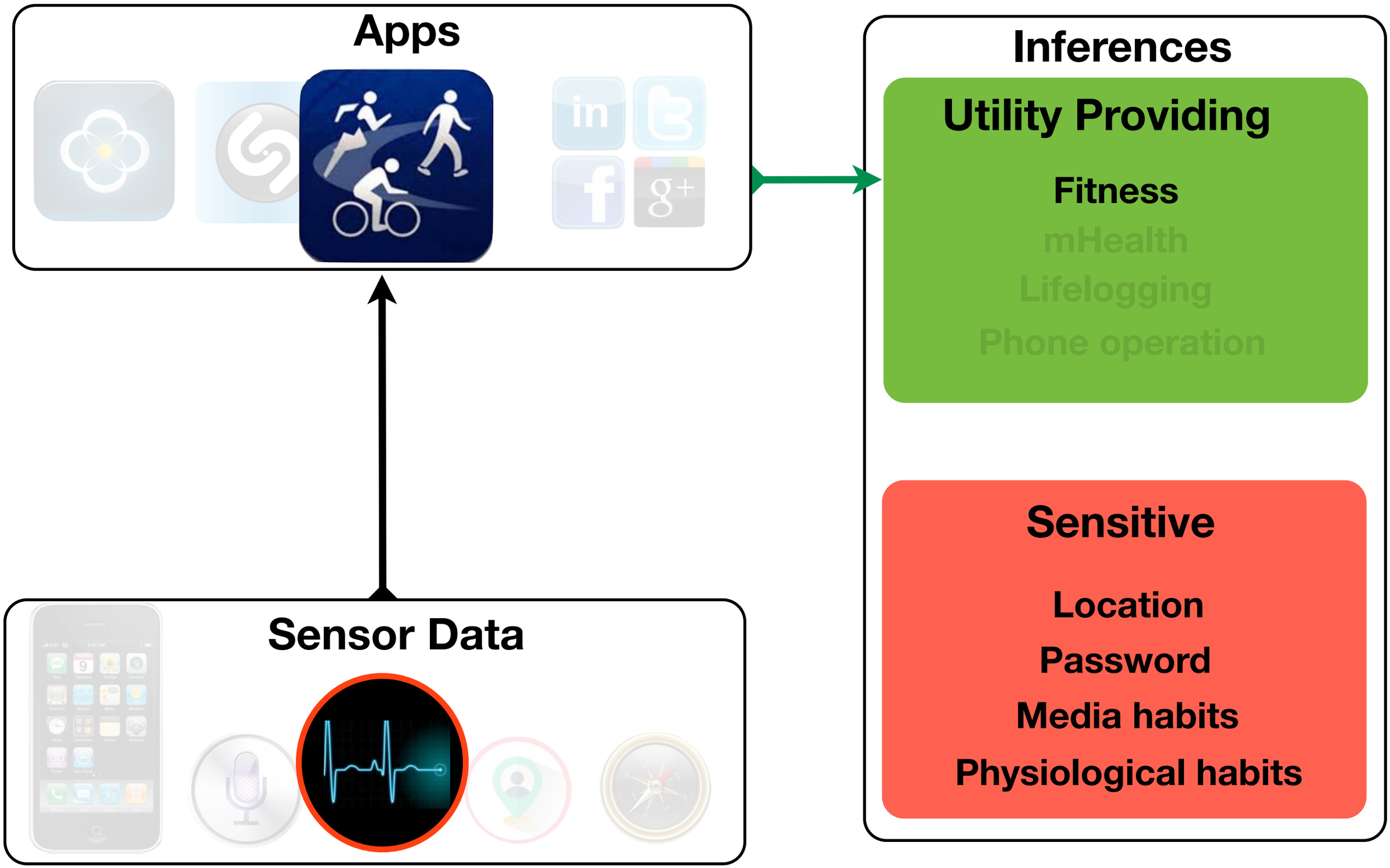
# From sensor data to inferences



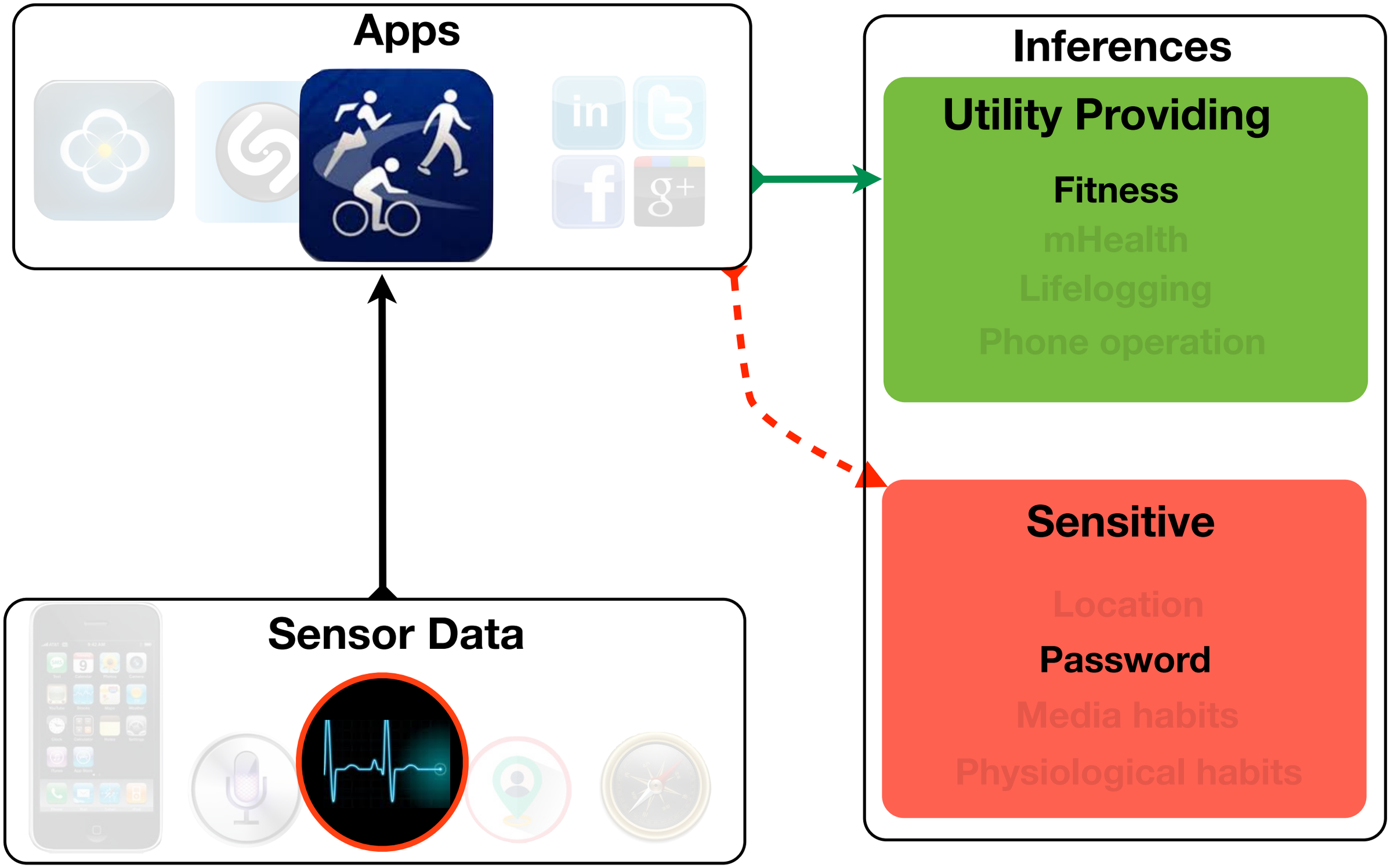
# From sensor data to inferences



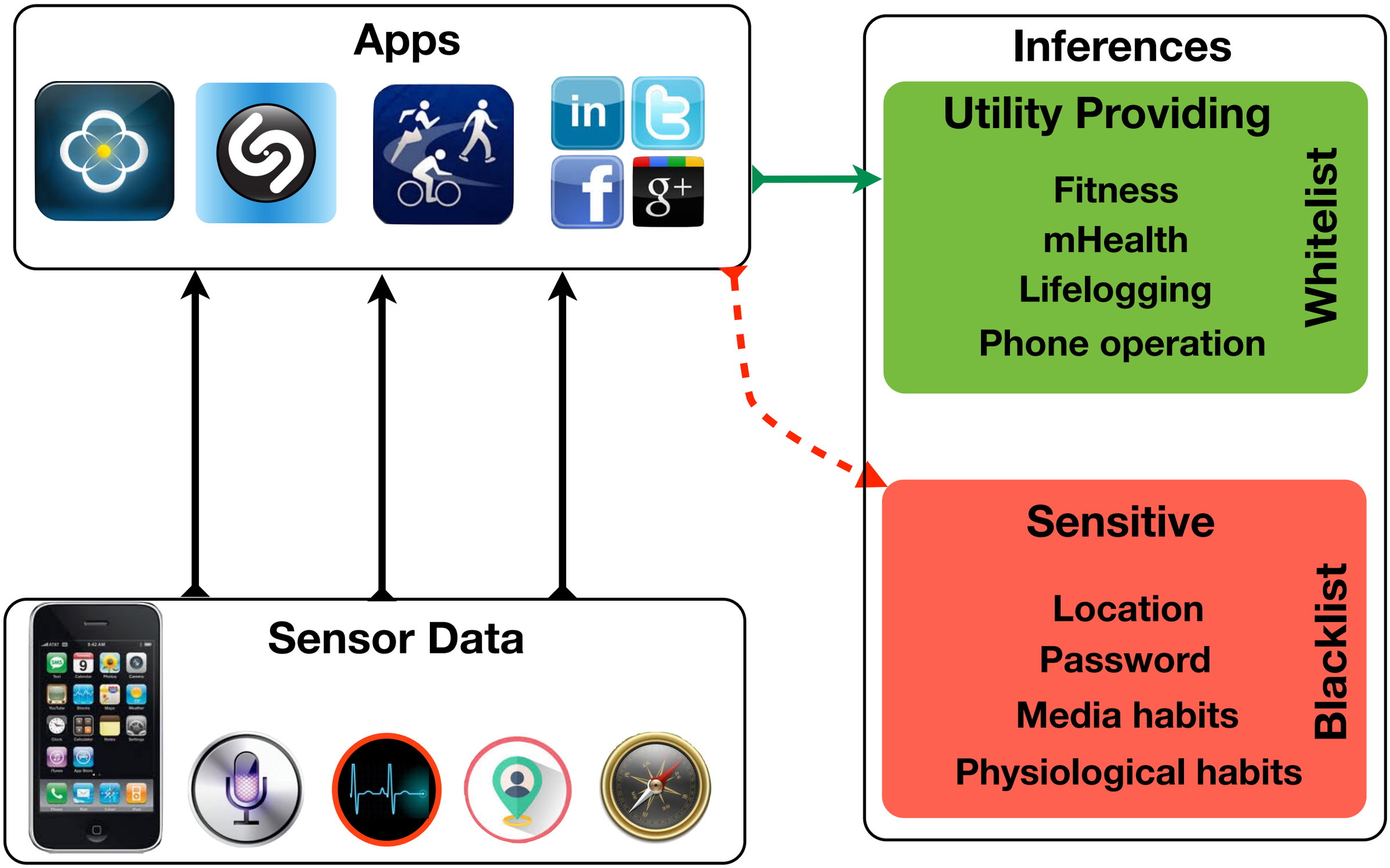
# From sensor data to inferences



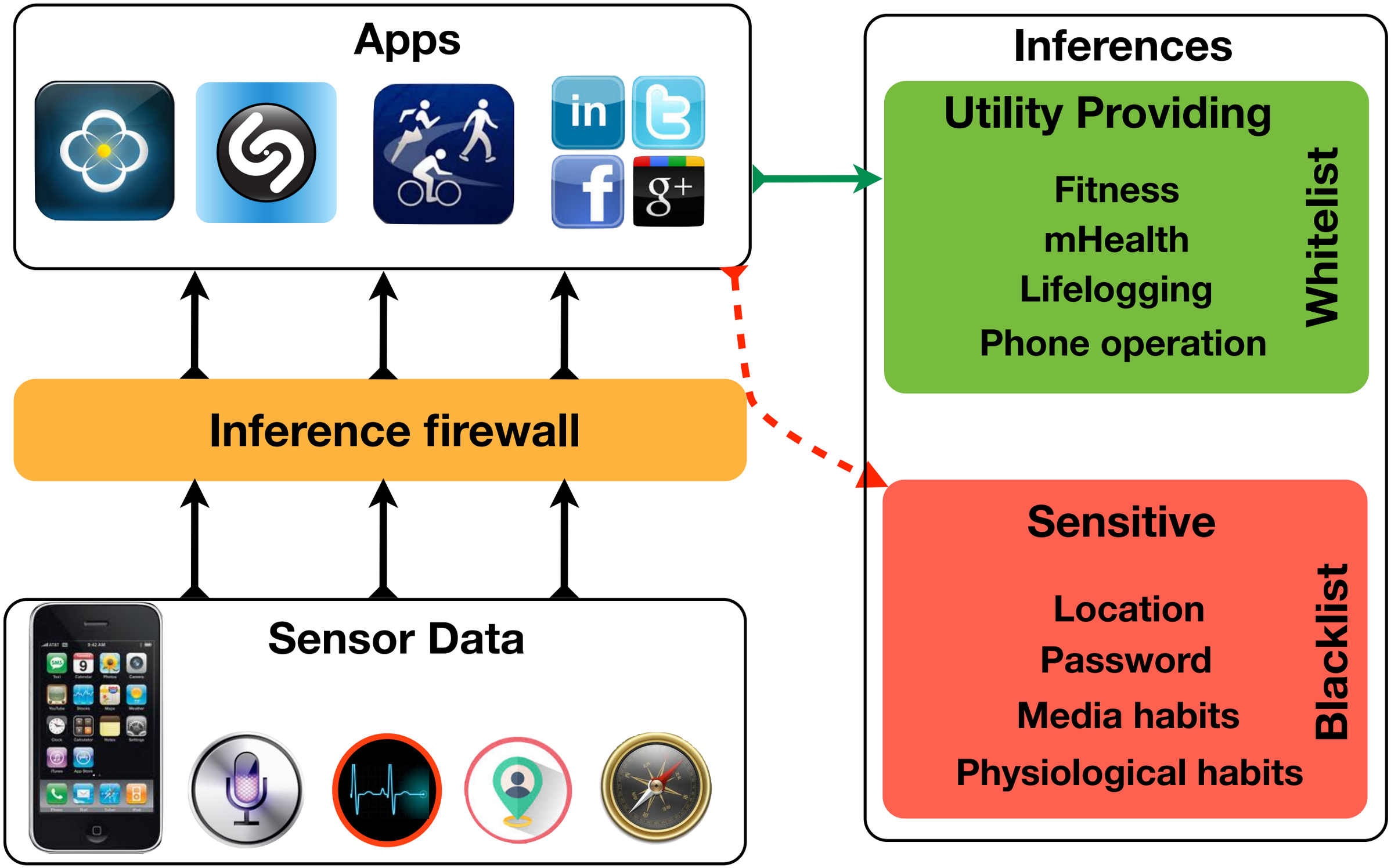
# From sensor data to inferences



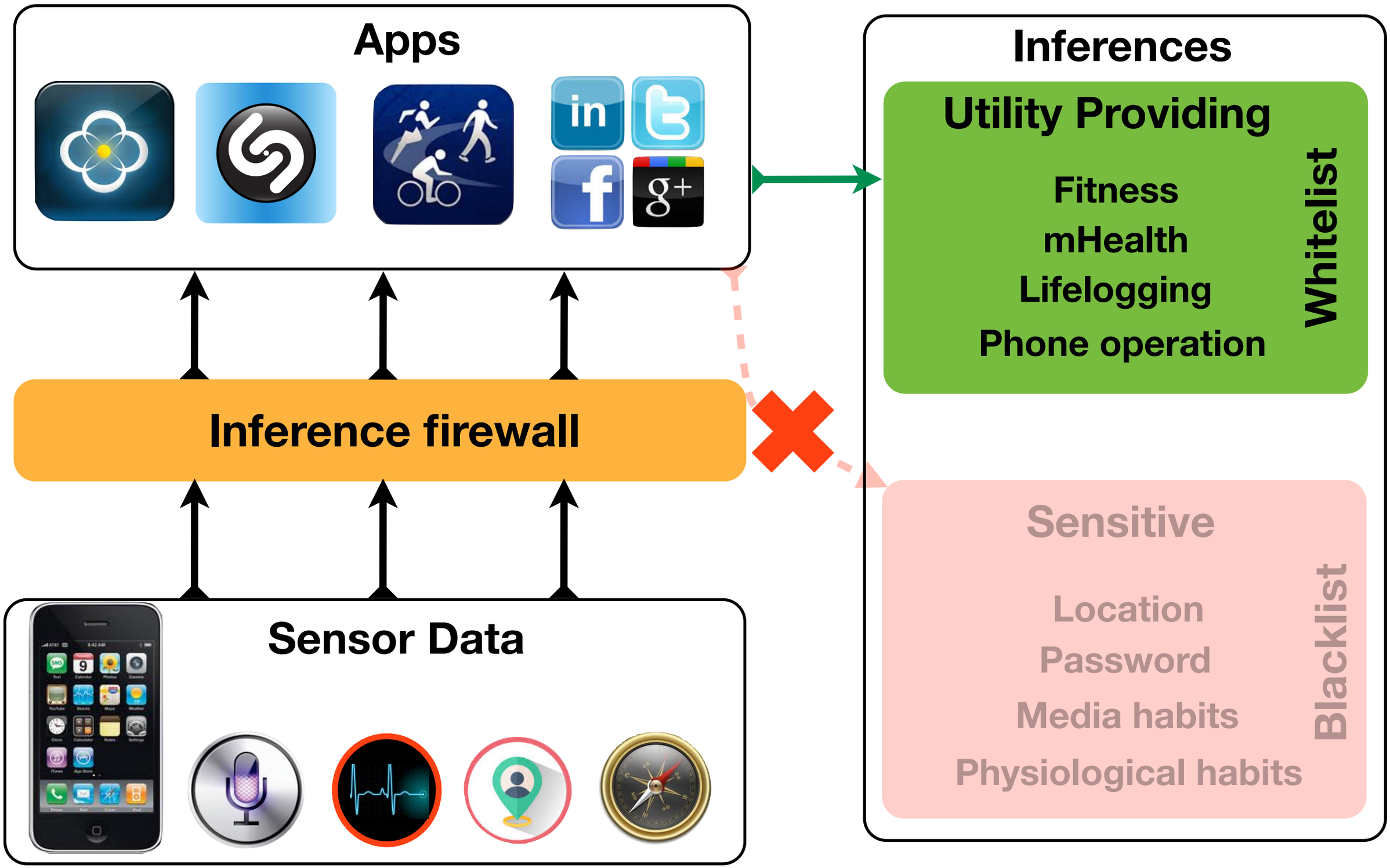
# Protecting inference privacy while providing utility



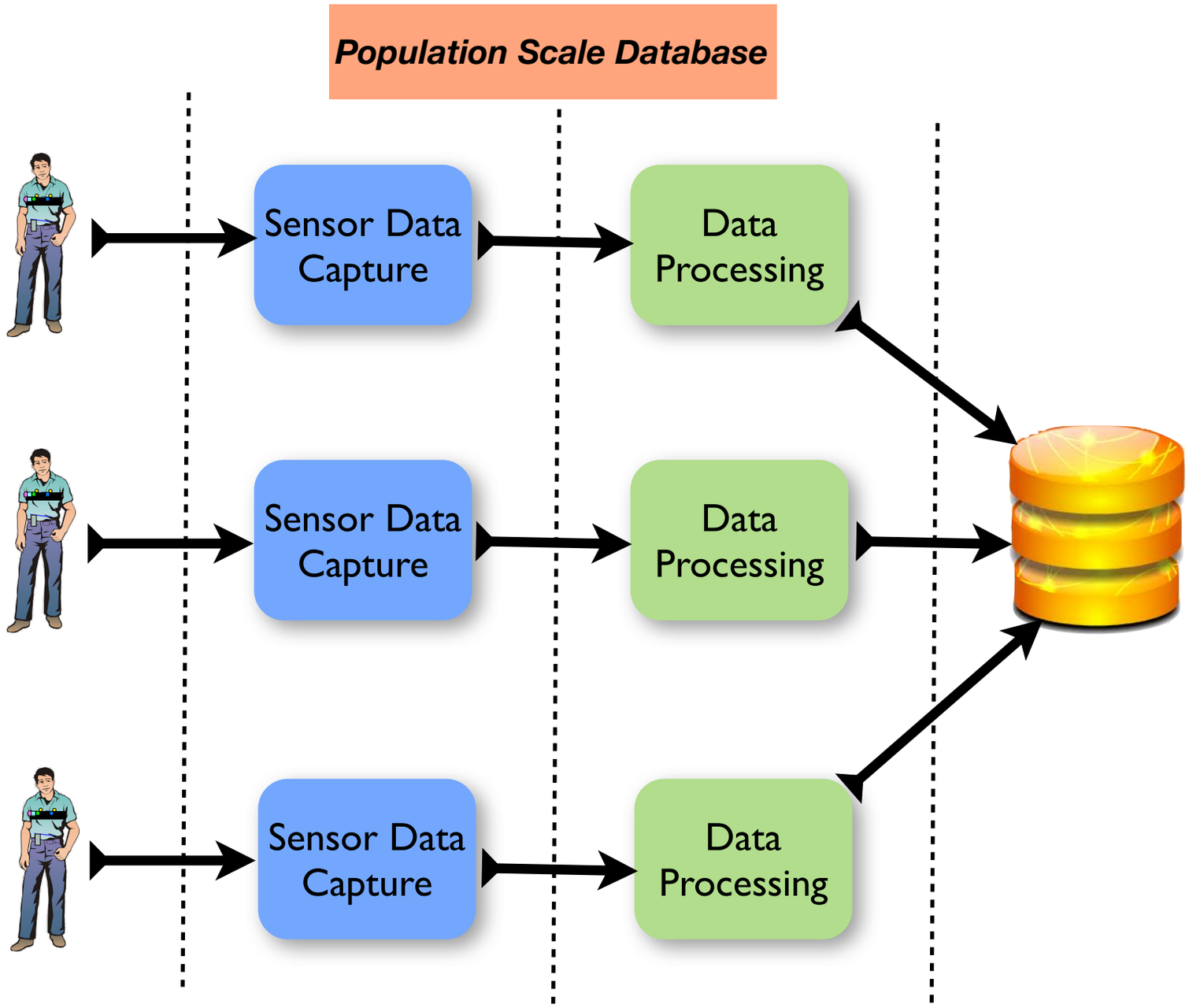
# Protecting inference privacy while providing utility



# Protecting inference privacy while providing utility

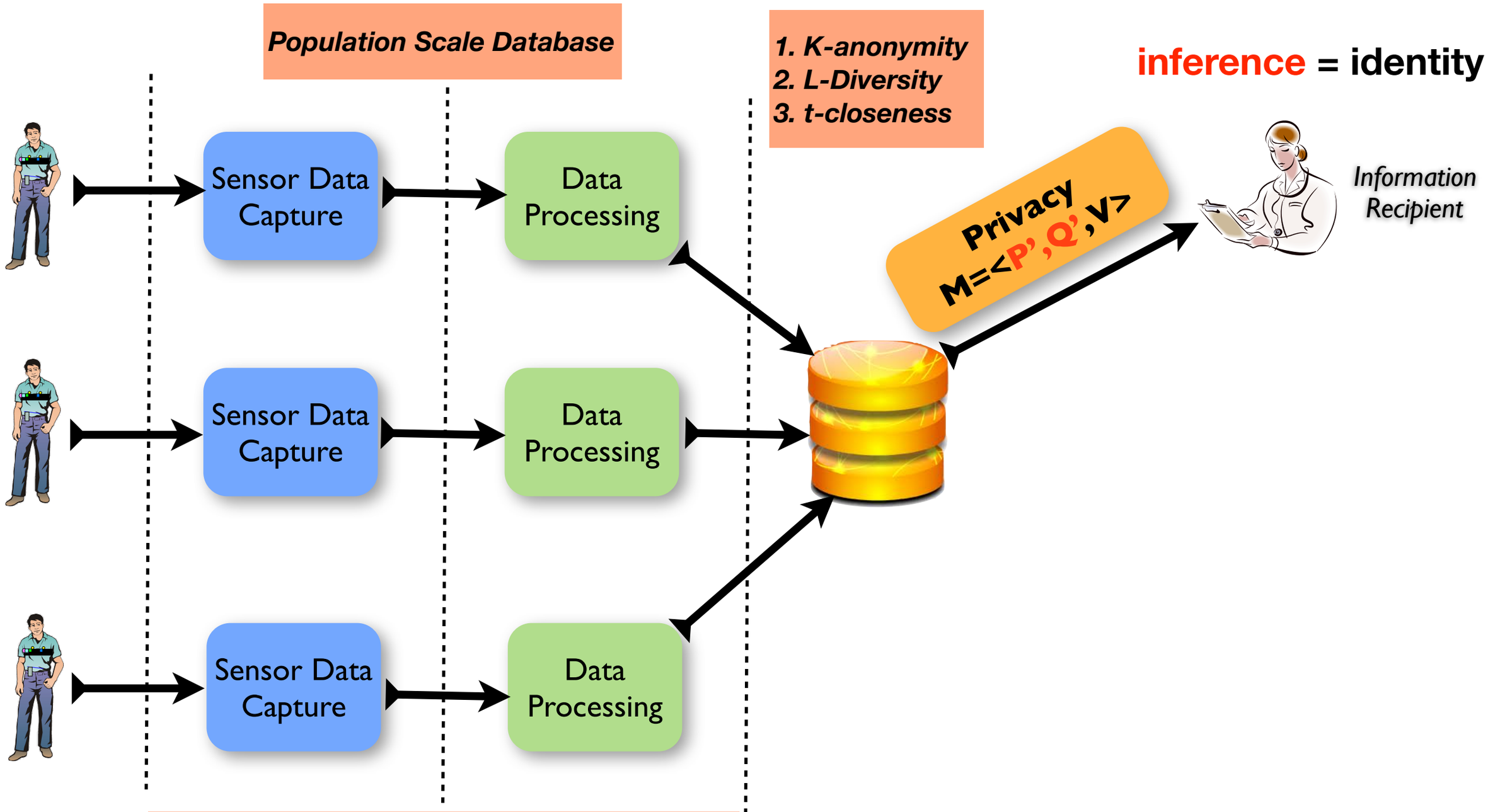


# Prior notions of privacy in databases



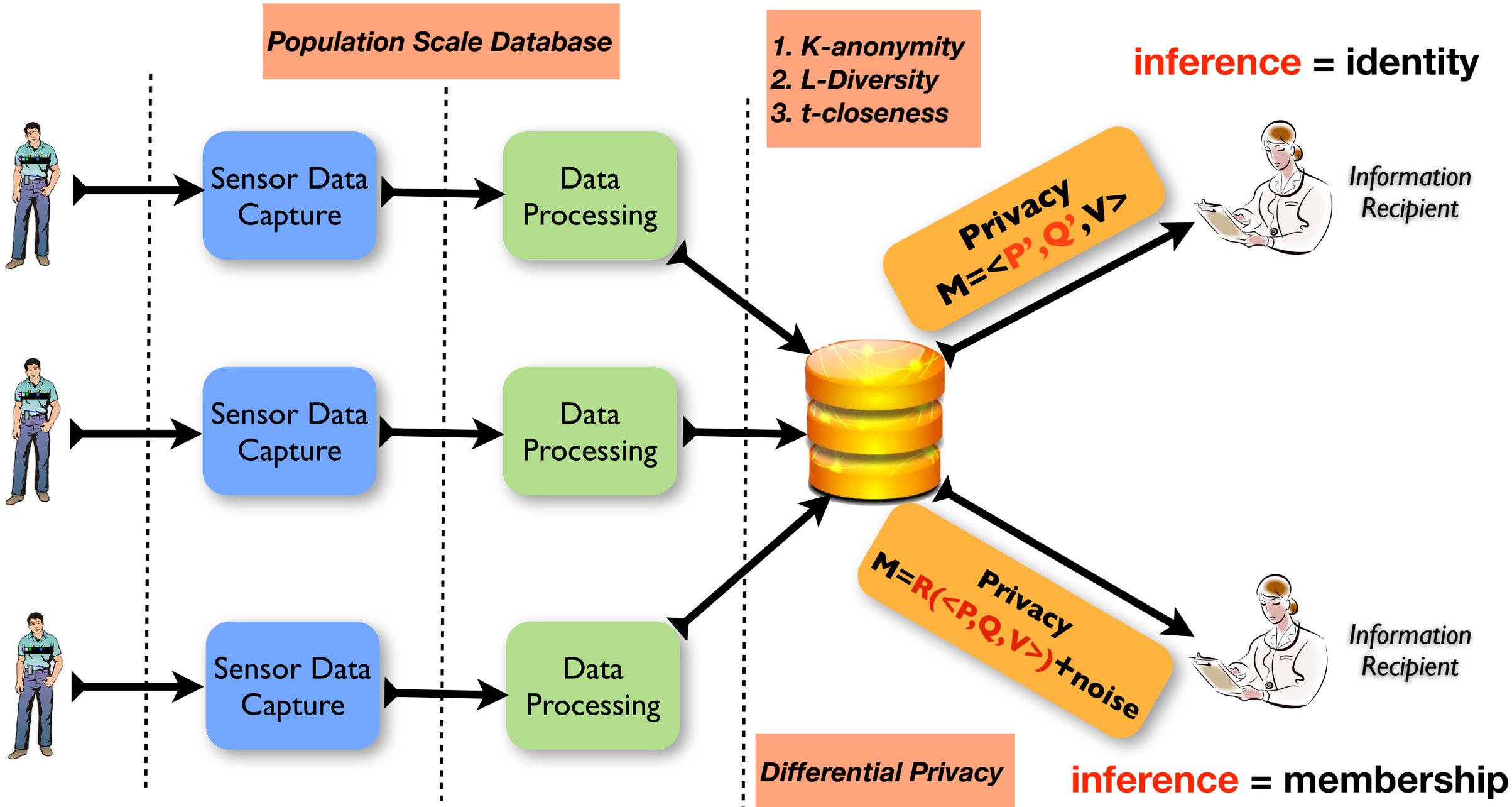

D := {
**P** = **personal identifiers**, (name, ID)
**Q** = **quasi identifiers**, (age, zip code)
**V** = **measurement values** (sensor data)}

# Prior notions of privacy in databases



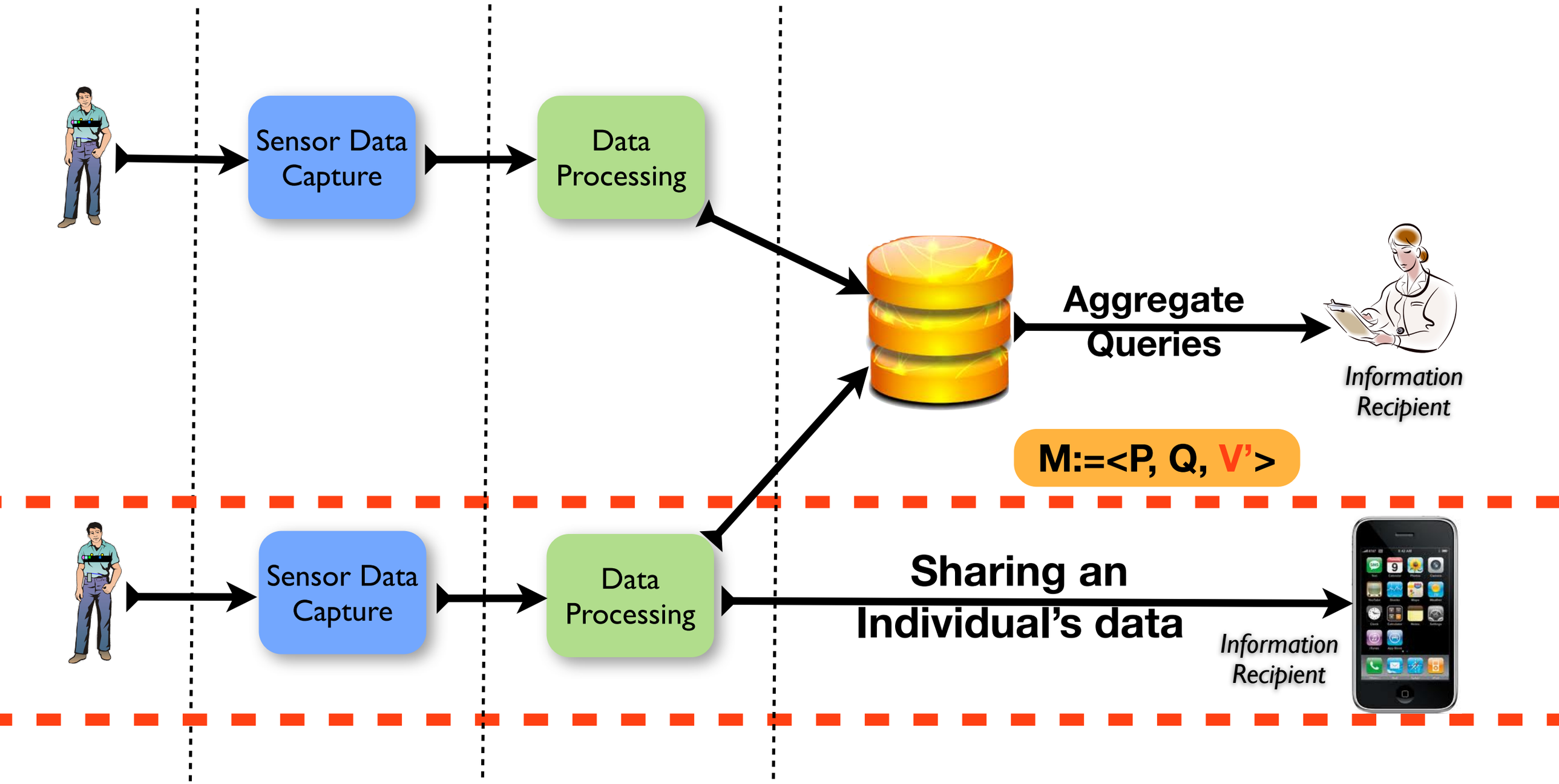

D := {
**P** = **personal identifiers**, (name, ID)
**Q** = **quasi identifiers**, (age, zip code)
**V** = **measurement values** (sensor data)}

# Prior notions of privacy in databases

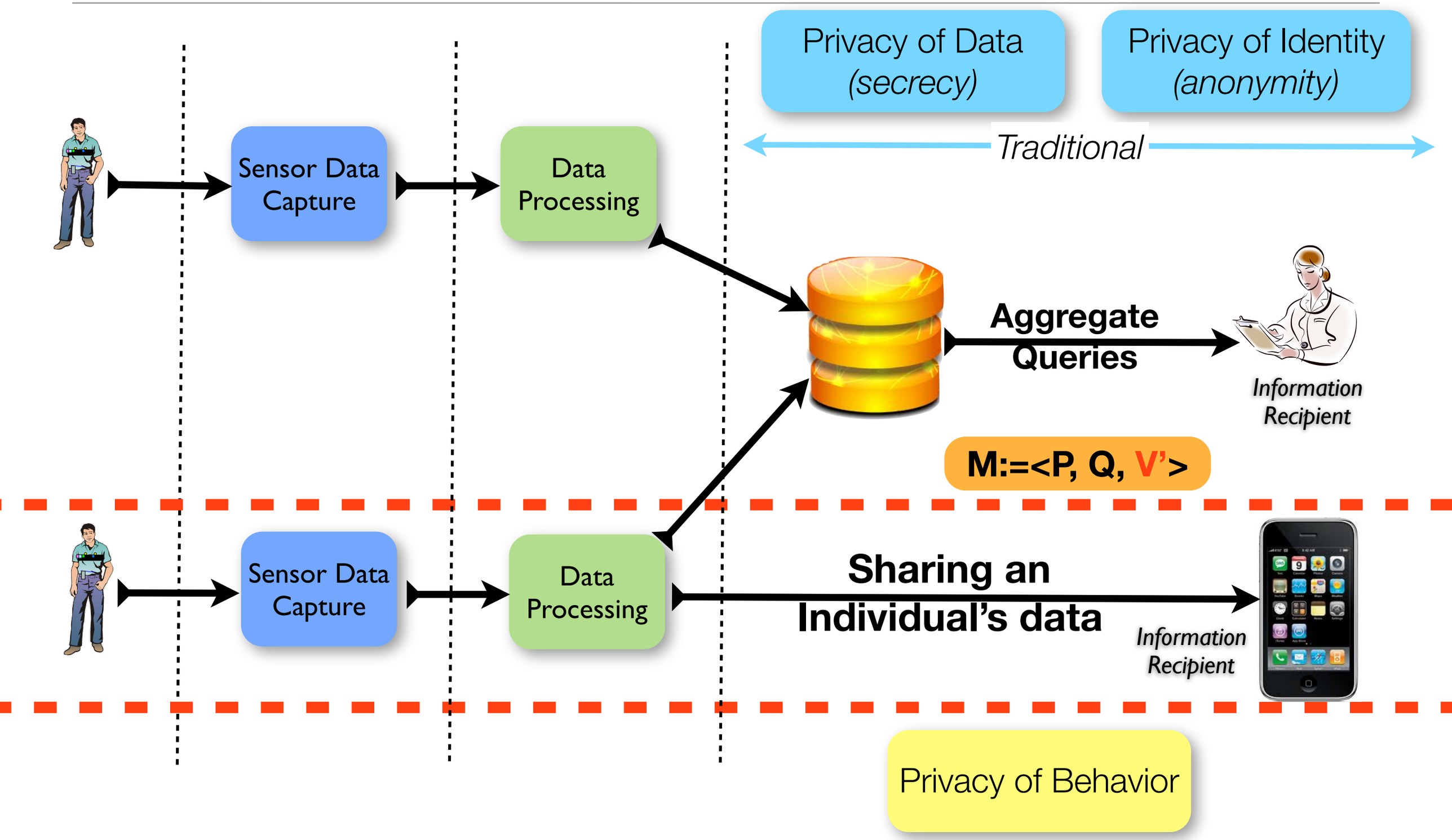



D := {
**P** = **personal identifiers**, (name, ID)
**Q** = **quasi identifiers**, (age, zip code)
**V** = **measurement values** (sensor data)}

# Prior notions of privacy in databases



# Prior notions of privacy in databases



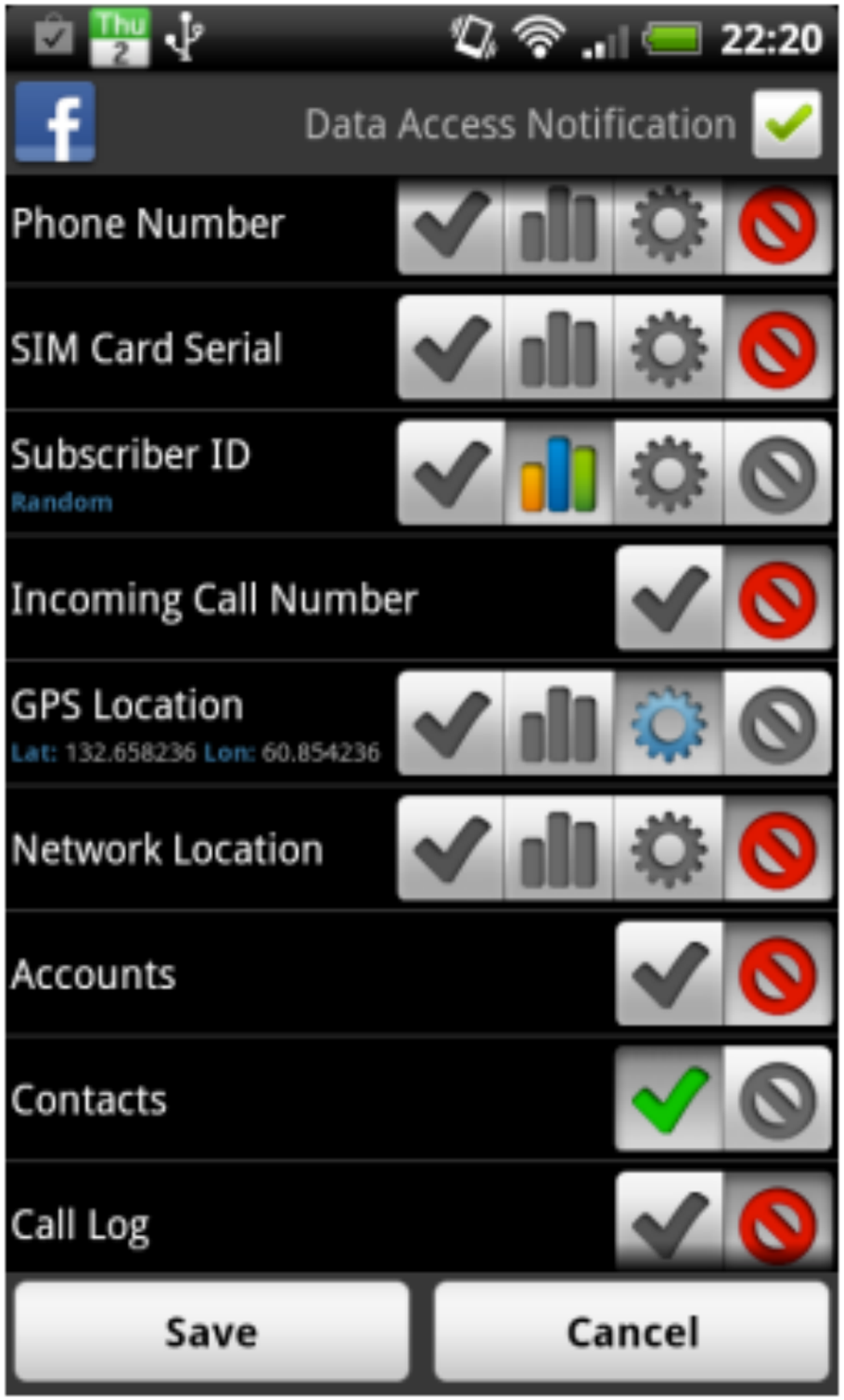
# Controls provided by current systems are insufficient

Android Manifest


Binary Policies

# Controls provided by current systems are insufficient

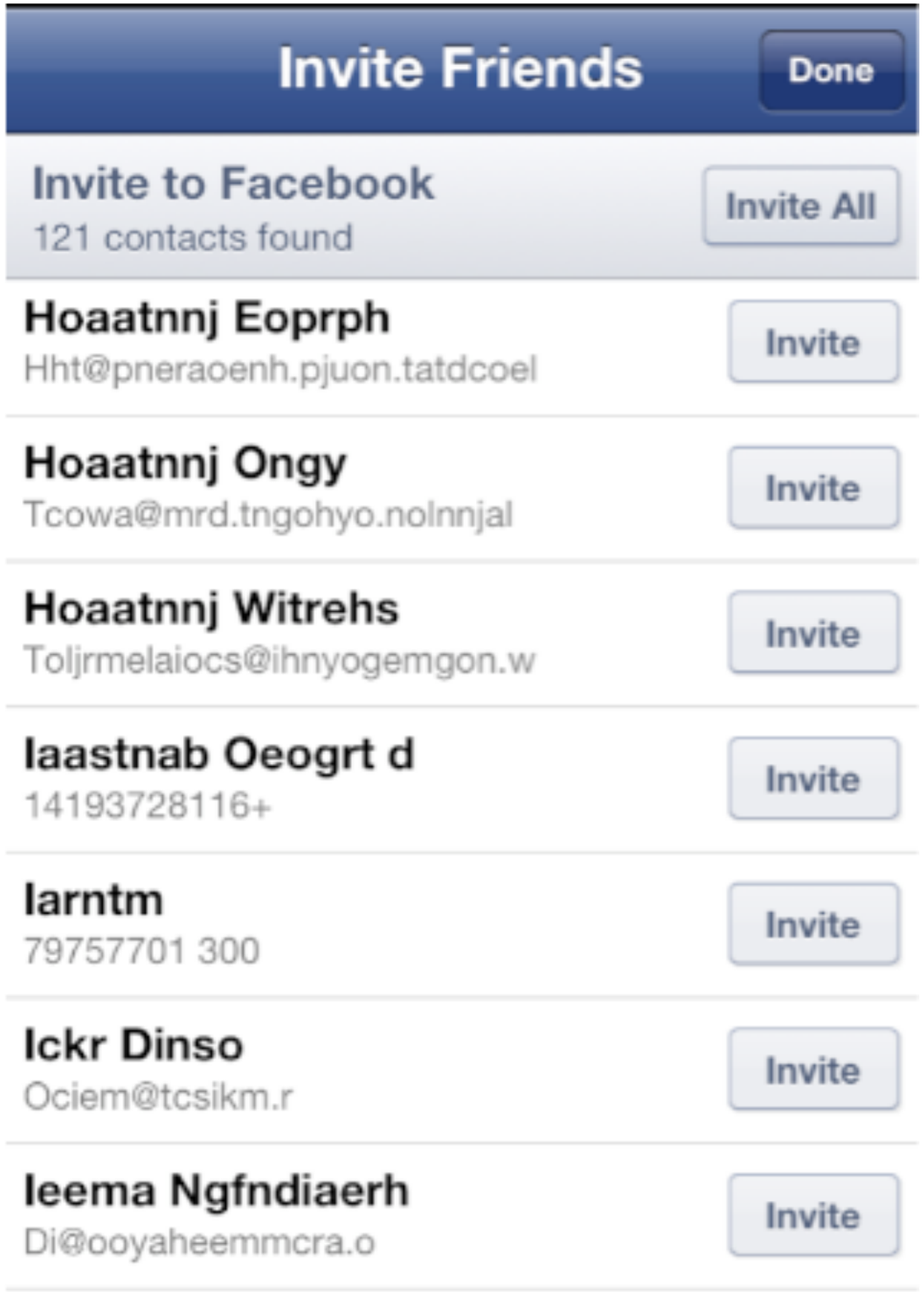
pDroid




Static Policies

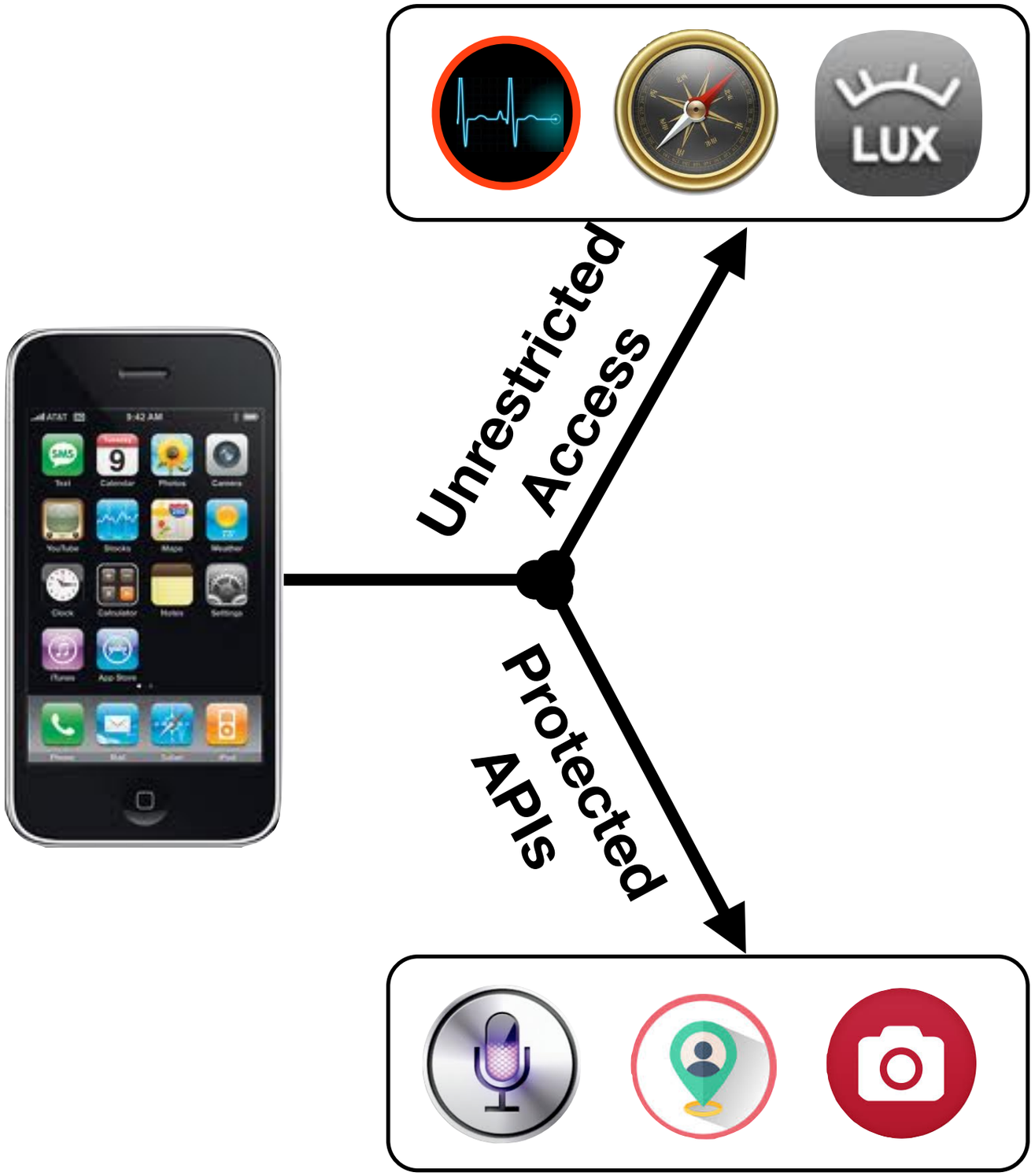
# Controls provided by current systems are insufficient

ProtectMyPrivacy

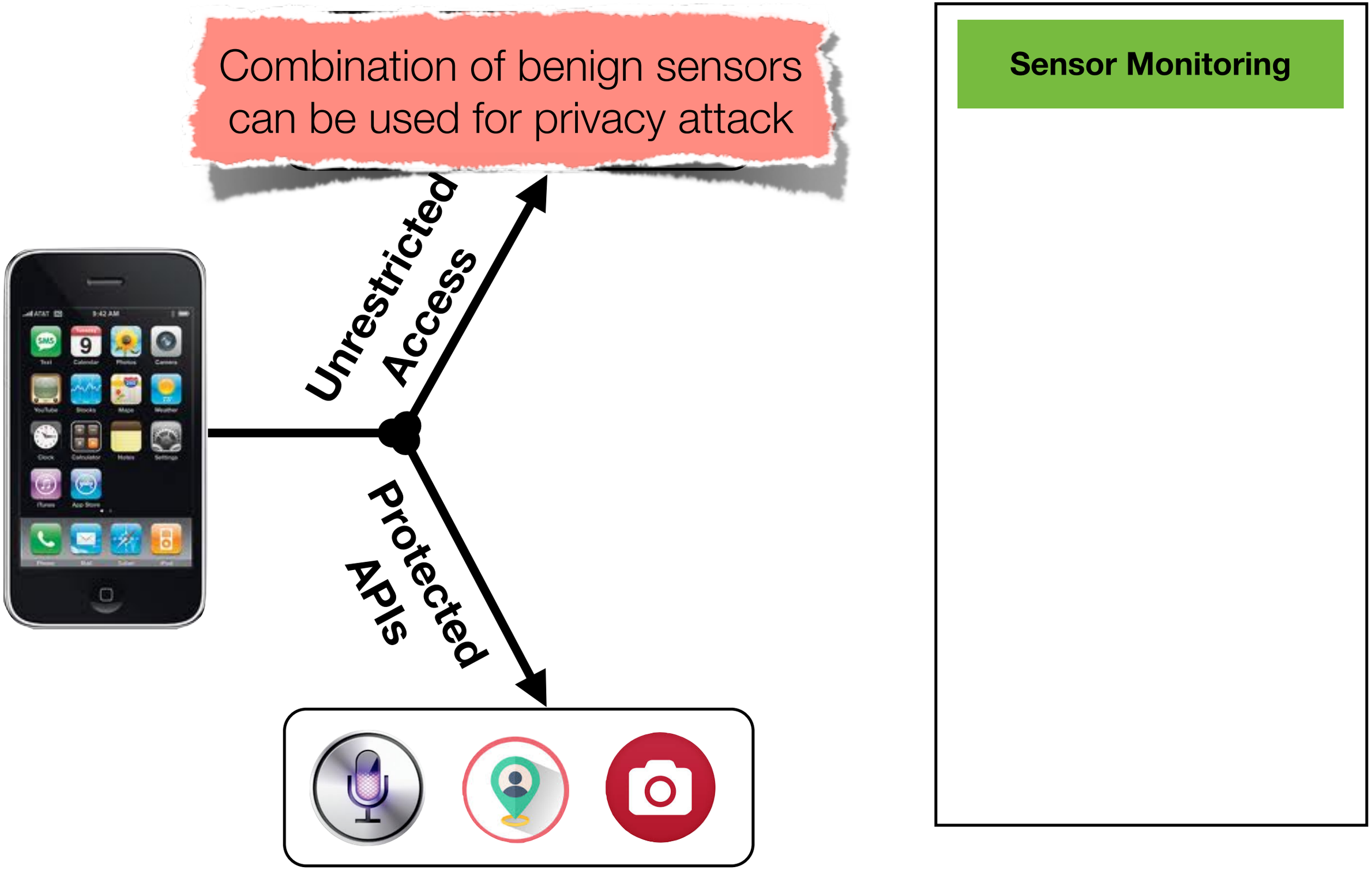



Share Random Data

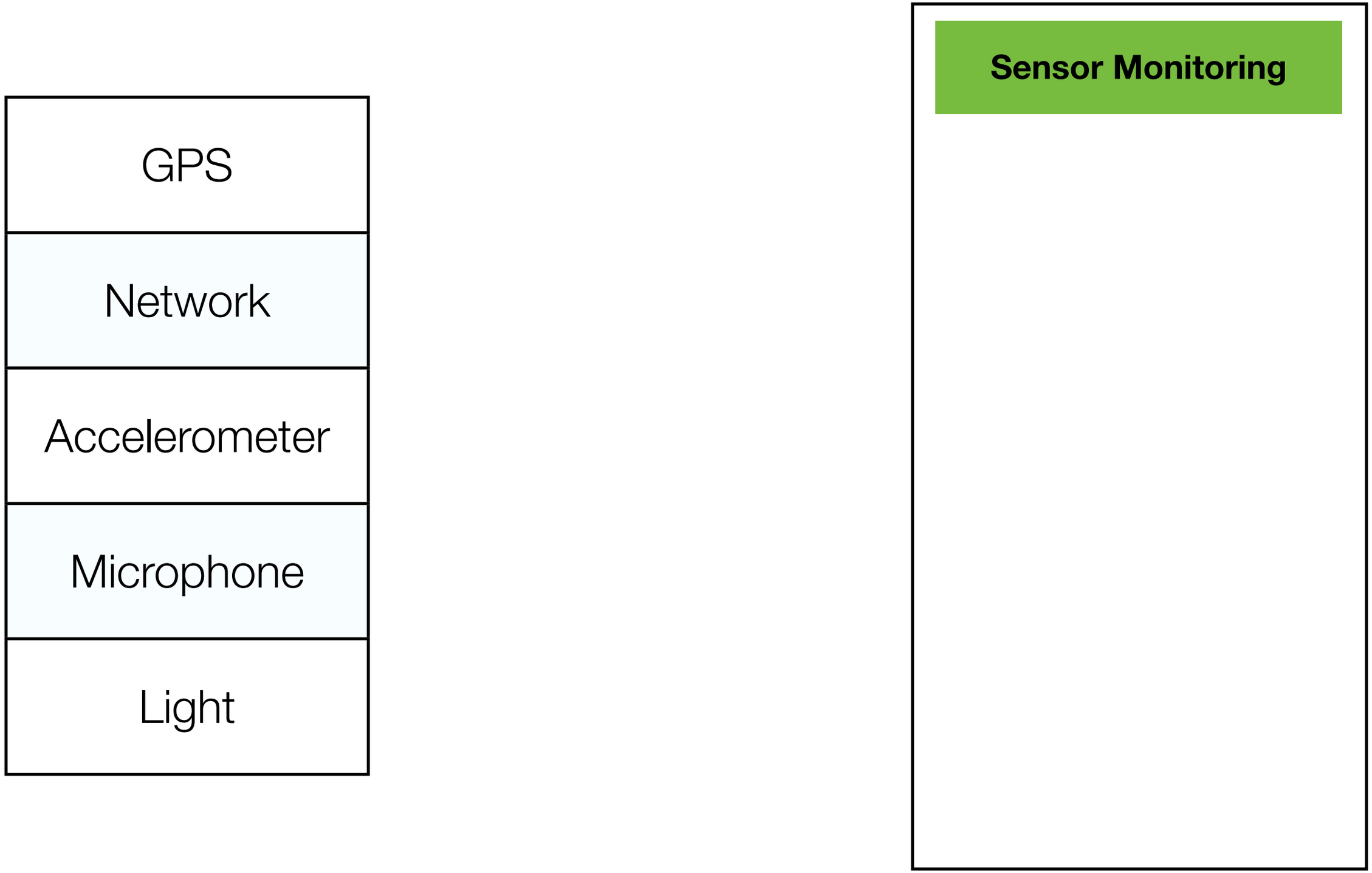
# Design requirements of ipShield



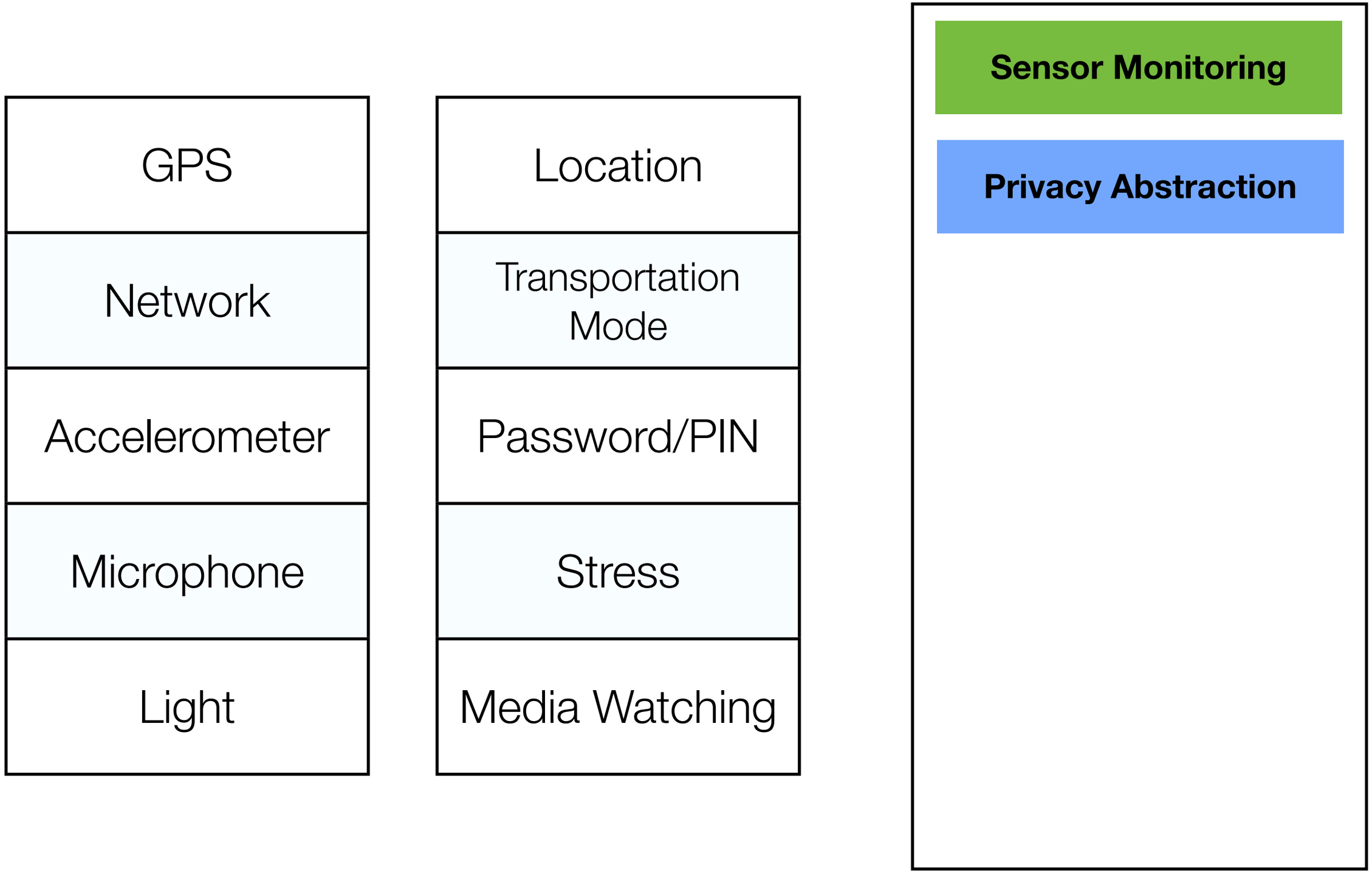
# Design requirements of ipShield



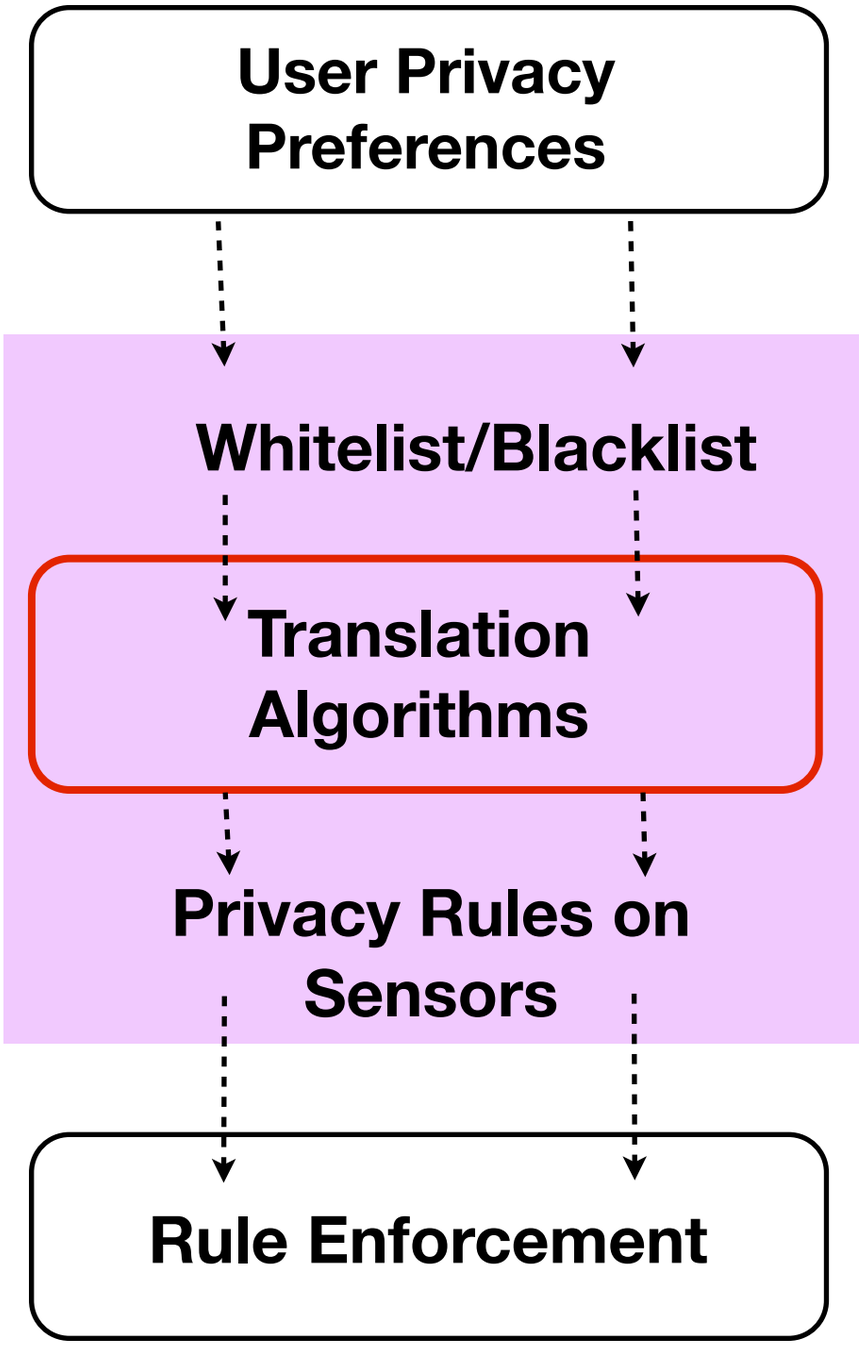
# Design requirements of ipShield

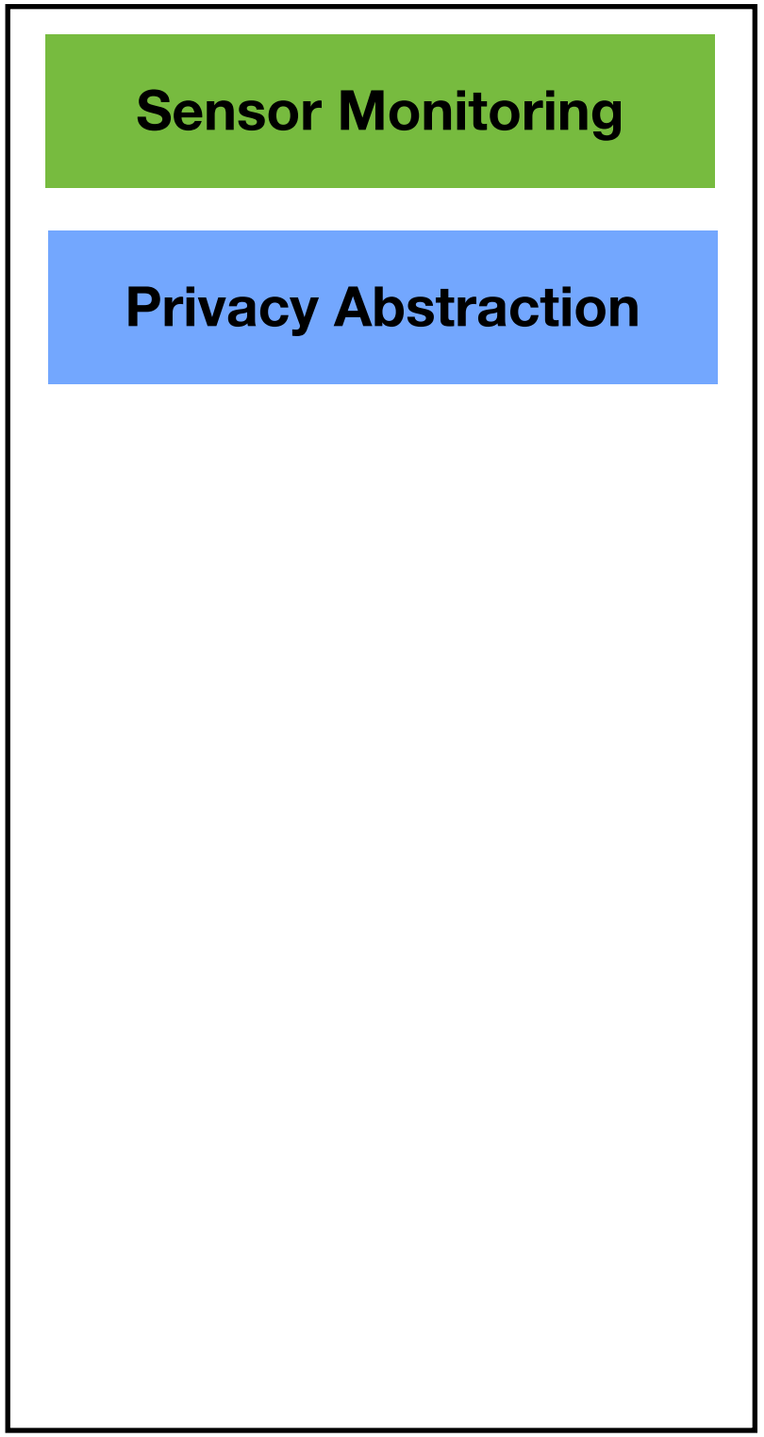


# Design requirements of ipShield

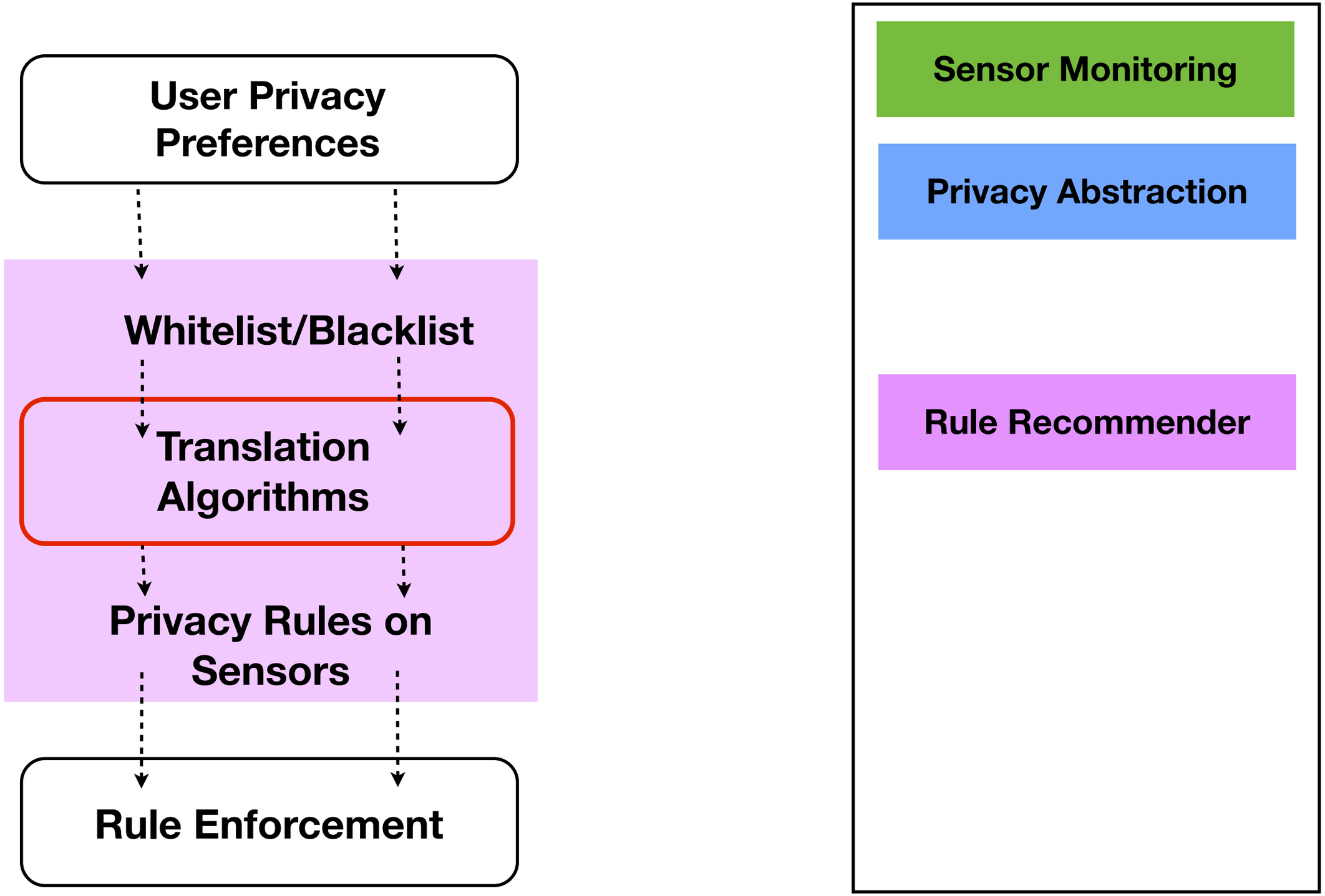


# Design requirements of ipShield

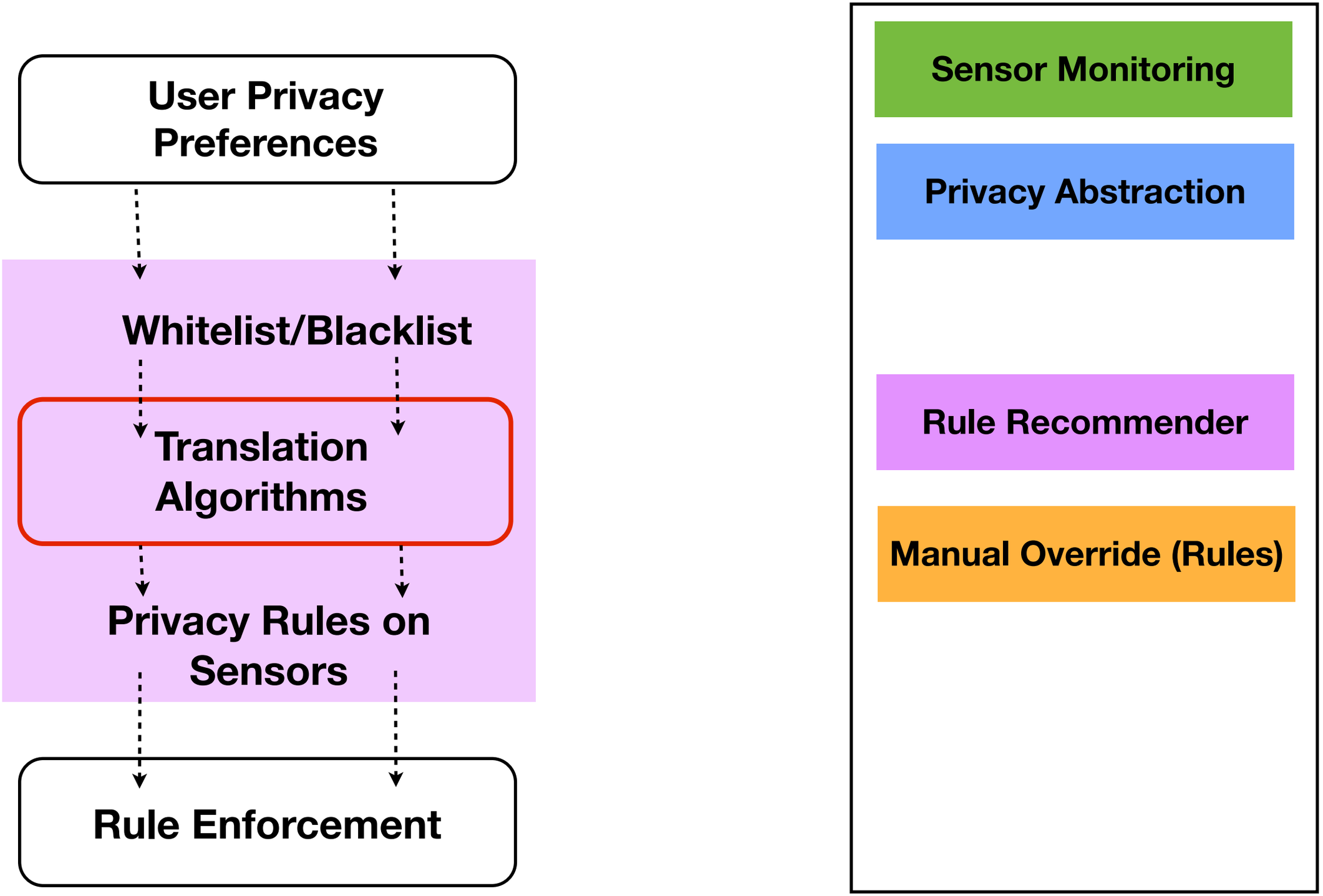


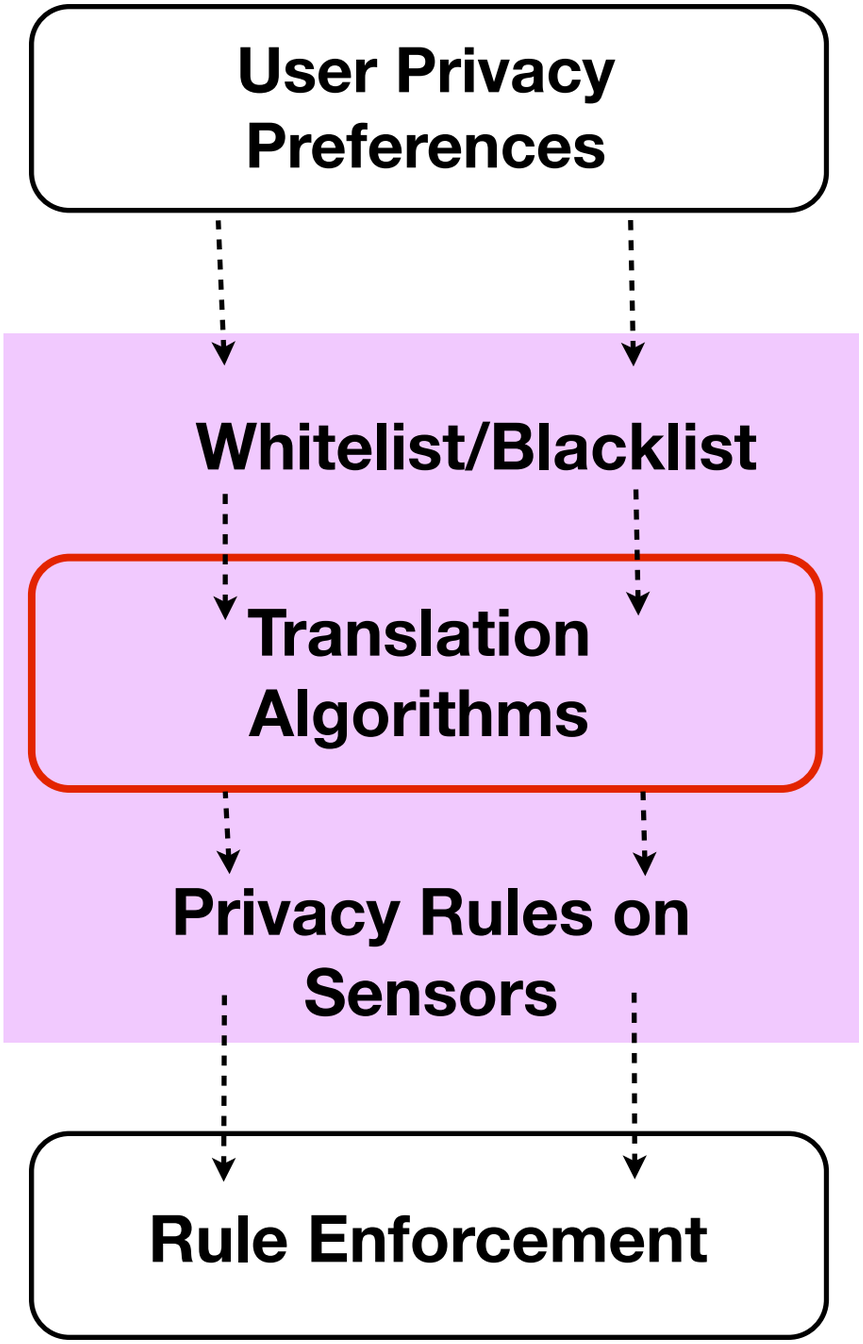
# Design requirements of ipShield

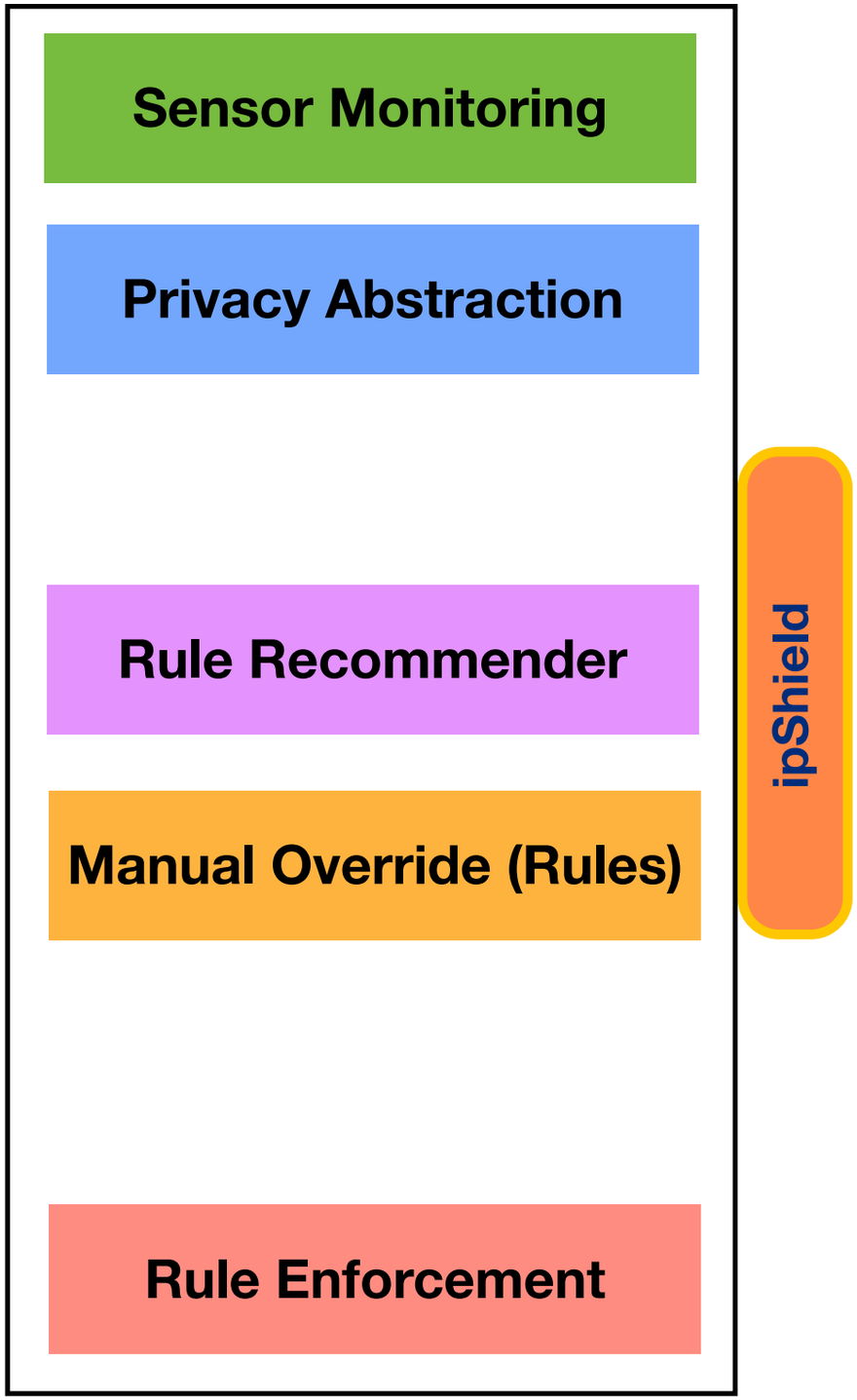


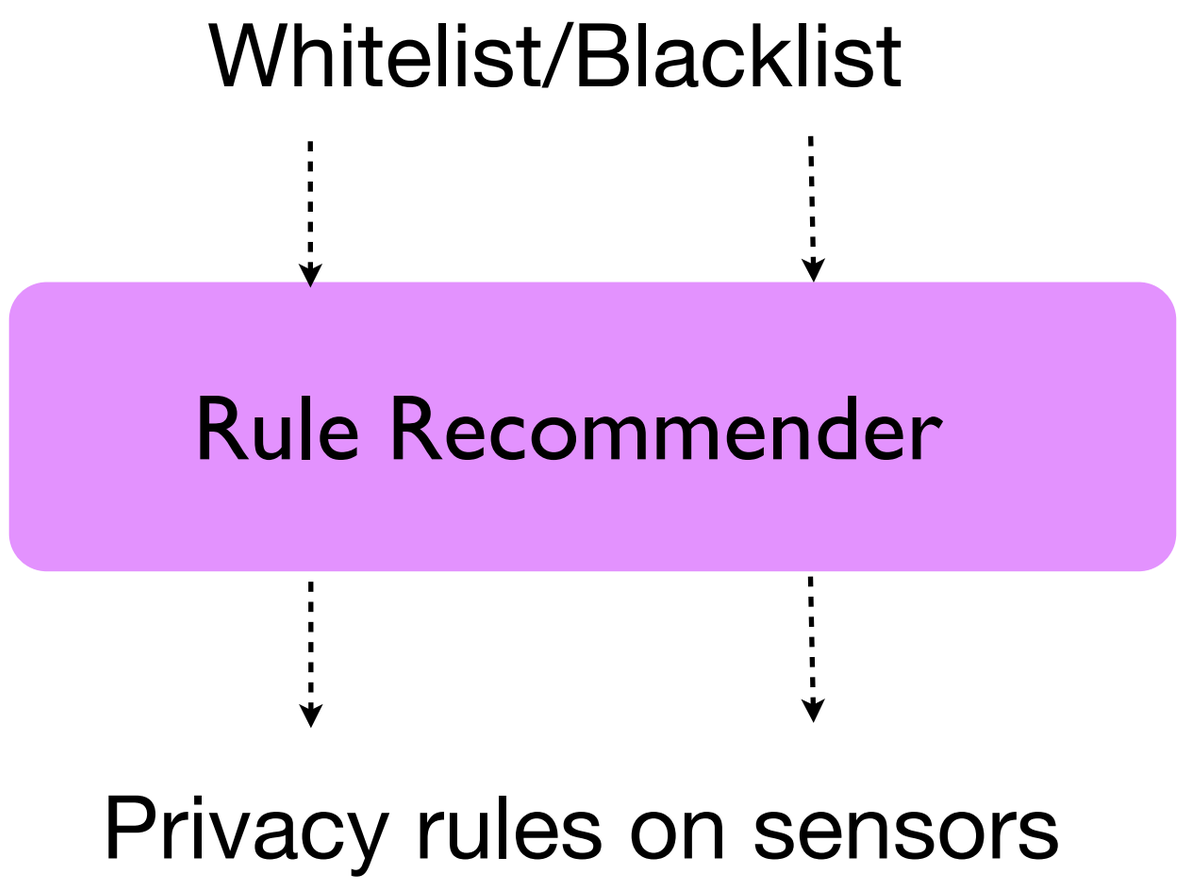
# Design requirements of ipShield



# Design requirements of ipShield




Whitelist/Blacklist
Rule Recommender
Privacy rules on sensors

# Recommender objective

Generate a plan for

**context-aware obfuscation of sensor data**

depending on the

**prioritized whitelist and blacklist**

such that

**accuracy of whitelist is maximized and**
**accuracy of blacklist is minimized.**

# Divide-and-conquer strategy

Recommend a plan containing

**allow/deny rules for sensors**

depending on the

**prioritized whitelist and blacklist**

such that

**accuracy of whitelist is maximized and accuracy of blacklist is minimized.**

# Divide-and-conquer strategy

Recommend a plan containing
**allow/deny rules for sensors**
depending on the
**prioritized whitelist and blacklist**
such that
**accuracy of whitelist is maximized and accuracy of blacklist is minimized.**

+

Support manual override/configuration of
**fine-grained context-aware rules**

# Elements of the problem: accuracy

# Elements of the problem: accuracy

| Inference Database (A) | Activity | Location | OnScreen Taps |
|---|---|---|---|
| GPS+Acc+Gyro | 95% | 97% | 80% |
| GPS+WiFi | 83.1% | 97% | 0% |
| GPS+GSM | 81.7% | 98.2% | 0% |
| GSM+WiFi | 72.9% | 94.03% | 0% |

# Elements of the problem: accuracy

Inference Database (A)

| | Activity | Location | OnScreen Taps |
|---|---|---|---|
| GPS+Acc+Gyro | 95% | 97% | 80% |
| Sensor Combination | | 97% | 0% |
| GPS+GSM | 81.7% | 98.2% | 0% |
| GSM+WiFi | 72.9% | 94.03% | 0% |

# Elements of the problem: accuracy

Inference Database (A)

| | Activity | Inference Type | OnScreen Taps |
|---|---|---|---|
| GPS+Acc+Gyro | 95% | | 80% |
| Sensor Combination | | 97% | 0% |
| GPS+GSM | 81.7% | 98.2% | 0% |
| GSM+WiFi | 72.9% | 94.03% | 0% |

# Elements of the problem: accuracy

Inference Database (A)

| | Activity | Inference Type | OnScreen Taps |
|---|---|---|---|
| GPS+Acc+Gyro | 95% | | 80% |
| Sensor Combination | | Accuracy of Prediction | 0% |
| GPS+GSM | 81.7% | 98.2% | 0% |
| GSM+WiFi | 72.9% | 94.03% | 0% |

# Elements of the problem: priority

**Priority (p)**

$$Priority = (p_{activity}, p_{location}, p_{tap})$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

priority = {10, 4, 10}

# Elements of the problem: priority

**Priority (p)**

$$Priority = (p_{activity}, p_{location}, p_{tap})$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

priority = {10, 4, 10}

**Whitelisted inferences**

priority ↑ ⇒ allow whitelisted inferences

**Blacklisted inferences**

priority ↑ ⇒ block blacklisted inferences

# Rule recommender in ipShield

$$\max_{\Phi \in 2^N} \sum_{l \in \mathcal{W}} A(\Phi, l) 2^{p_l} - \sum_{l \in \mathcal{B}} A(\Phi, l) 2^{p_l}$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{B}, p_l = p_{max}} A(\Psi, l) = 0$$

$\mathcal{W}$ = whitelist, $\mathcal{B}$ = blacklist, $p_l$ = priority, and
$\Phi$ = Sensor combination

# Rule recommender in ipShield

$$\max_{\Phi \in 2^N} \sum_{l \in \mathcal{W}} A(\Phi, l) 2^{p_l} - \sum_{l \in \mathcal{B}} A(\Phi, l) 2^{p_l}$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{B}, p_l = p_{max}} A(\Psi, l) = 0$$

$\mathcal{W}$ = whitelist, $\mathcal{B}$ = blacklist, $p_l$ = priority, and $\Phi$ = Sensor combination

$$\Downarrow$$

**Over all sensor combinations**

# Rule recommender in ipShield

$$\max_{\Phi \in 2^N} \sum_{l \in \mathcal{W}} A(\Phi, l) 2^{p_l} - \sum_{l \in \mathcal{B}} A(\Phi, l) 2^{p_l}$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{B}, p_l = p_{max}} A(\Psi, l) = 0$$

$\mathcal{W}$ = whitelist, $\mathcal{B}$ = blacklist, $p_l$ = priority, and $\Phi$ = Sensor combination

$$\Downarrow$$

**Over all sensor combinations**

**maximize accuracy of prioritized whitelist and**

# Rule recommender in ipShield

$$\max_{\Phi \in 2^N} \sum_{l \in \mathcal{W}} A(\Phi, l) 2^{p_l} - \sum_{l \in \mathcal{B}} A(\Phi, l) 2^{p_l}$$

$$\text{s.t.} \sum_{l \in \mathcal{B}, p_l = p_{max}} A(\Psi, l) = 0$$

$\mathcal{W}$ = whitelist, $\mathcal{B}$ = blacklist, $p_l$ = priority, and $\Phi$ = Sensor combination

$$\Downarrow$$

**Over all sensor combinations**

**maximize accuracy of prioritized whitelist and**

**minimize accuracy of prioritized blacklist**

# Rule recommender in ipShield

$$\max_{\Phi \in 2^N} \sum_{l \in \mathcal{W}} A(\Phi, l) 2^{p_l} - \sum_{l \in \mathcal{B}} A(\Phi, l) 2^{p_l}$$

$$\text{s.t.} \sum_{l \in \mathcal{B}, p_l = p_{max}} A(\Psi, l) = 0$$

$\mathcal{W}$ = whitelist, $\mathcal{B}$ = blacklist, $p_l$ = priority, and
$\Phi$ = Sensor combination

$$\Downarrow$$

**Over all sensor combinations**

**maximize accuracy of prioritized whitelist and**

**minimize accuracy of prioritized blacklist**

**such that highest priority blacklists are always blocked.**

# Rule recommender at work

| | Activity | Location | OnScreen Taps |
|---|---|---|---|
| GPS+Acc+Gyro | 95% | 97% | 80% |
| GPS+WiFi | 83.1% | 97% | 0% |
| GPS+GSM | 81.7% | 98.2% | 0% |
| GSM+WiFi | 72.9% | 94.03% | 0% |

# Rule recommender at work

| | Activity | Location | OnScreen Taps |
|---|---|---|---|
| GPS+Acc+Gyro | 95% | 97% | 80% |
| GPS+WiFi | 83.1% | 97% | 0% |
| GPS+GSM | 81.7% | 98.2% | 0% |
| GSM+WiFi | 72.9% | 94.03% | 0% |

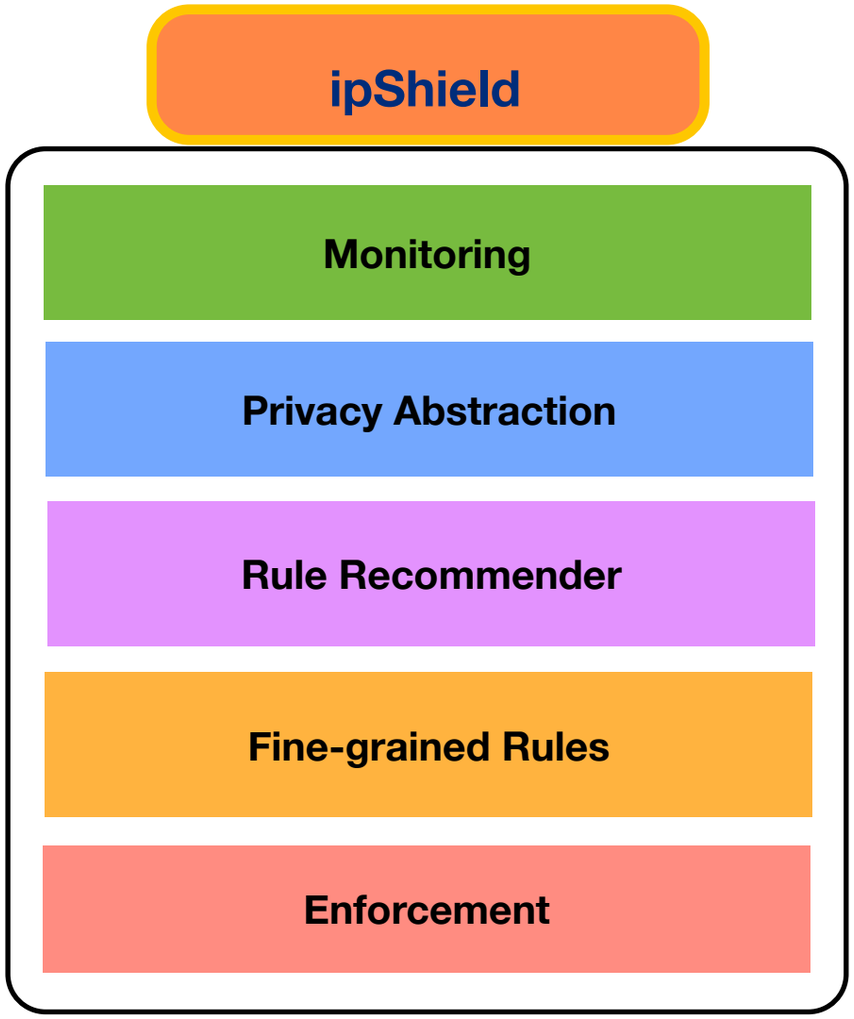
# Rule recommender at work

| | Activity | Location | OnScreen Taps | Priority1 {10, 4, 10} |
|---|---|---|---|---|
| **GPS+Acc+Gyro** | **95%** | **97%** | **80%** | 0 |
| **GPS+WiFi** | **83.1%** | **97%** | **0%** | **835.4** |
| **GPS+GSM** | **81.7%** | **98.2%** | **0%** | 820.0 |
| **GSM+WiFi** | **72.9%** | **94.03%** | **0%** | 731.45 |

# Rule recommender at work

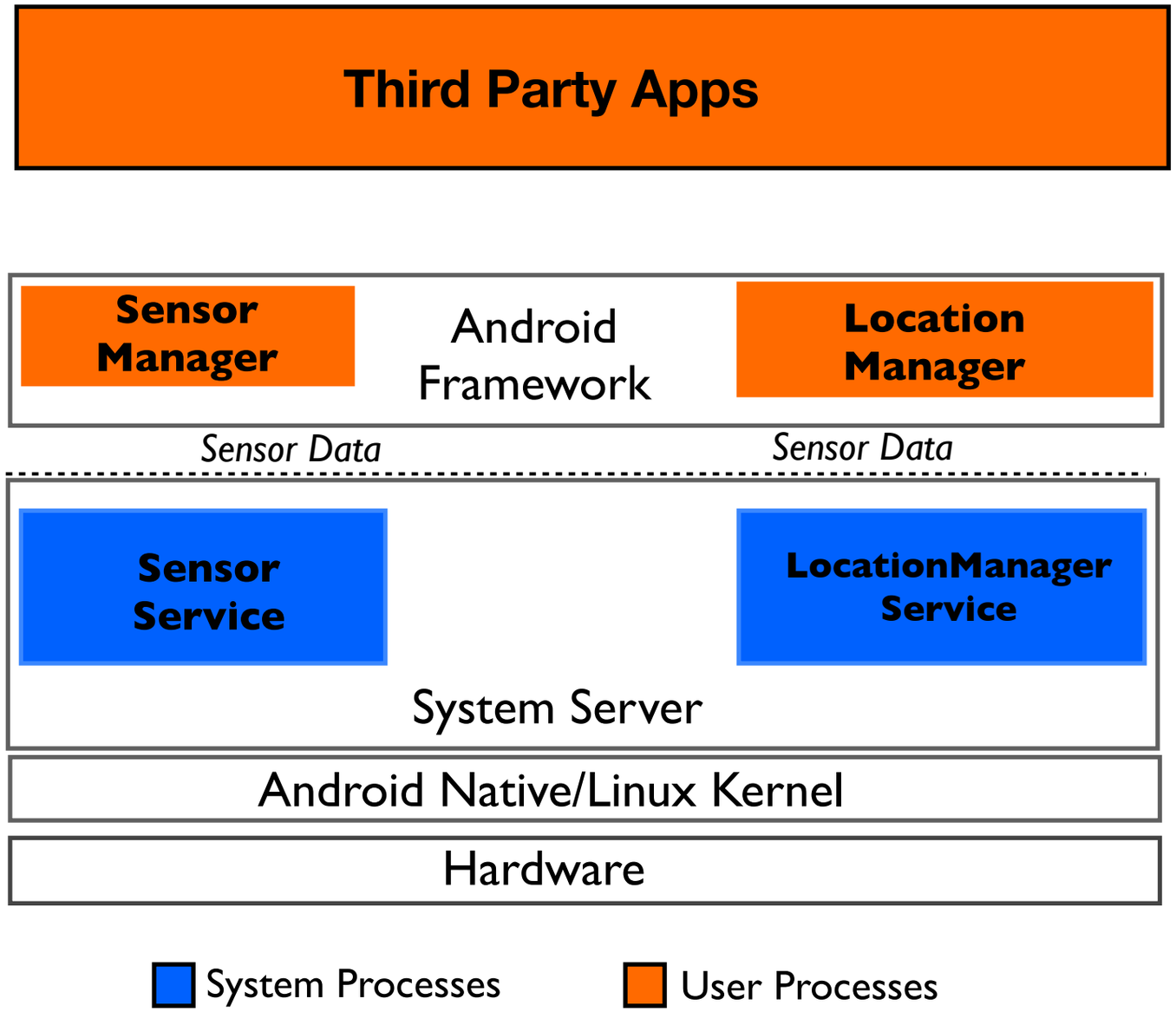
| | Activity | Location | OnScreen Taps | Priority1 {10, 4, 10} |
|---|---|---|---|---|
| GPS+Acc+Gyro | 95% | 97% | 80% | 0 |
| GPS+WiFi | 83.1% | 97% | 0% | 835.4 |
| GPS+GSM | 81.7% | 98.2% | 0% | 820.0 |
| GSM+WiFi | 72.9% | 94.03% | 0% | 731.45 |

**Allow**

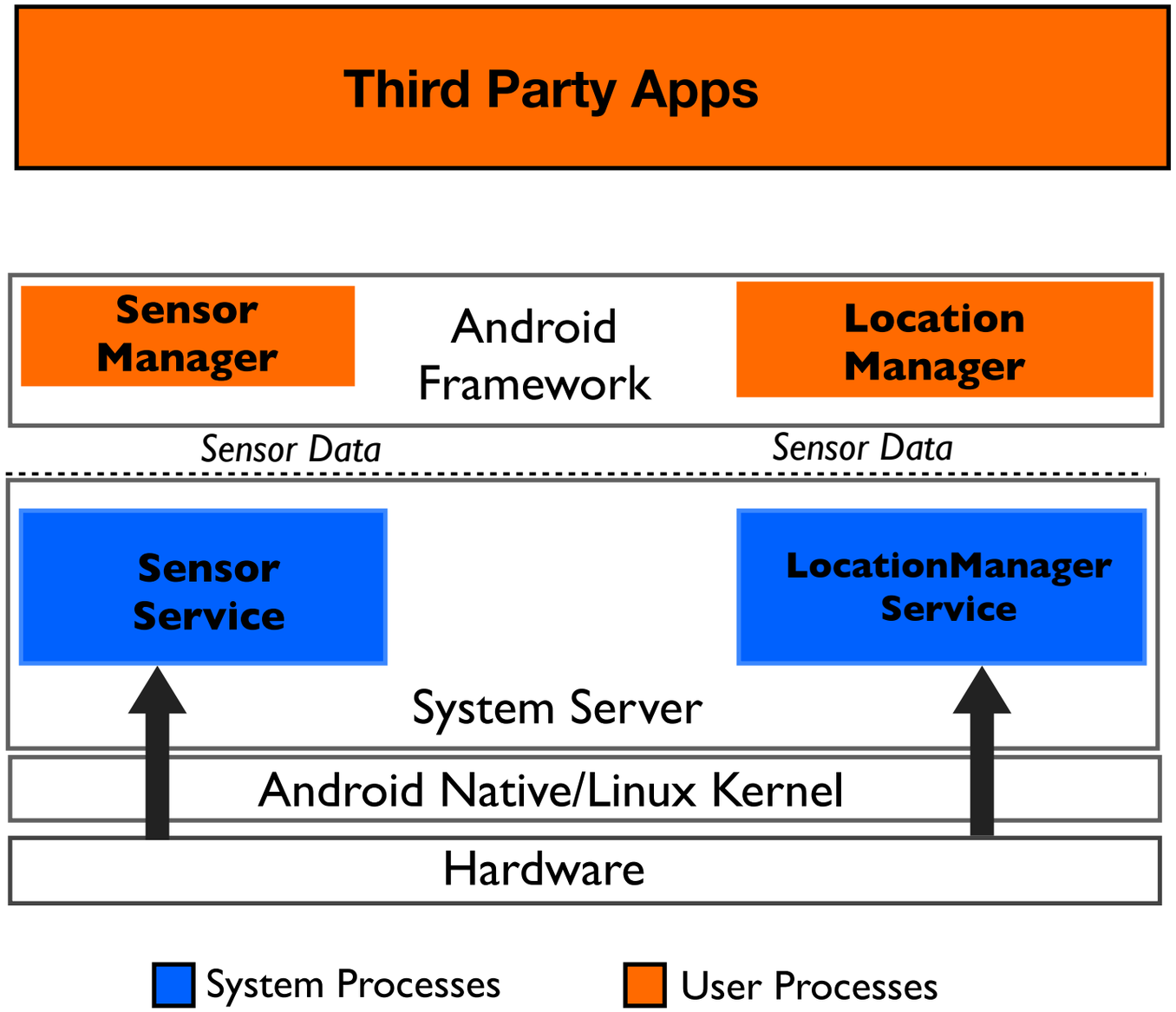



Prototype implementation on Android

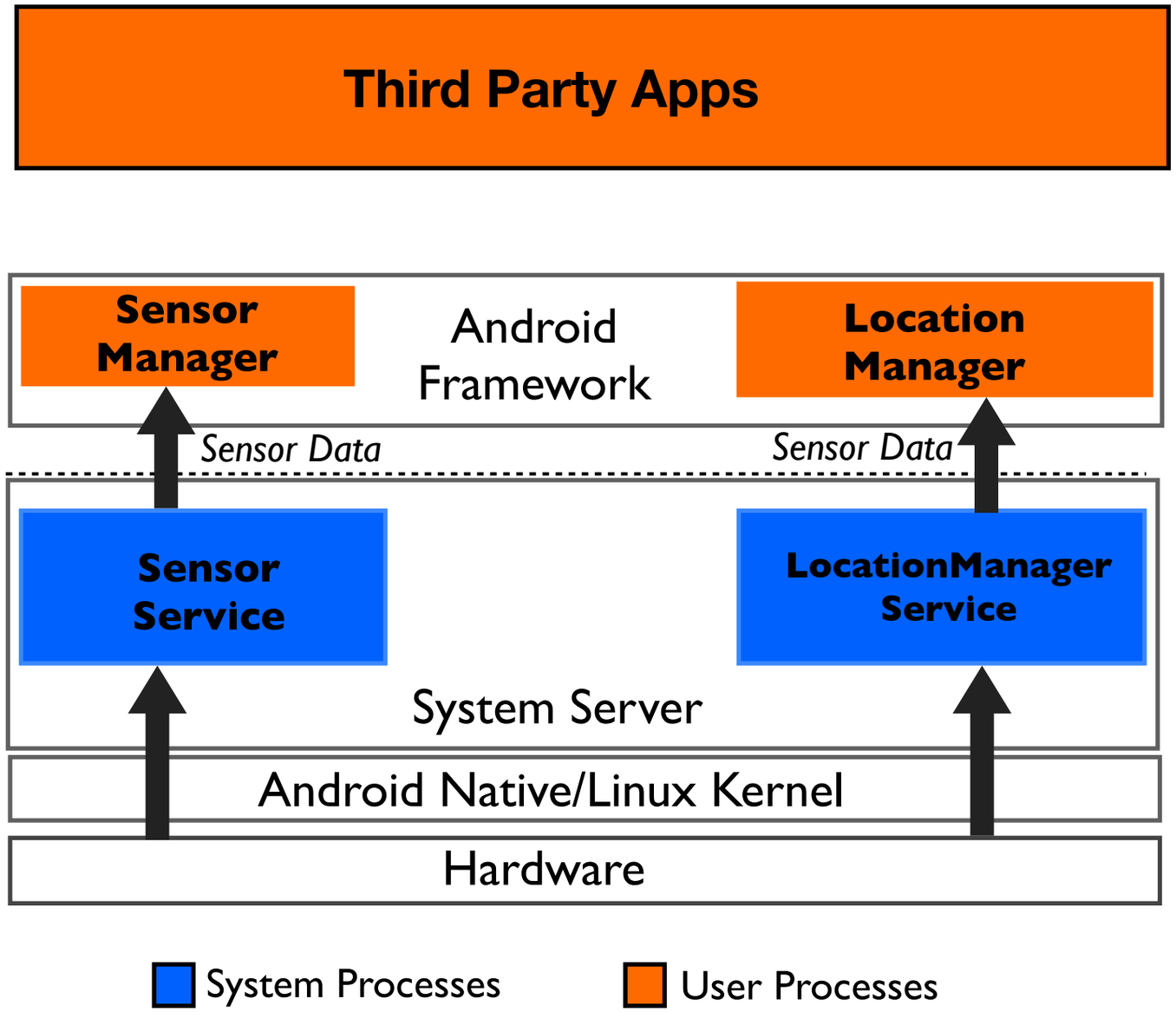
# Sensor subsystem in android and data interception



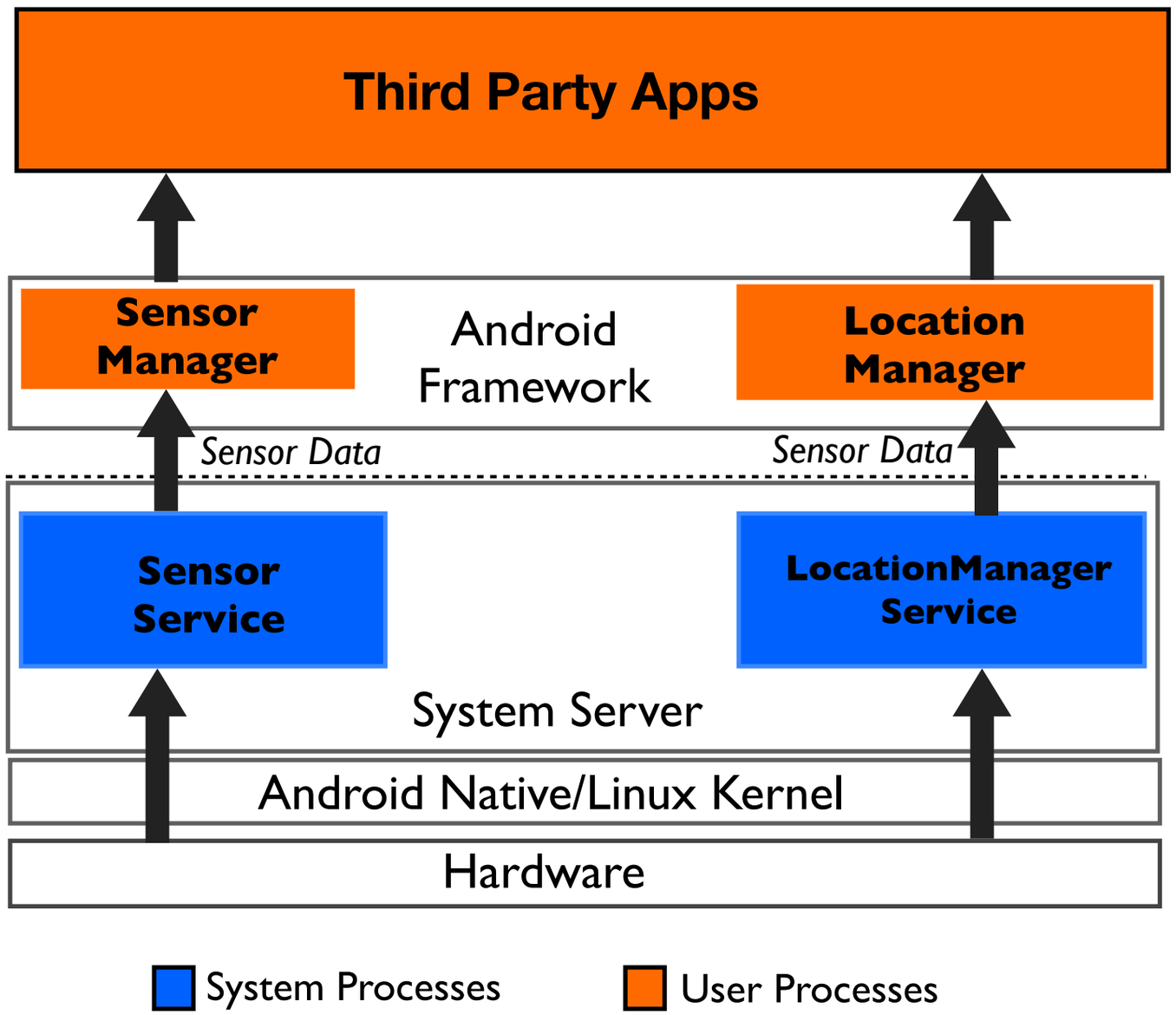
# Sensor subsystem in android and data interception



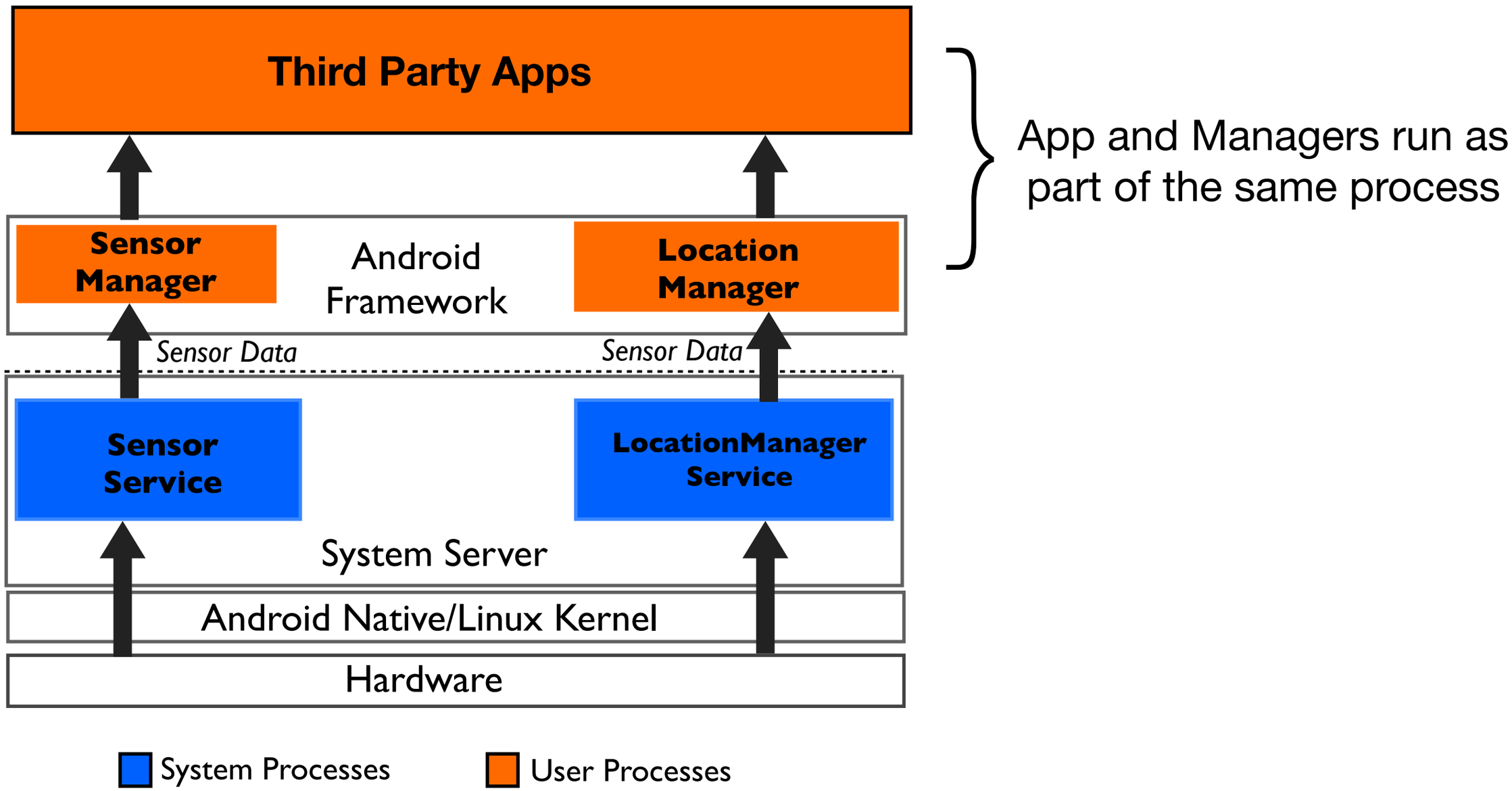
# Sensor subsystem in android and data interception



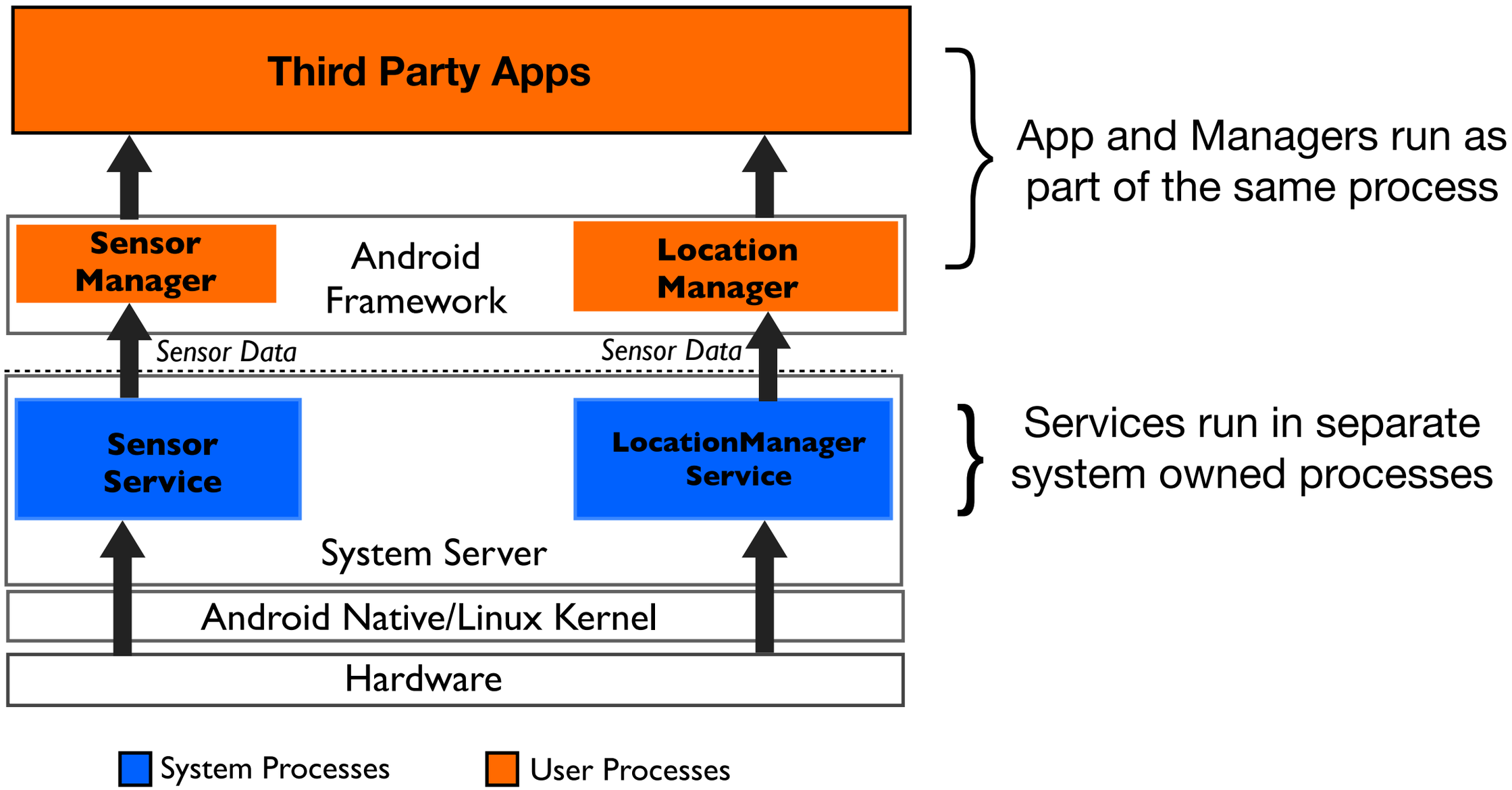
# Sensor subsystem in android and data interception



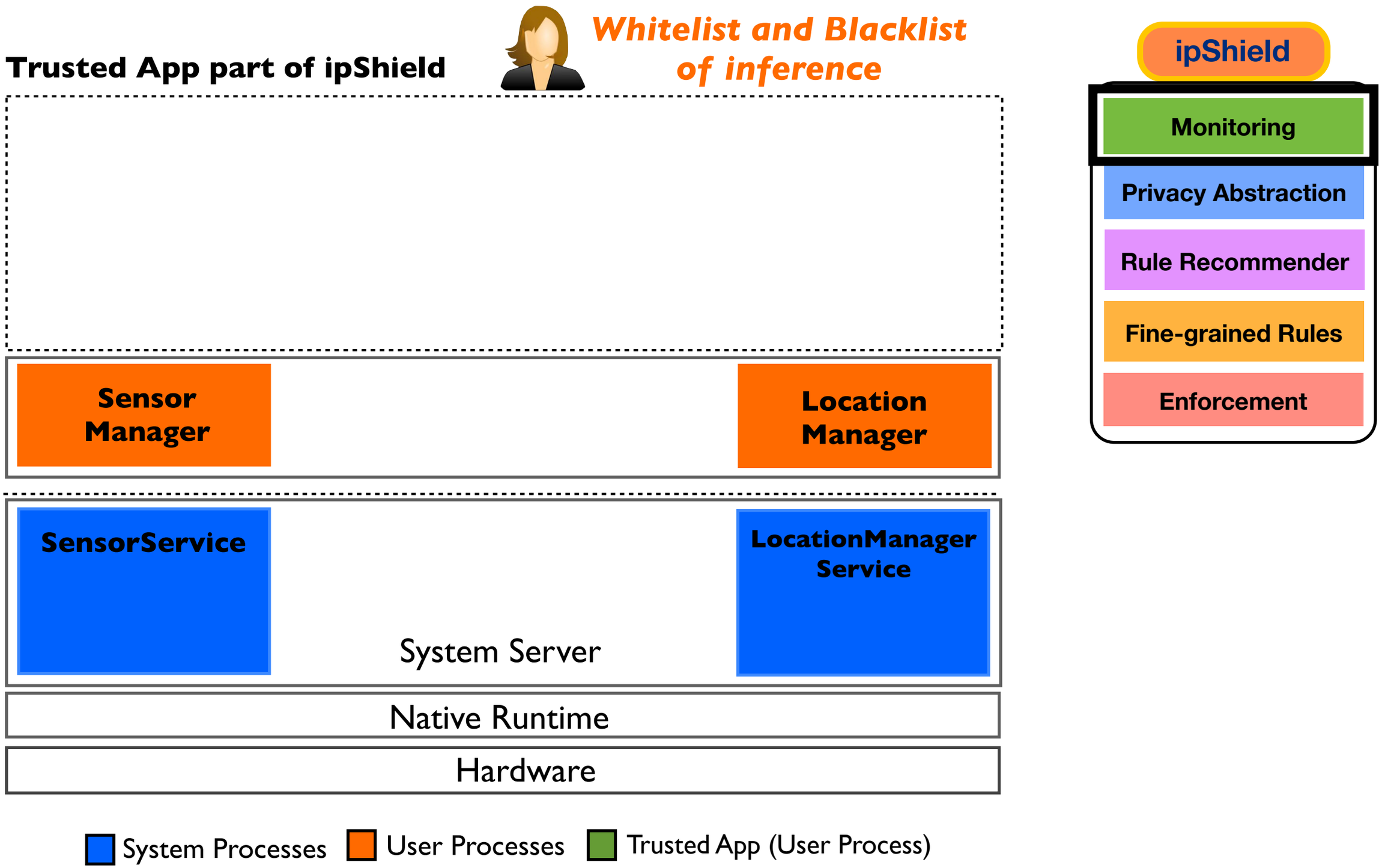
# Sensor subsystem in android and data interception



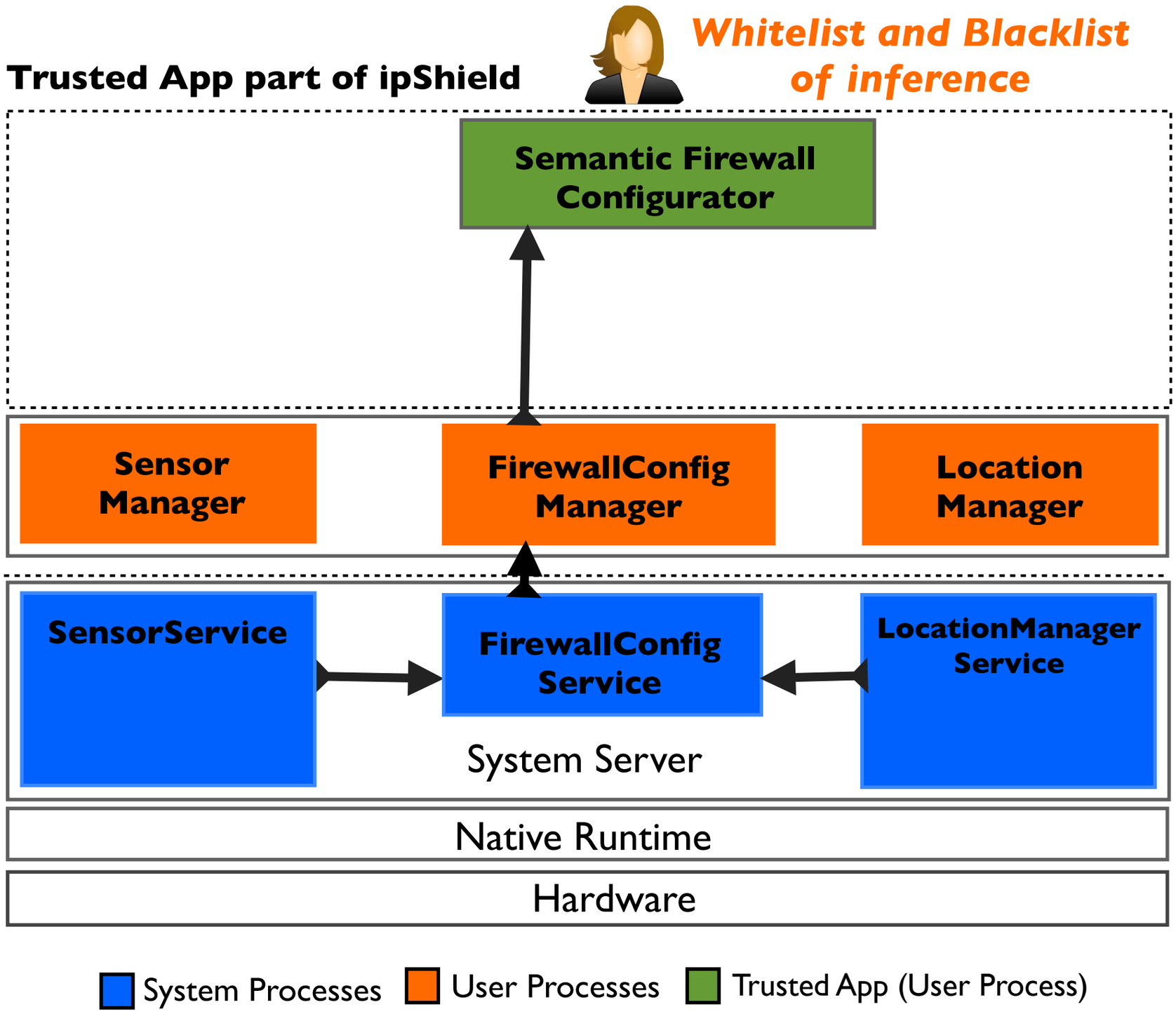
# Sensor subsystem in android and data interception

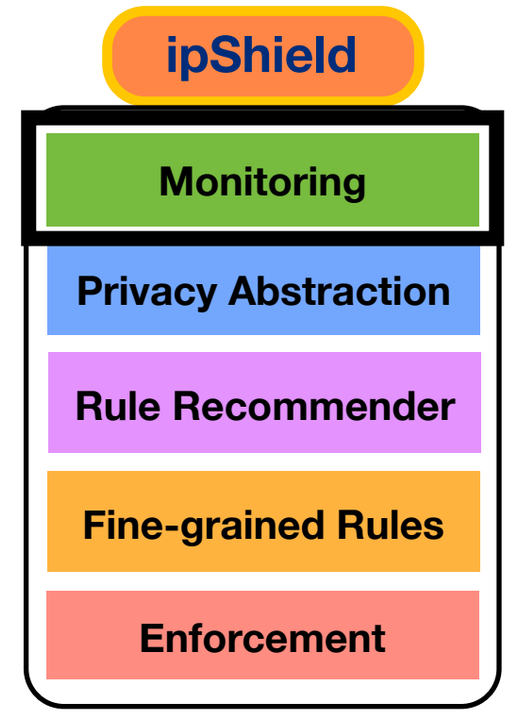


# Implementing ipShield

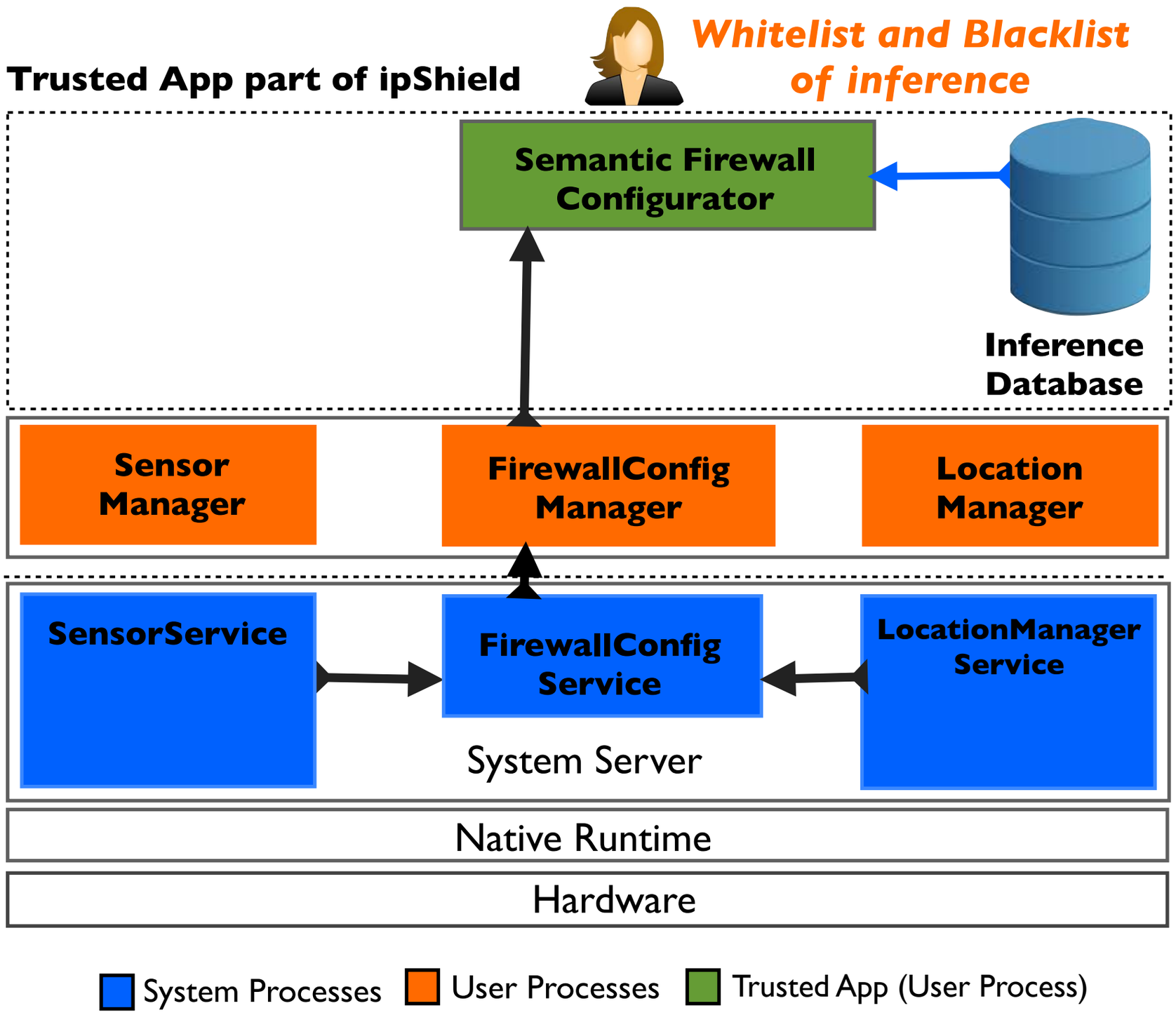


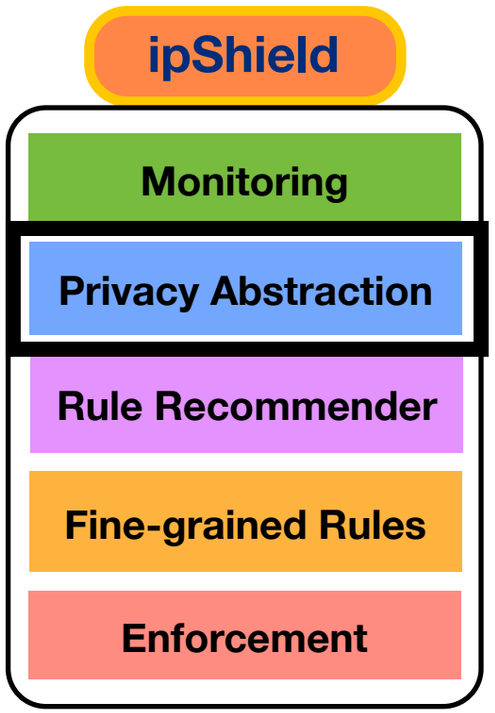
# Implementing ipShield



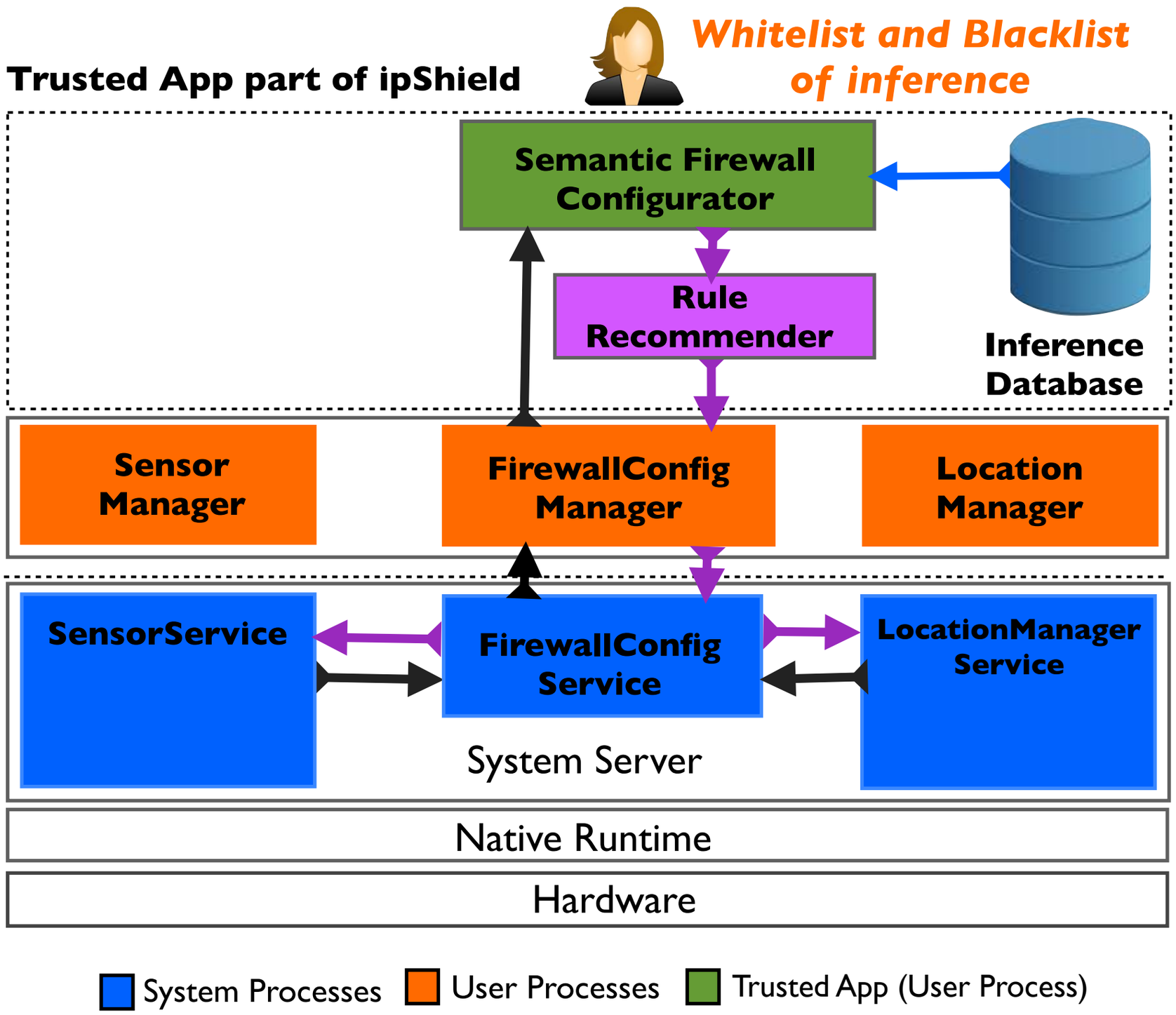



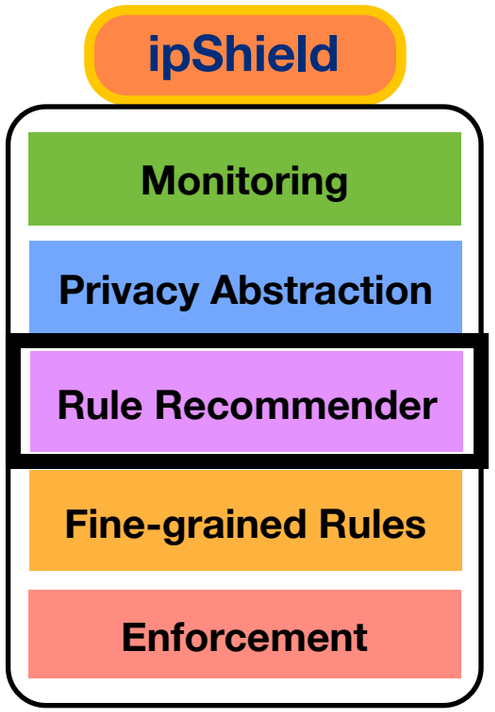
# Implementing ipShield



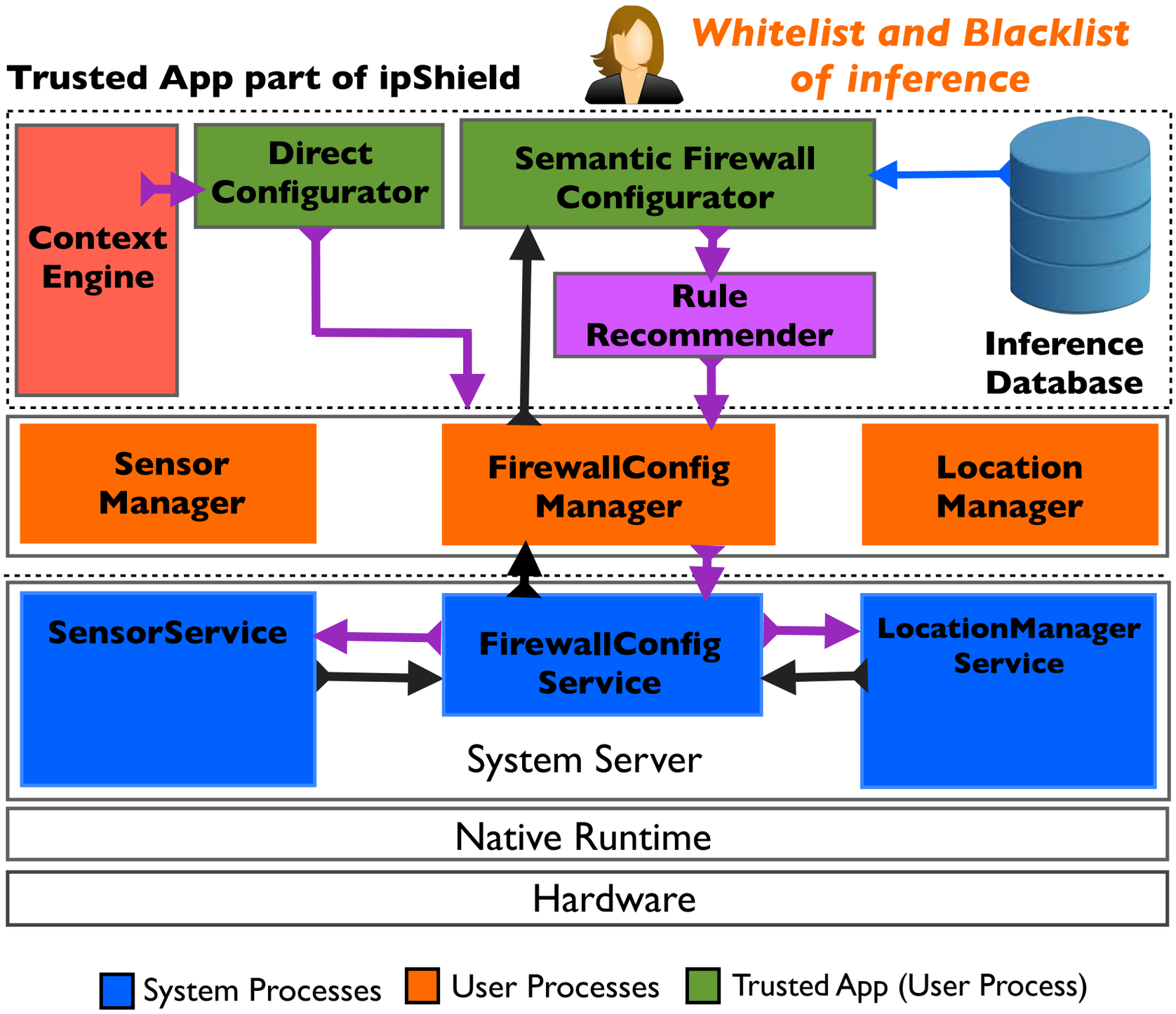



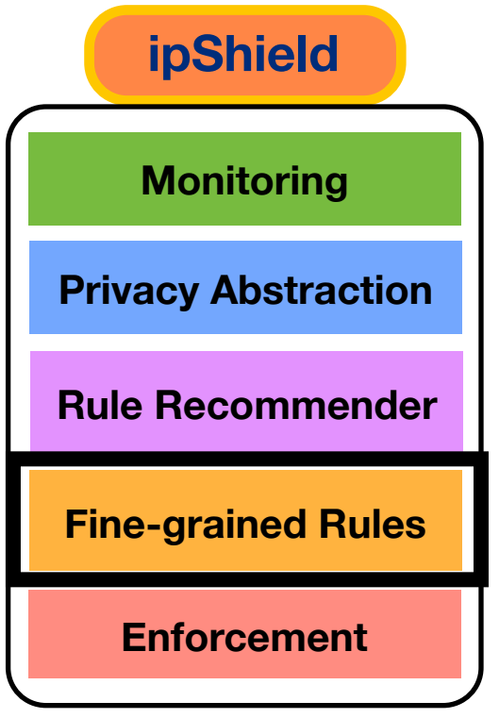
# Implementing ipShield



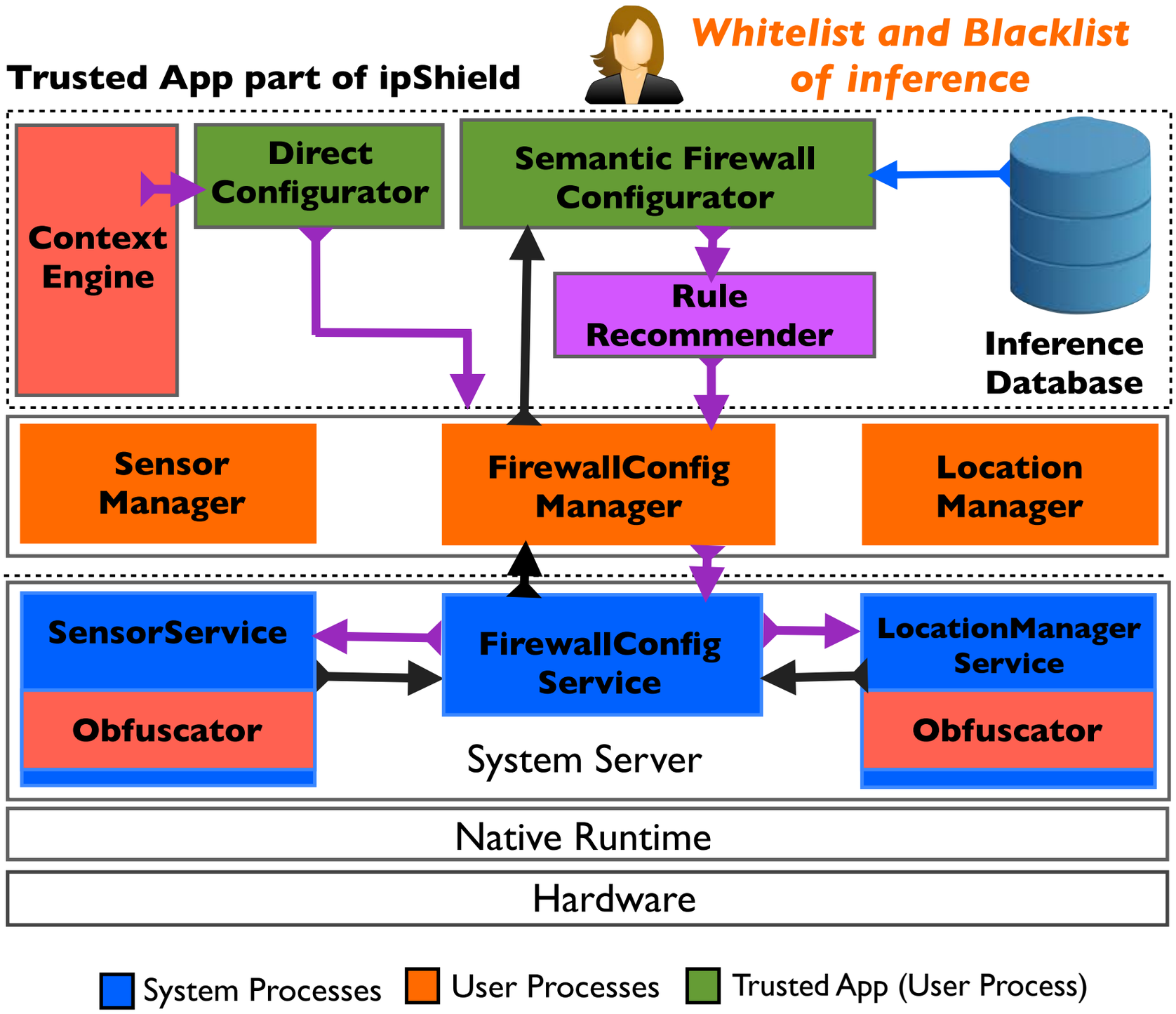



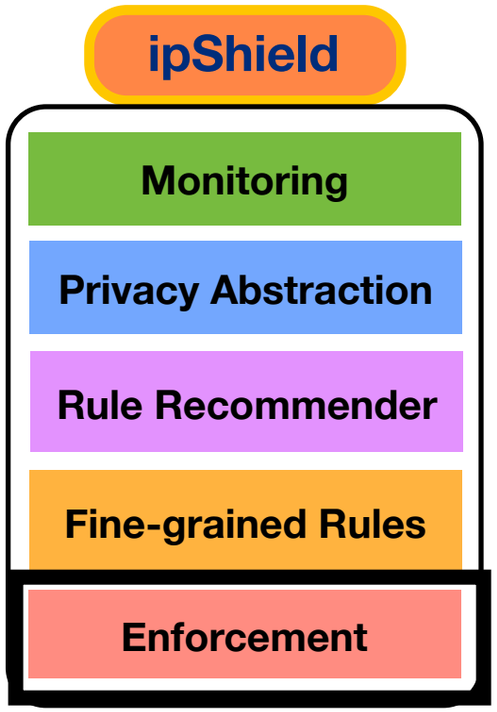
# Implementing ipShield



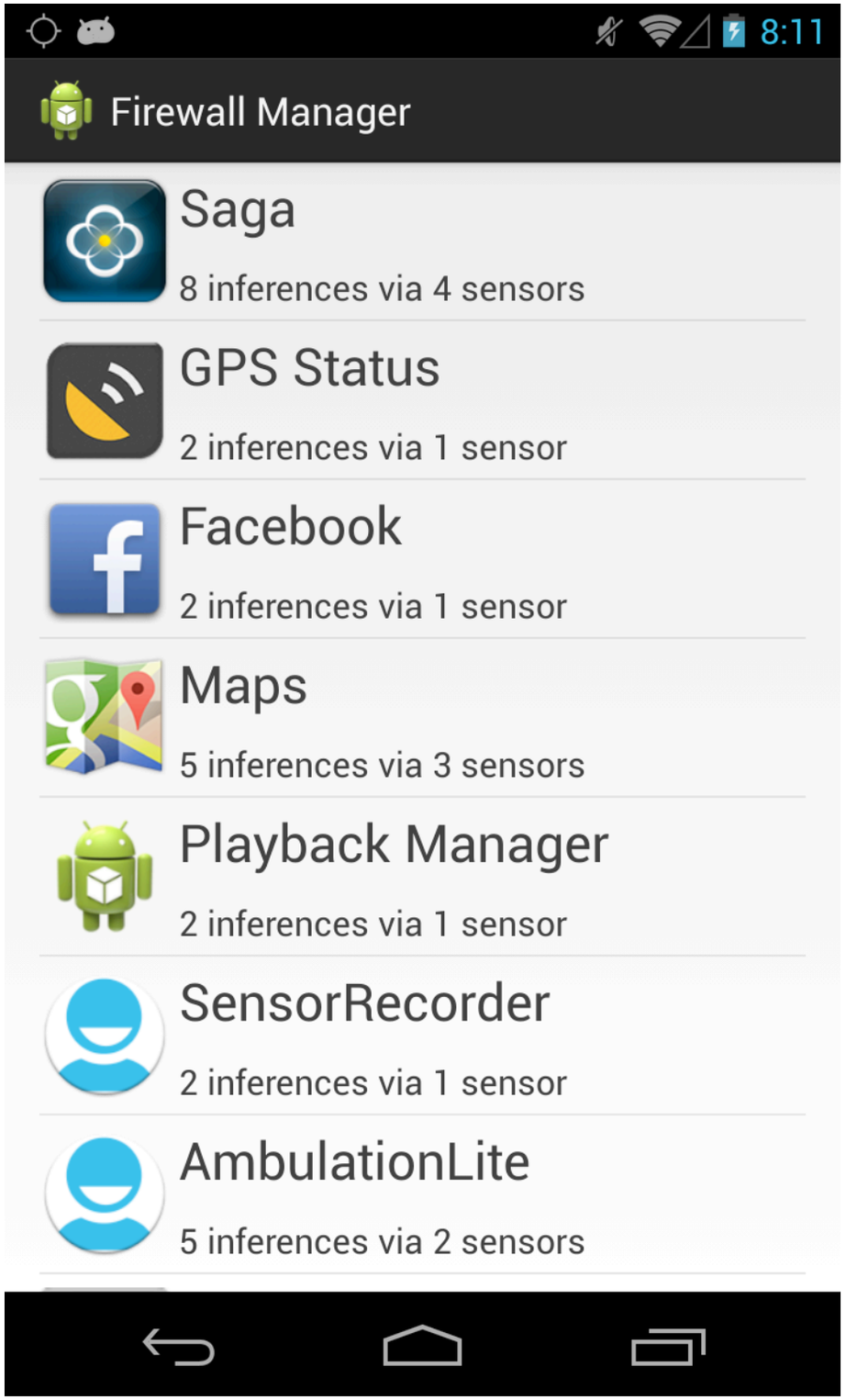



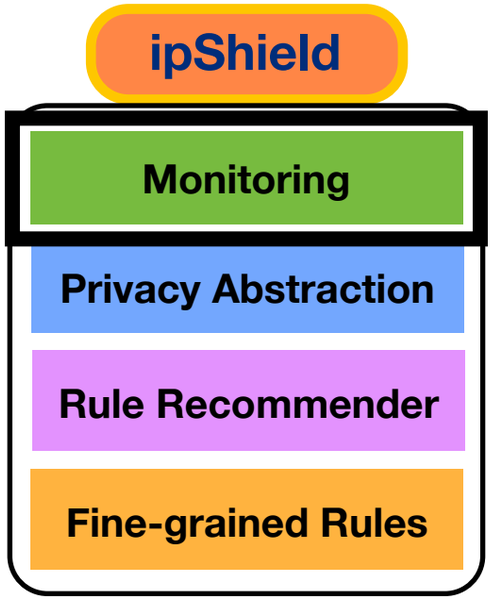
# Implementing ipShield



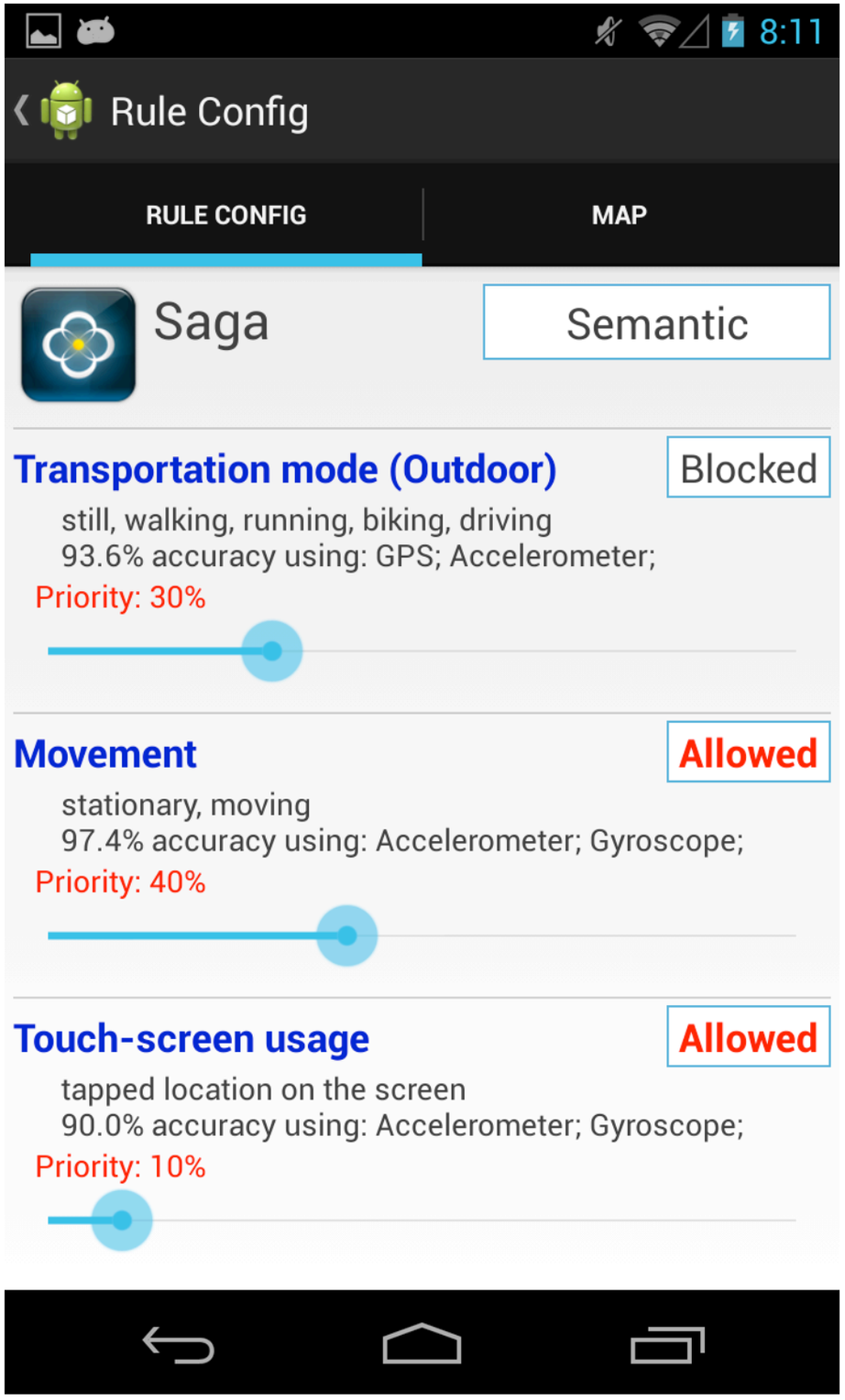



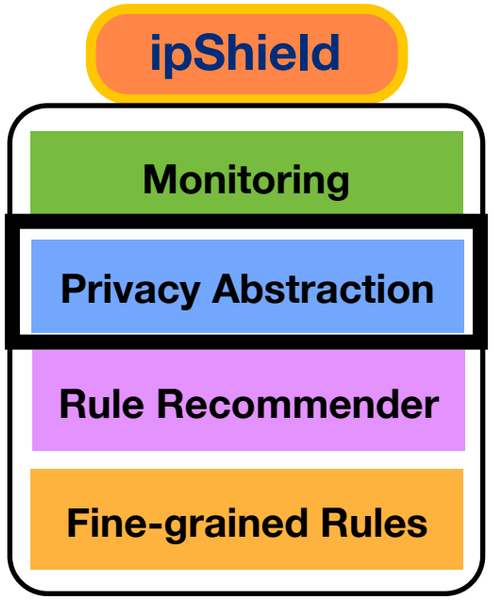
# User interaction with ipShield



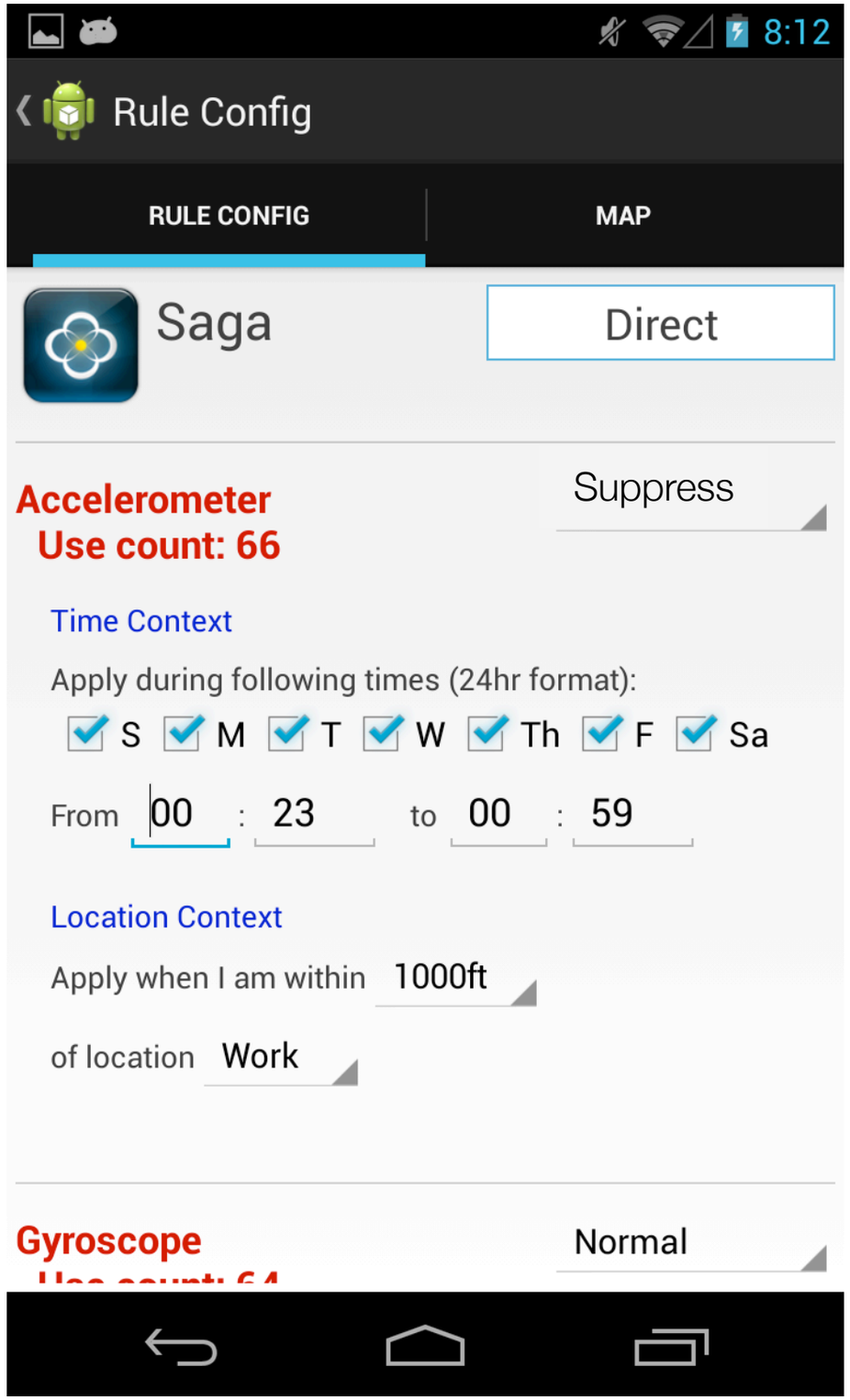



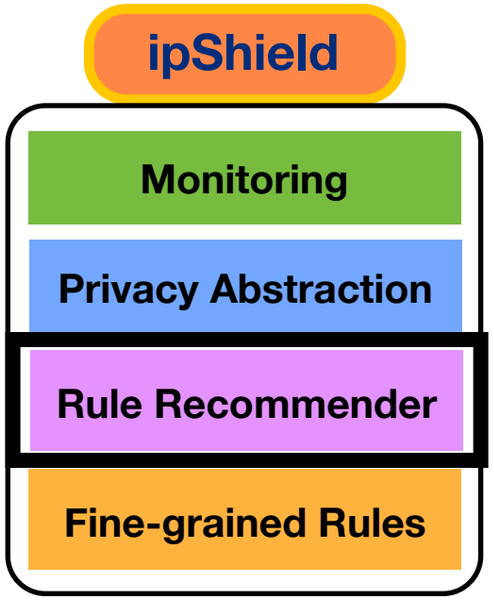
# User interaction with ipShield



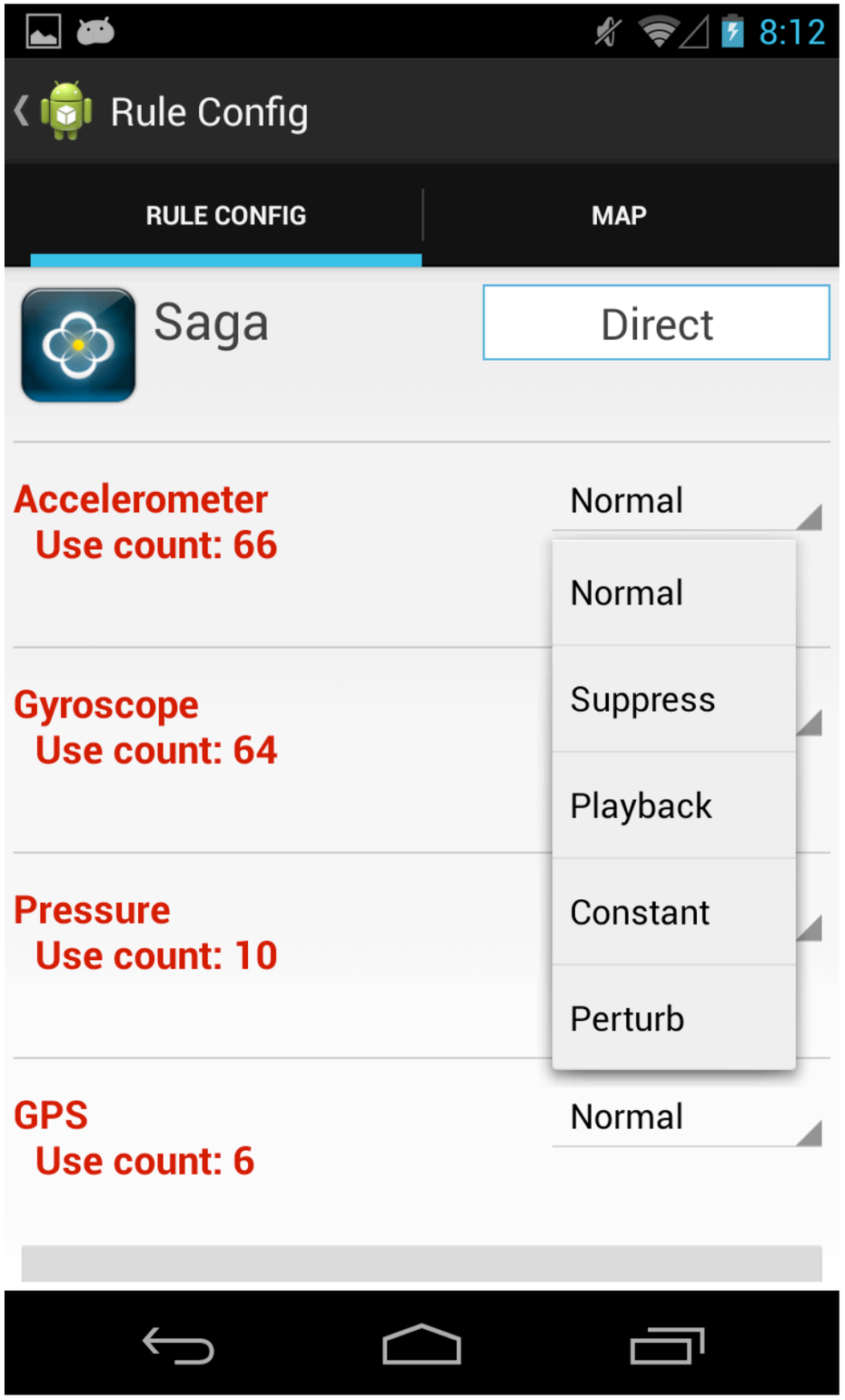



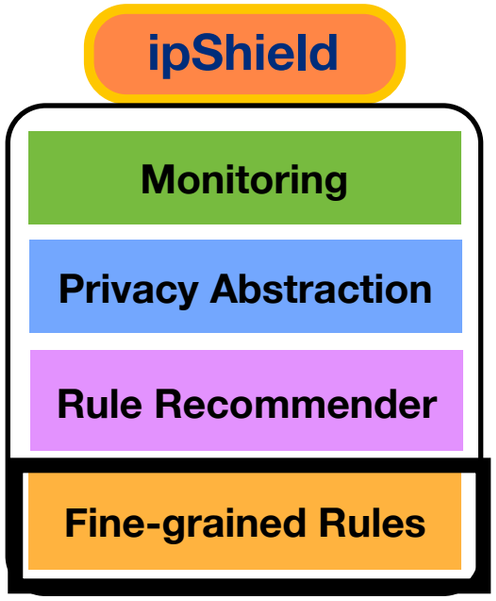
# User interaction with ipShield



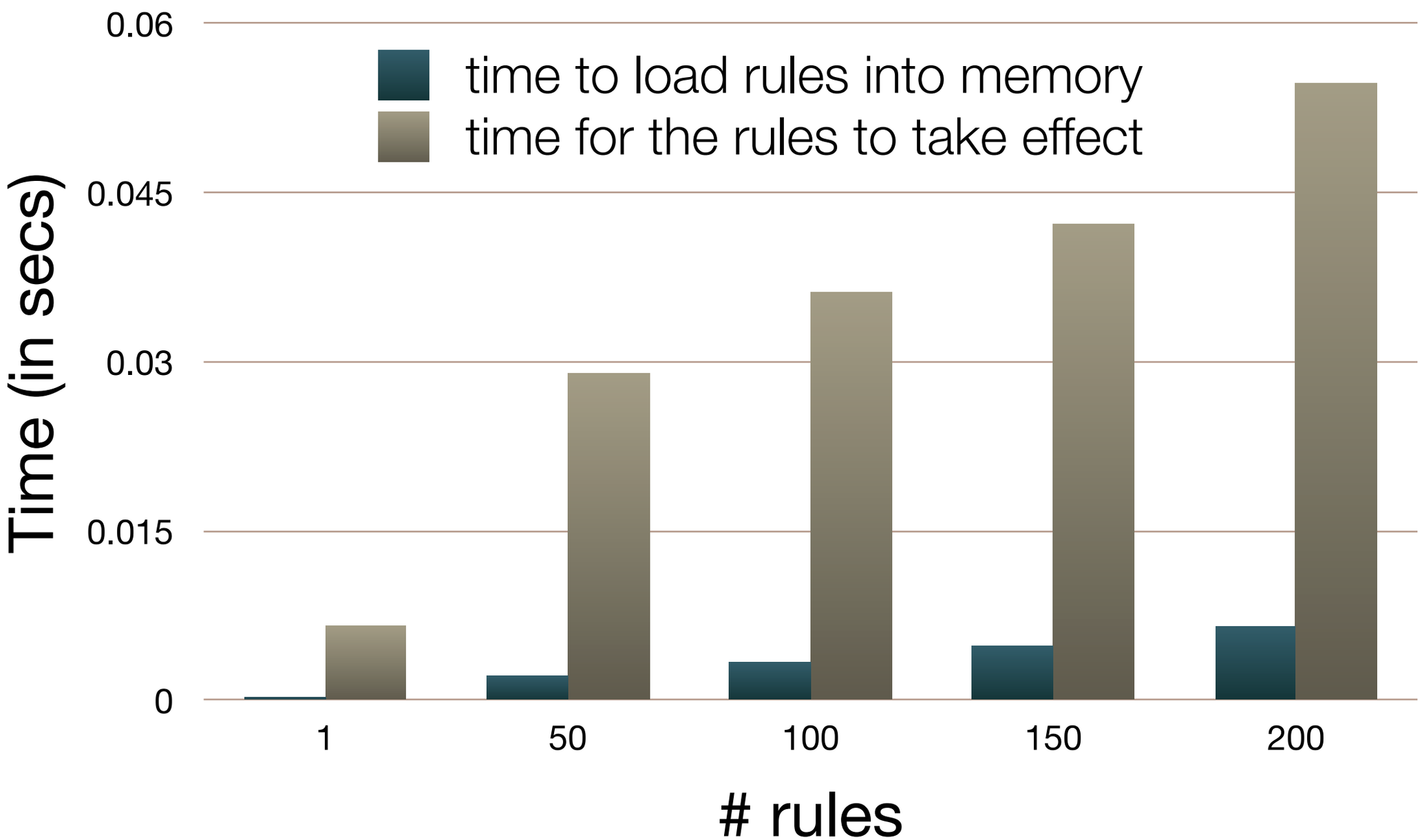



# User interaction with ipShield

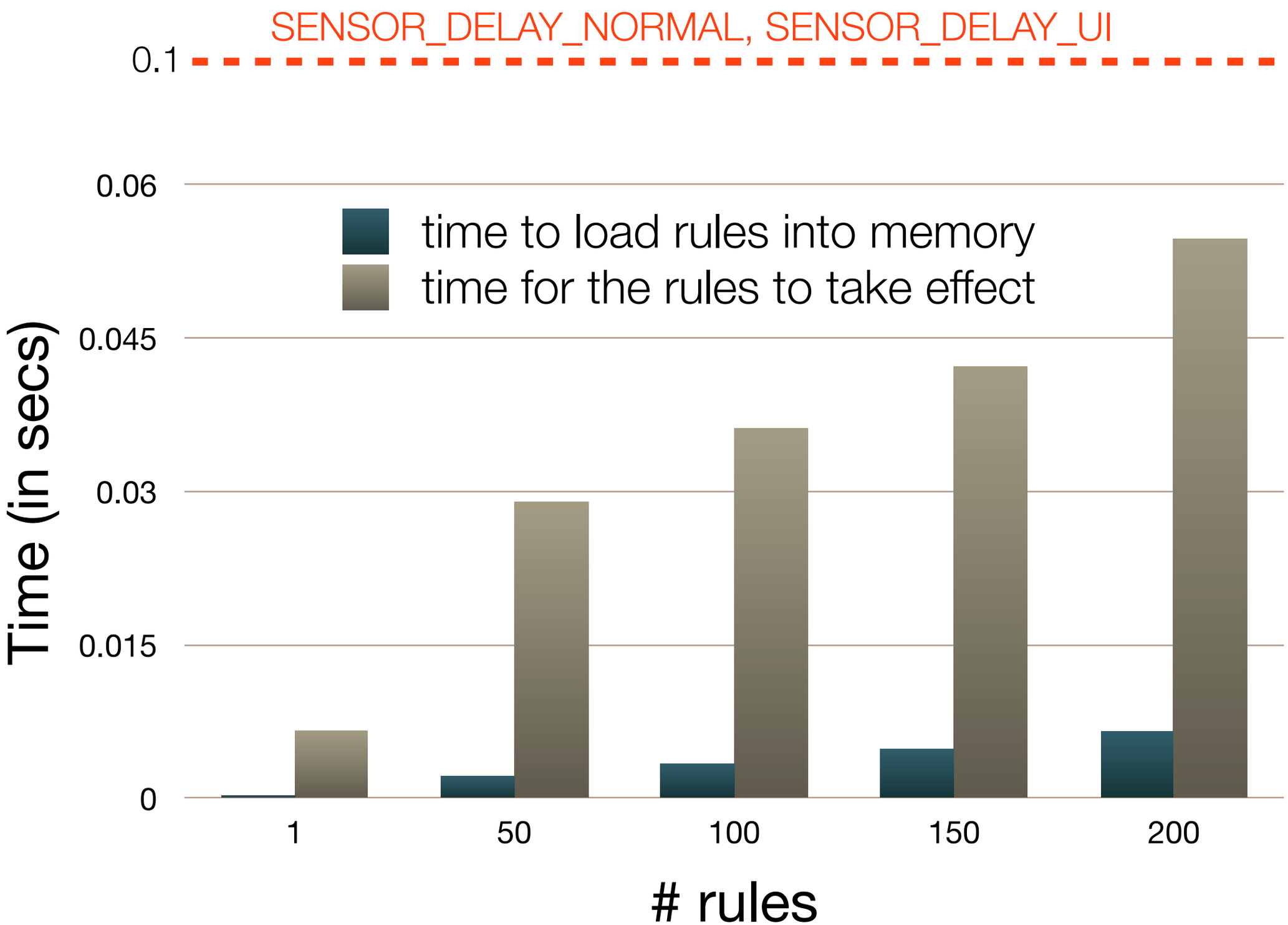


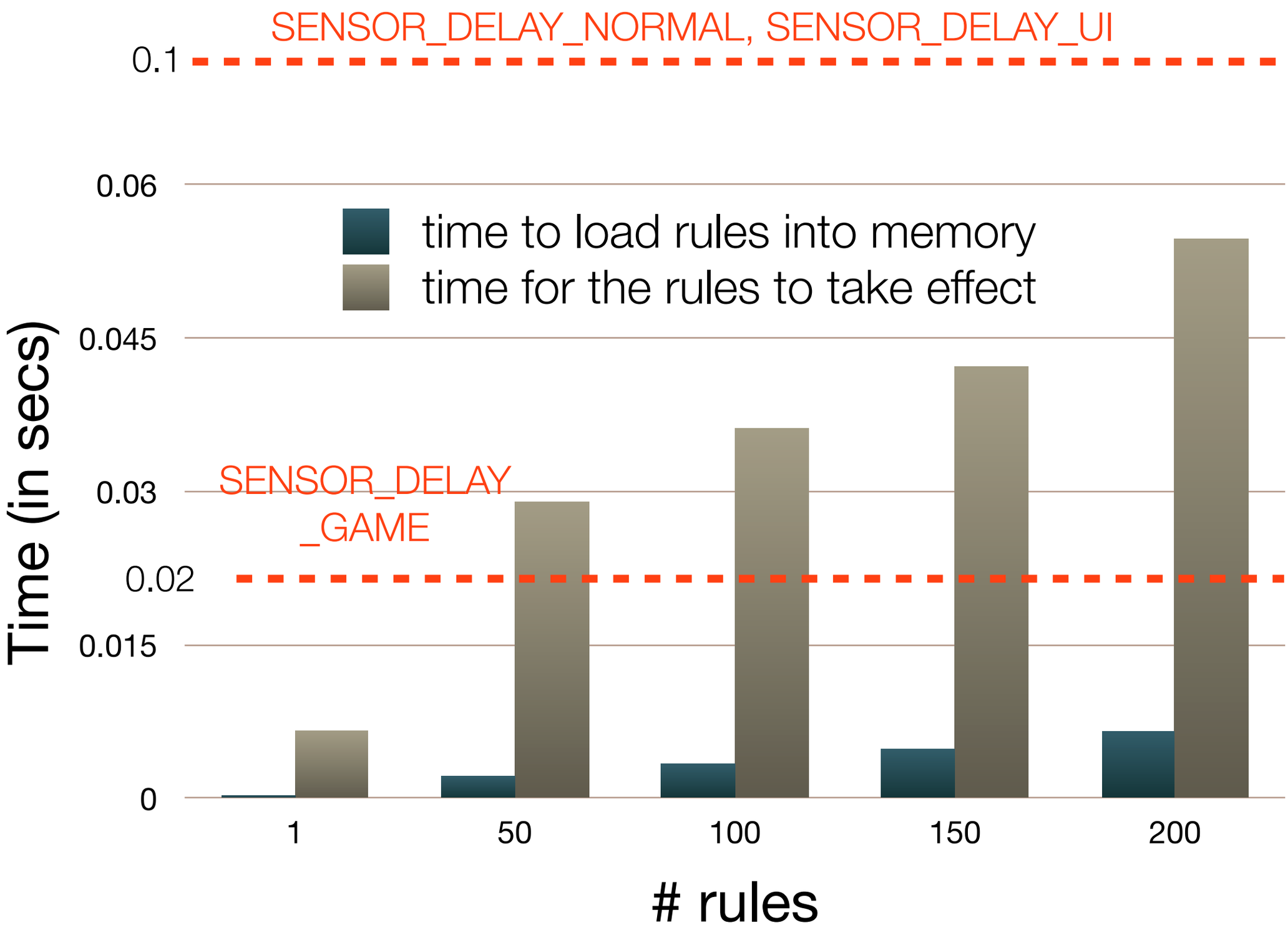
# Feasibility of running ipShield on mobile platforms



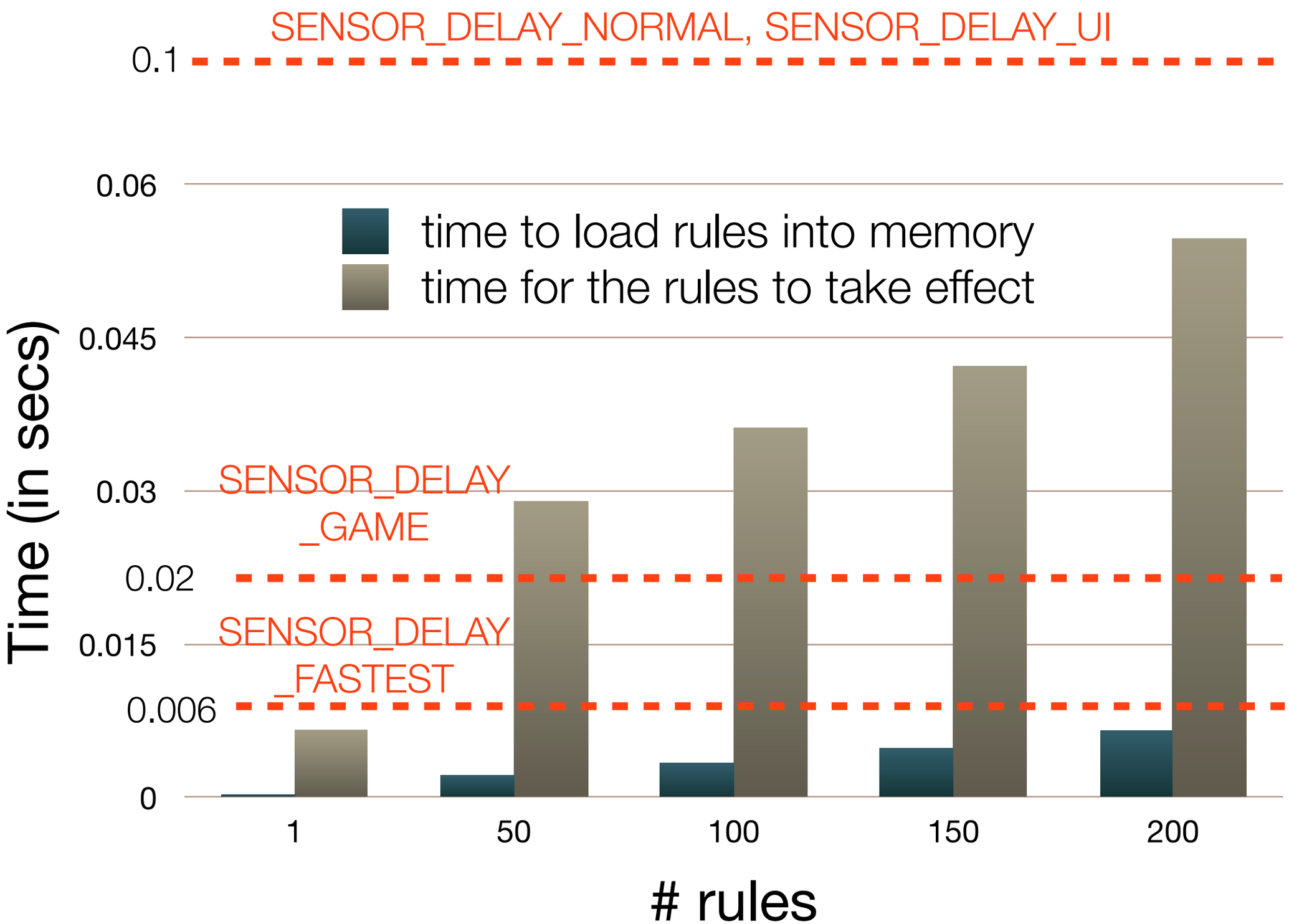
# Feasibility of running ipShield on mobile platforms



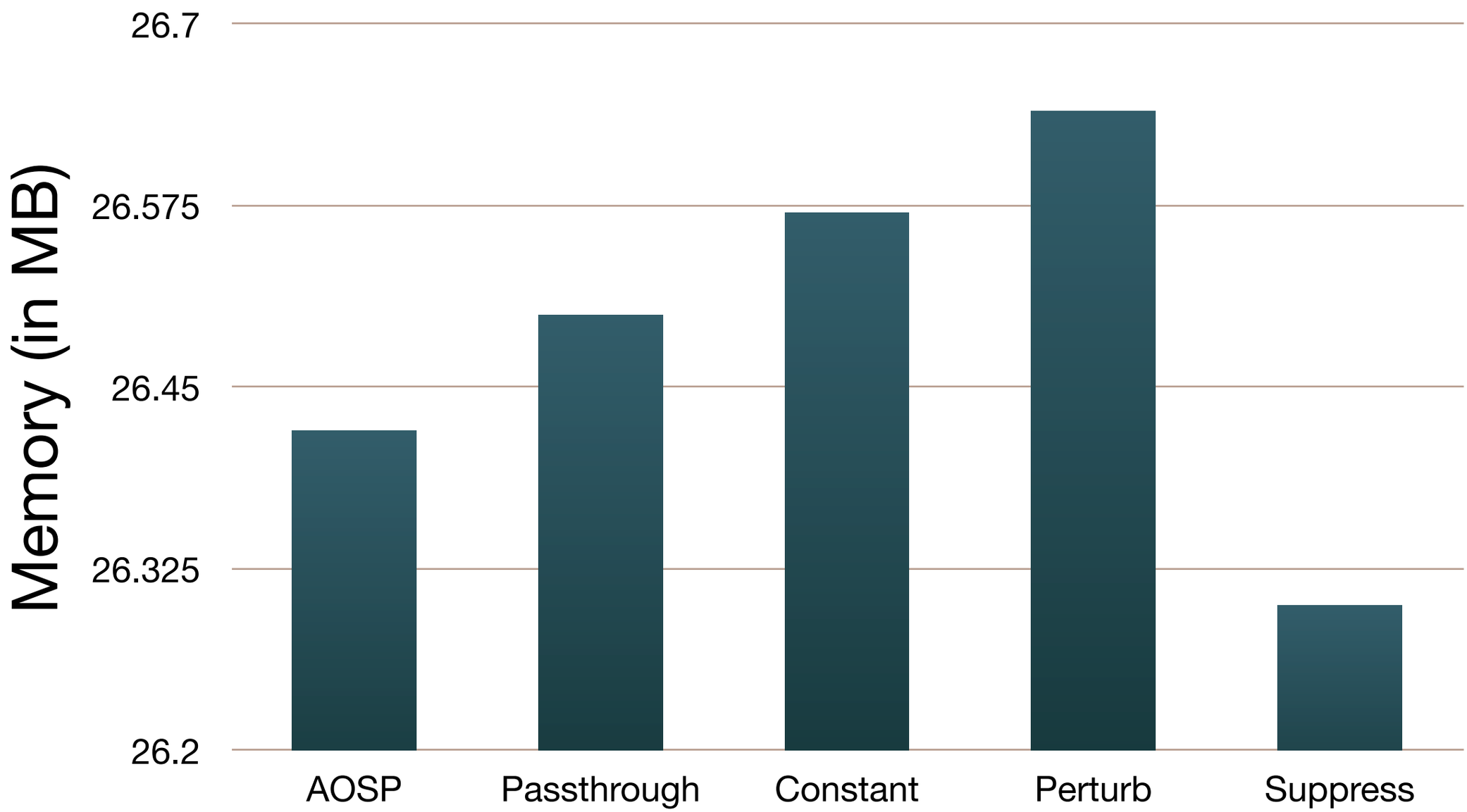
# Feasibility of running ipShield on mobile platforms



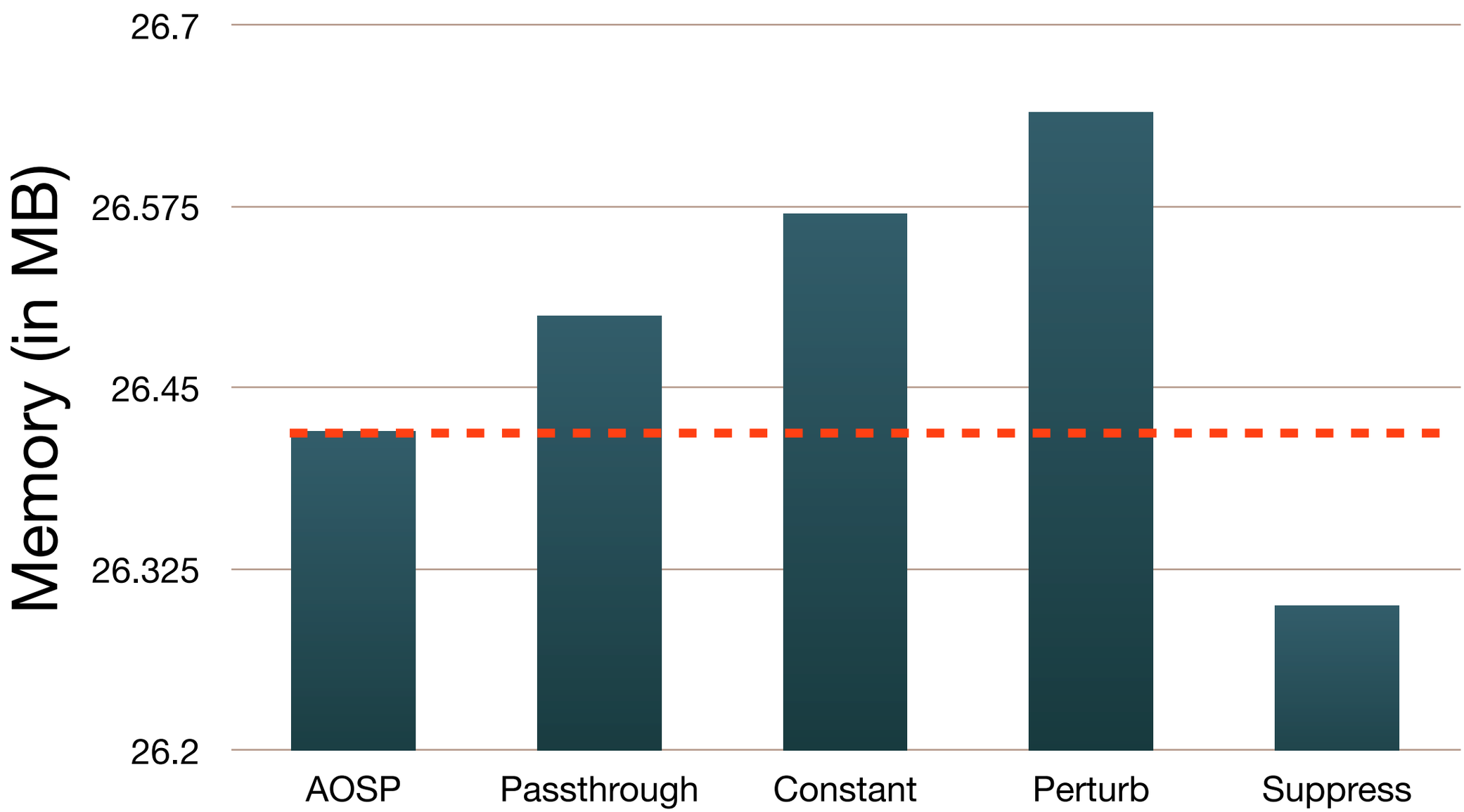
# Feasibility of running ipShield on mobile platforms



# Feasibility of running ipShield on mobile platforms



# Feasibility of running ipShield on mobile platforms



# Concluding Remarks

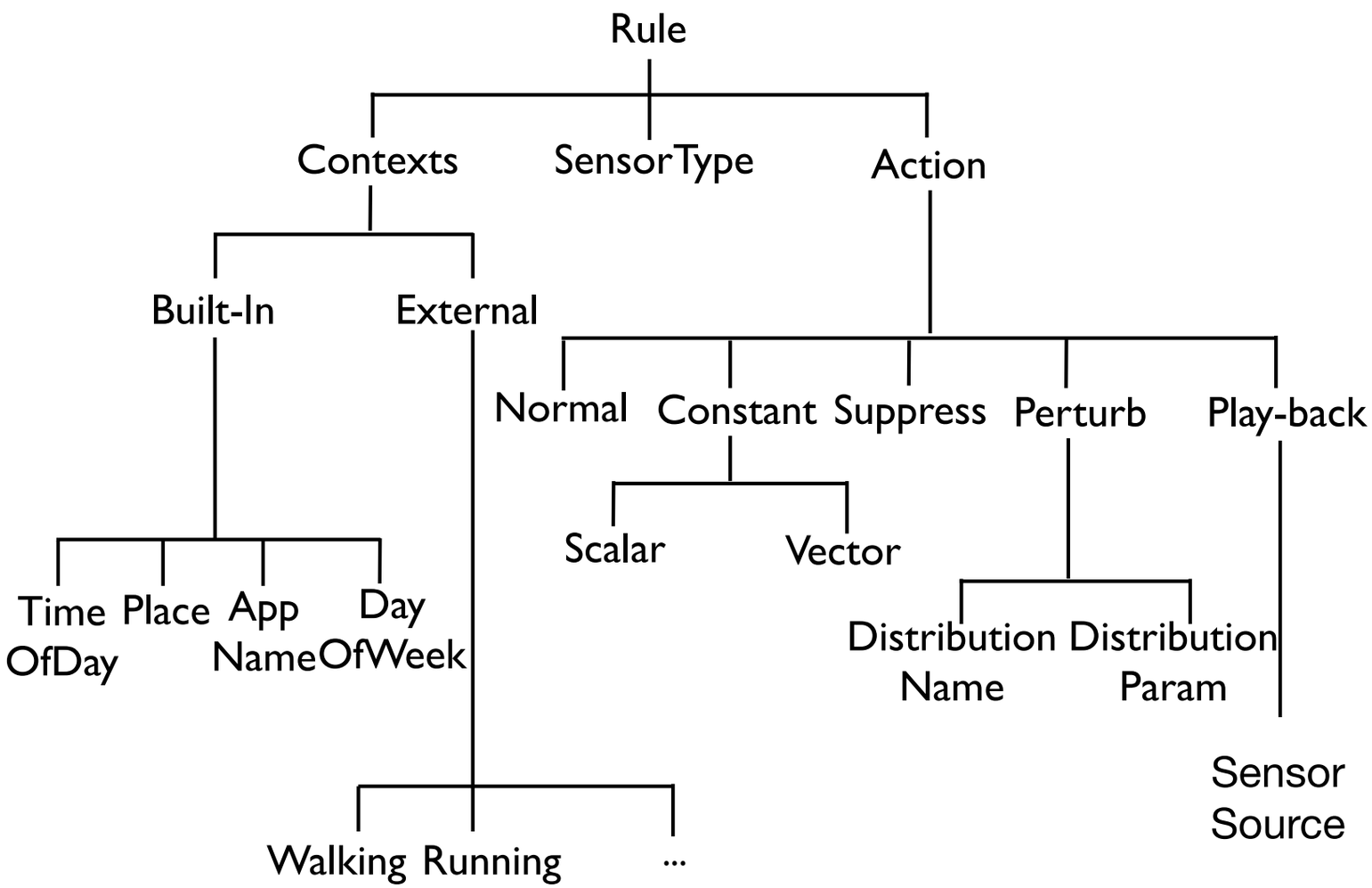
- We designed and implemented ipShield which
  - proposes the use of inferences as the currency for privacy and utility specification.
  - advocates that the burden of configuring fine-grained privacy rules should be shifted from the user to the system.
  - provides insight into how and what data is being used by apps and better visibility into potential risks and consequences of sharing data.

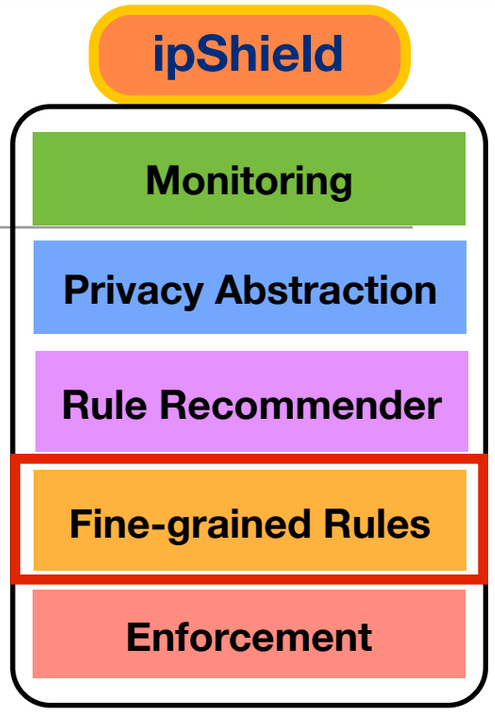
- Going forward we want to...
  - develop the rule recommender to generate rules for obfuscating data.
  - augment ipShield with ability to perform static analysis of app code to better understand the risks presented by the apps.
  - allow crowd-sourcing for bootstrapping of rules.

**ipShield can be downloaded at http://tinyurl.com/ipshieldgit**

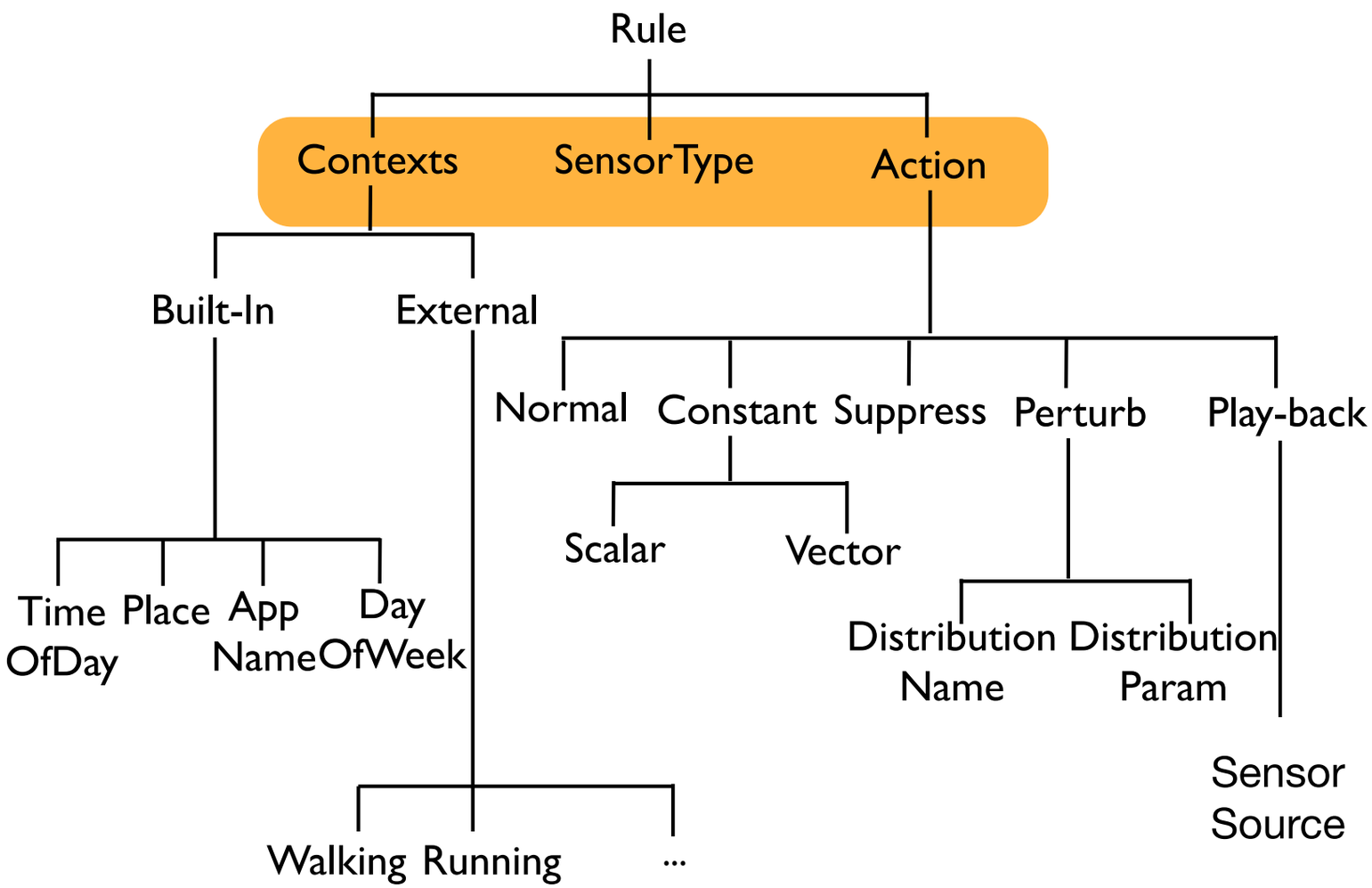
# Thank You

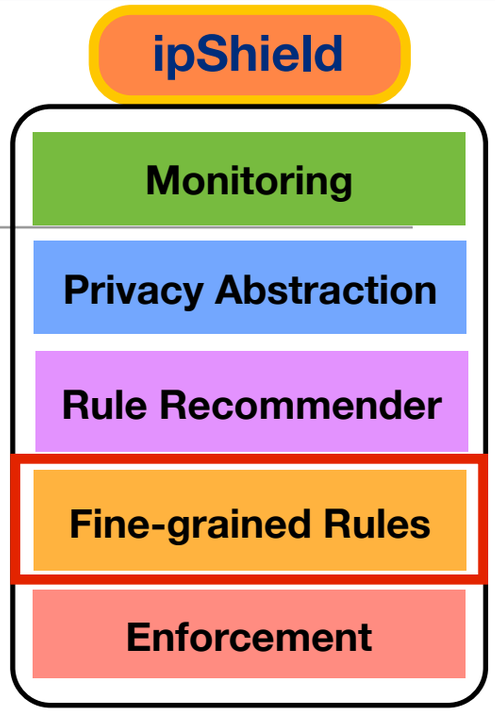
# Rules supported



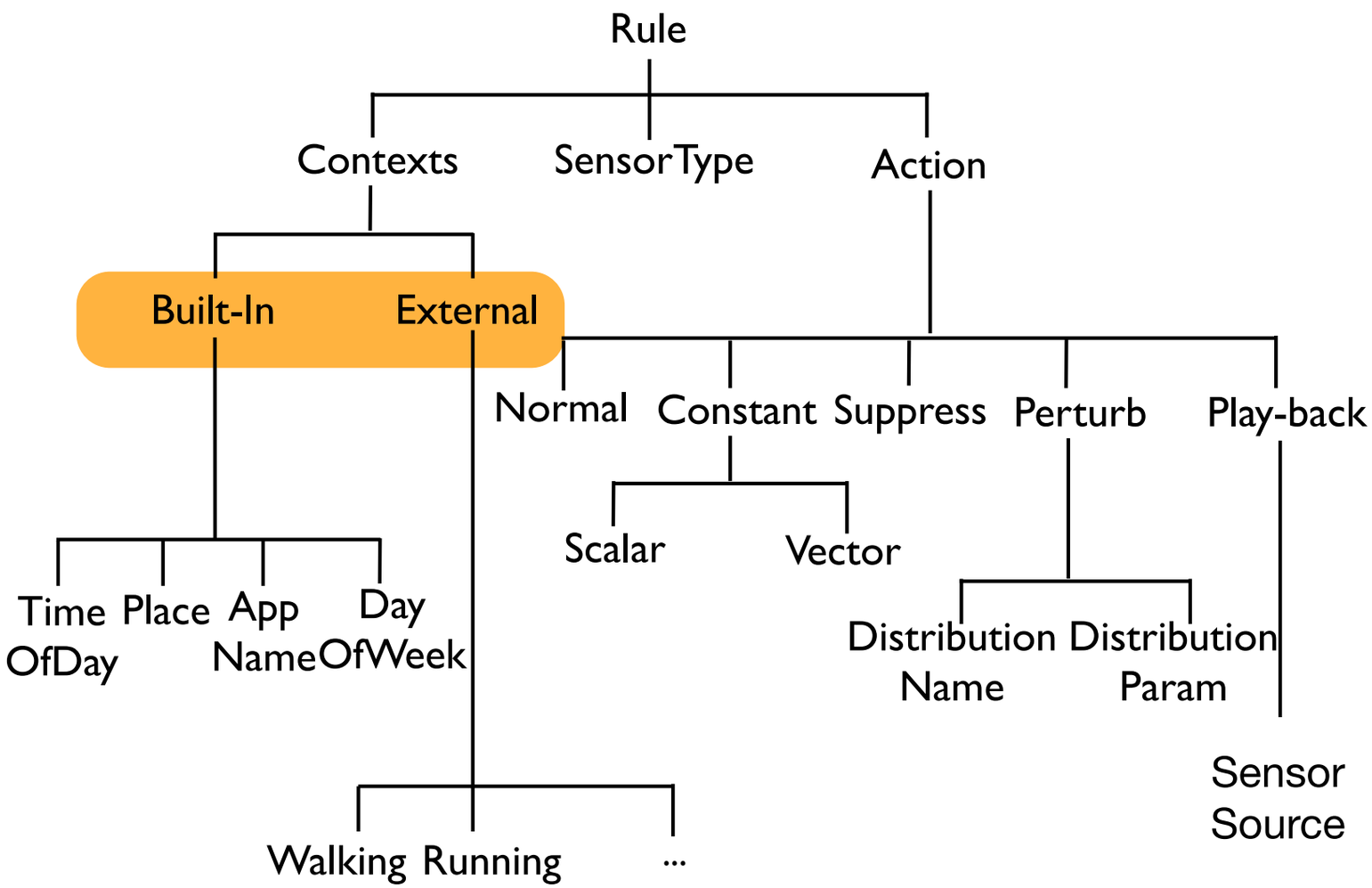



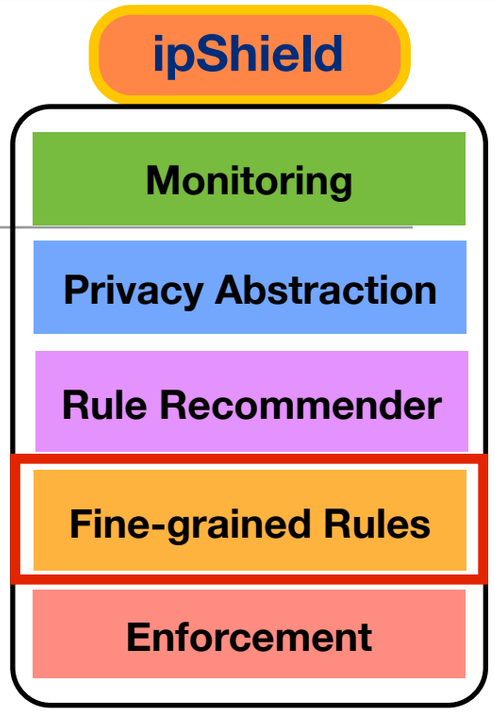
# Rules supported



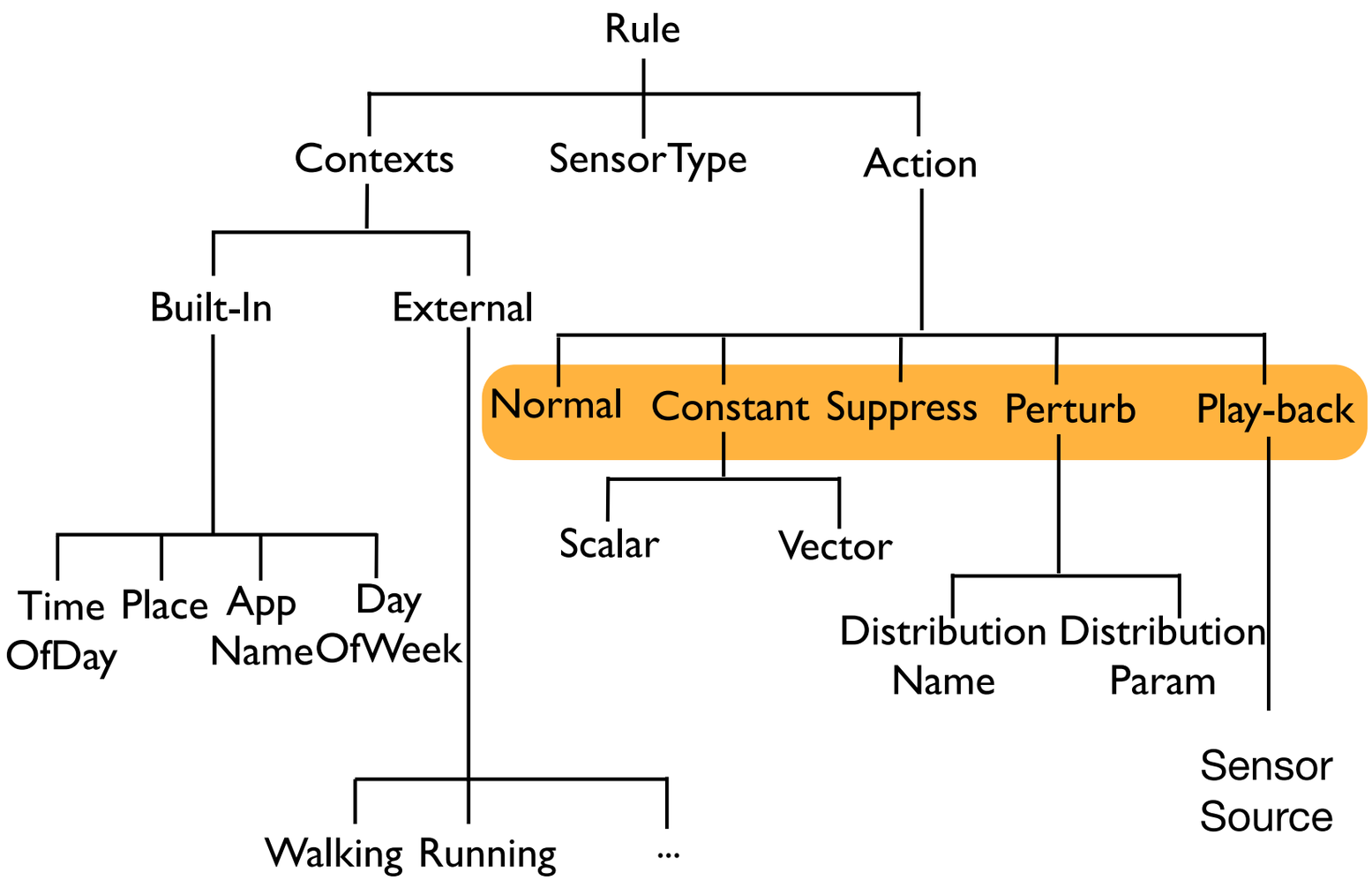



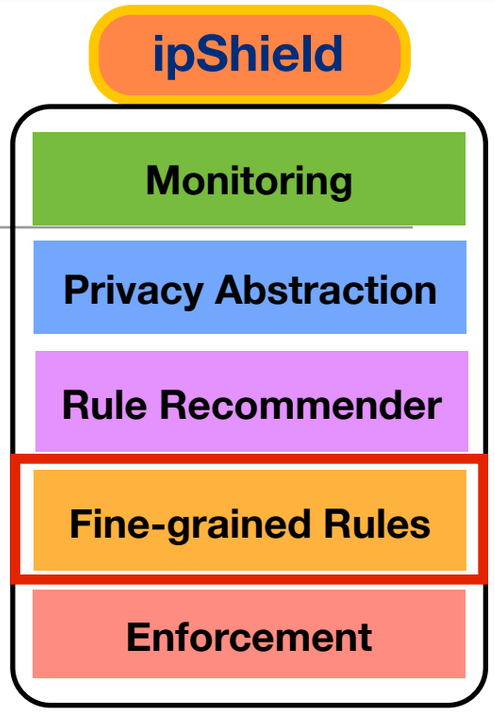
# Rules supported



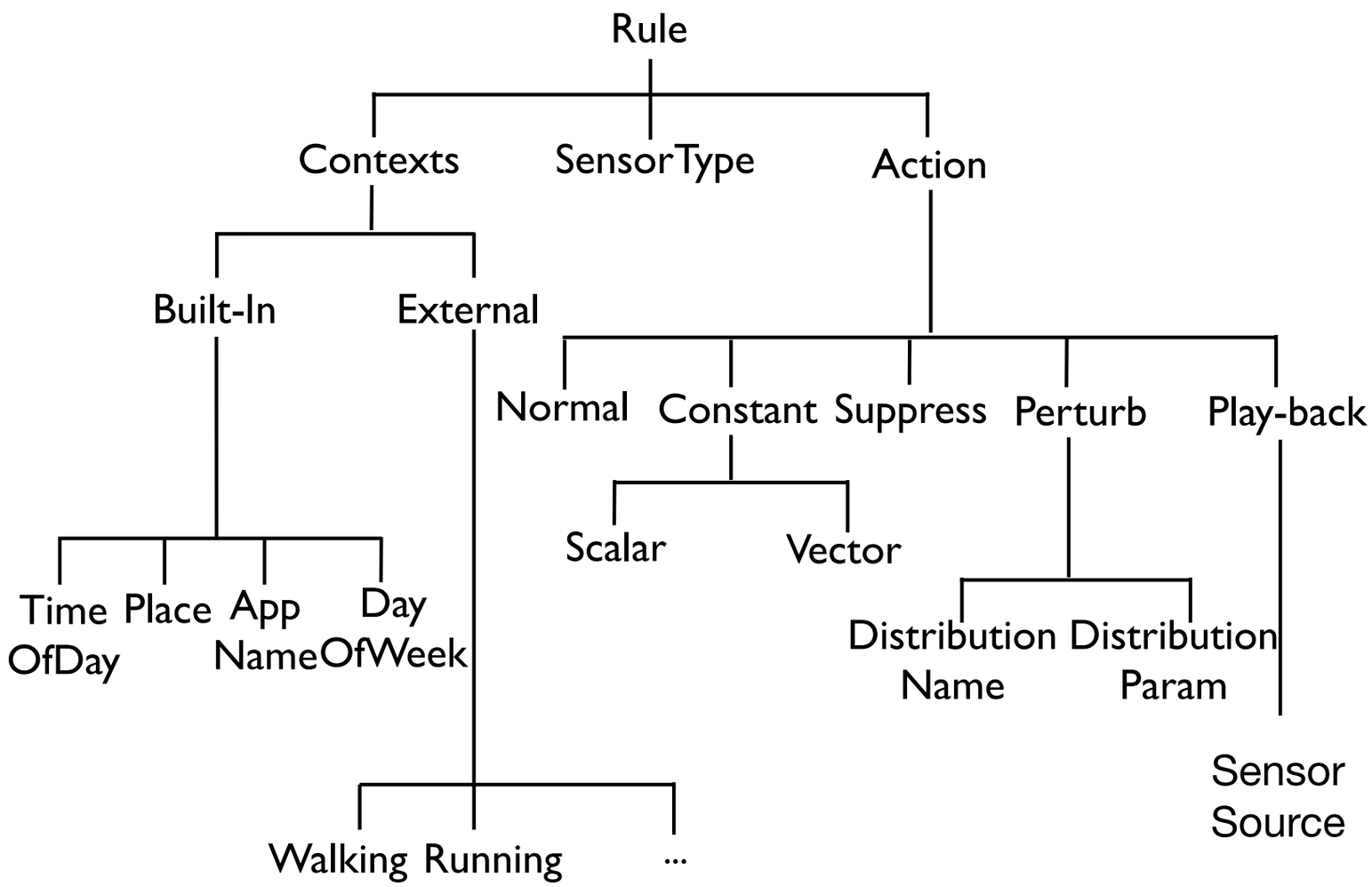



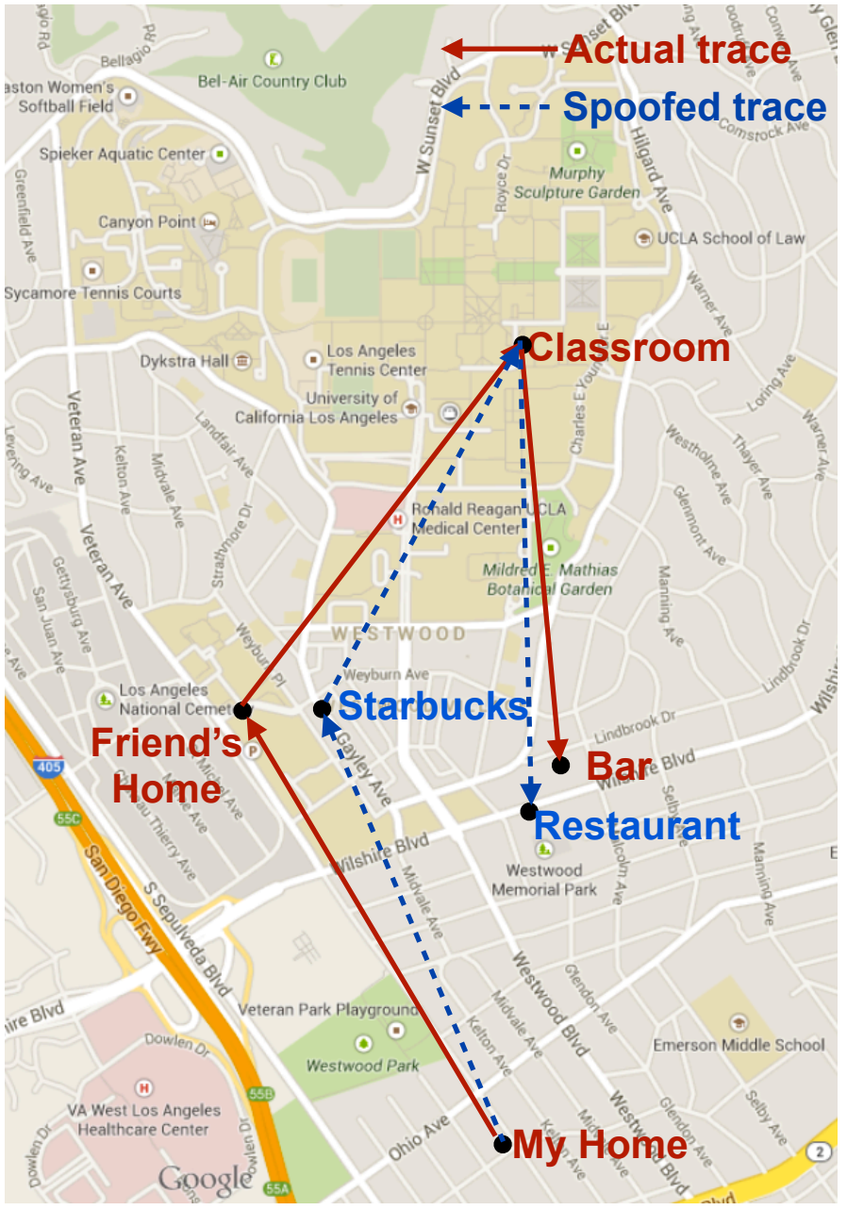
# Rules supported



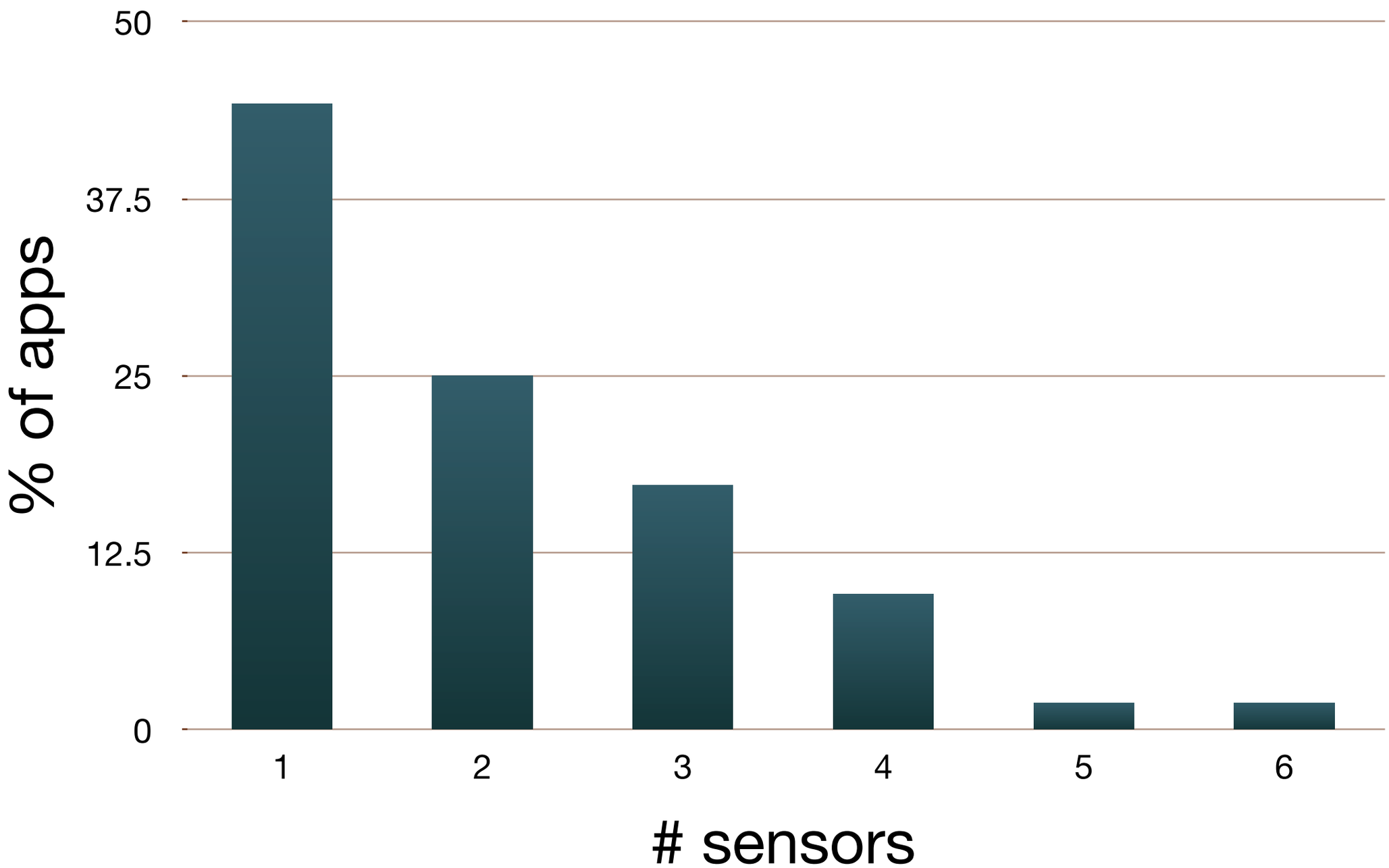
# Rules supported




```
Rule: If ((TimeOfDay in [12am-11:59pm]) and (Place=Bar) and (AppName=Saga)
then apply action = Constant and Value = Restaurant on SensorType = GPS;
```

# Sensor usage for apps



# Distribution of sensors by type