



# Supporting Disconnected Operations and Derived Data Control

Enrico Scalavino, Vaibhav Gowadia, Anandha Gopalan

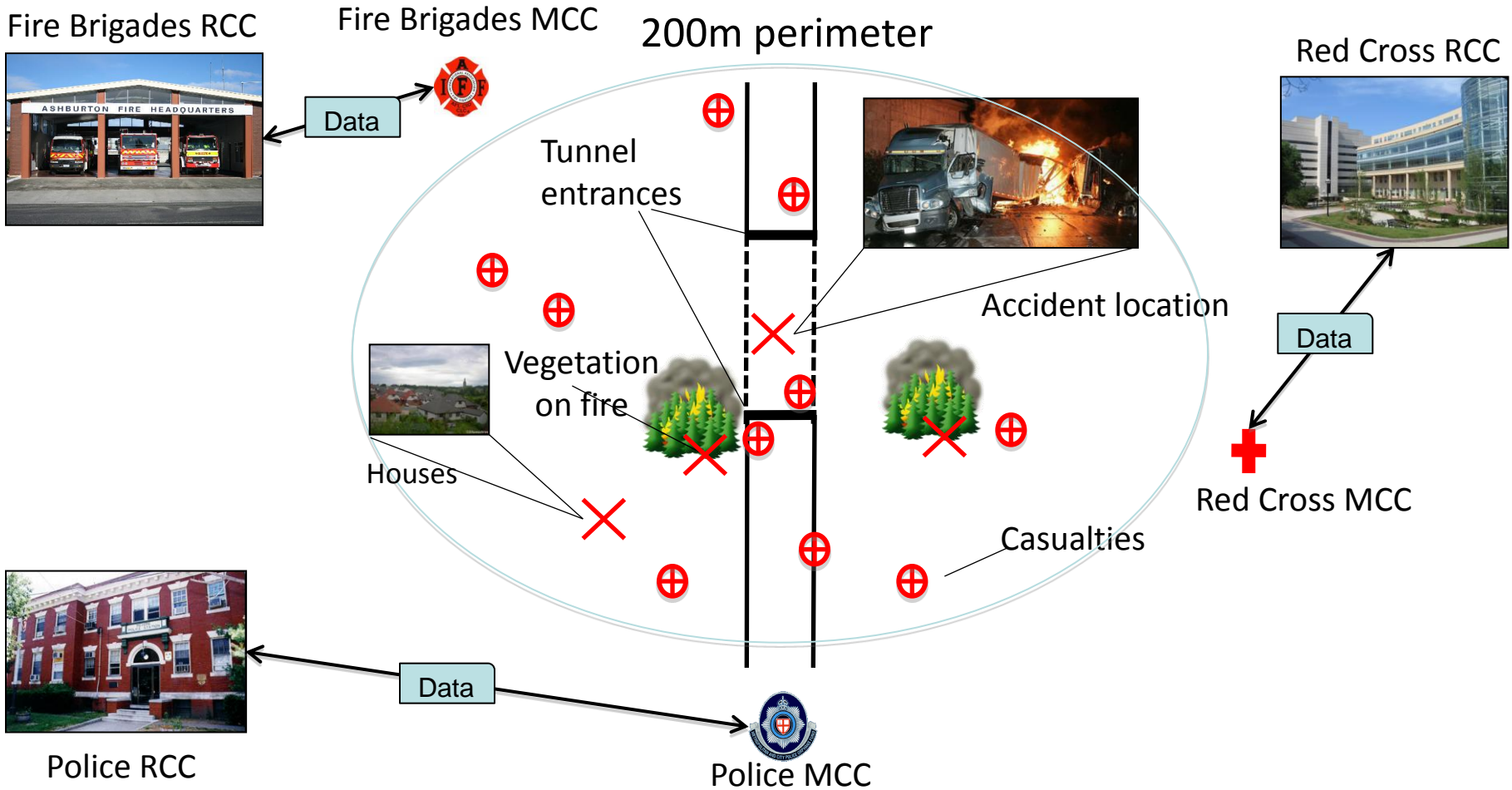
Emil Lupu

Imperial College London

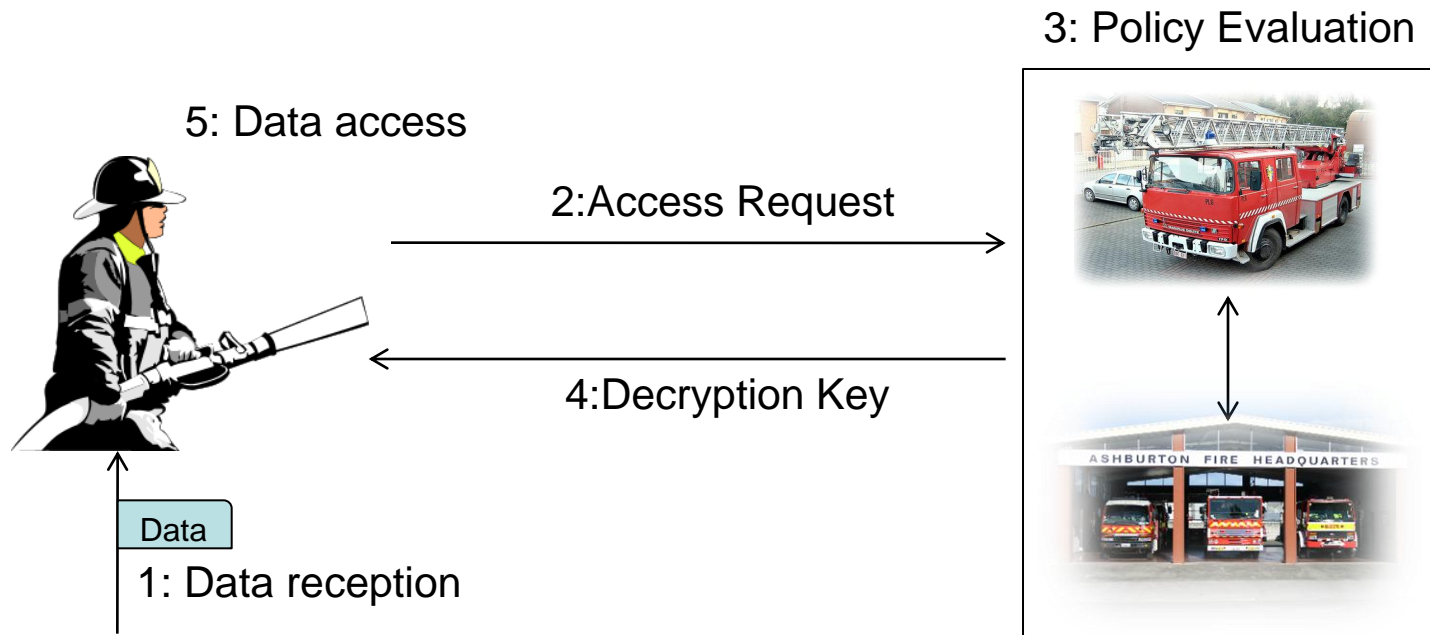
[e.scalavino@imperial.ac.uk](mailto:e.scalavino@imperial.ac.uk)



# A Crisis Scenario



- Disseminated information can be private, confidential, commercially sensitive etc.. and must be protected.
- DRM/ERM solutions: data is encrypted for an evaluation authority (EA) before dissemination. The EA releases the data protection key only after policies have been evaluated.



## **What happens when connectivity is not available?**

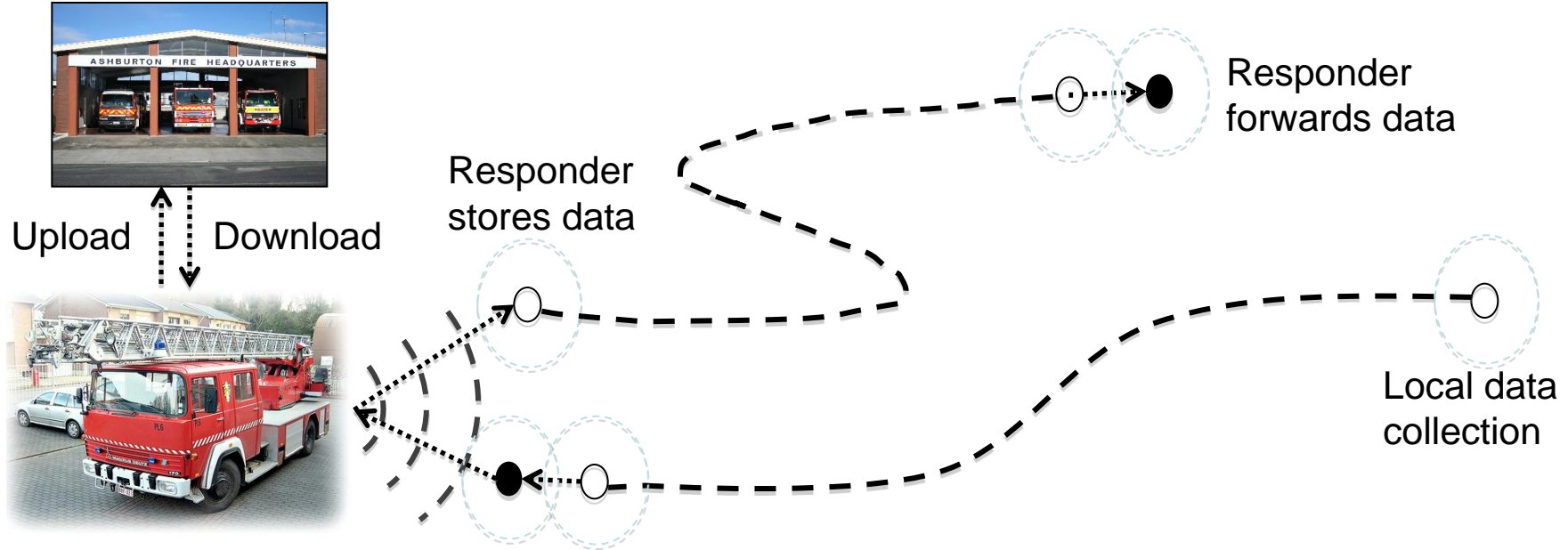
- The accident is underground : Mont Blanc Tunnel (1999). 39 victims in a 53-hours long rescue operation.
- Tactical deployment in remote or hostile regions.
- Failure of long-range communications.

## **What happens to information derived from already existing data?**

- Information on victims' conditions used to infer information on the emergency (e.g. toxic contaminations, fires etc..);
- Information on nearby care centers and maps of the area used to issue commands to rescuers.
- Etc...

# Lack of connectivity ...

requires opportunistic communication and devolution of authority

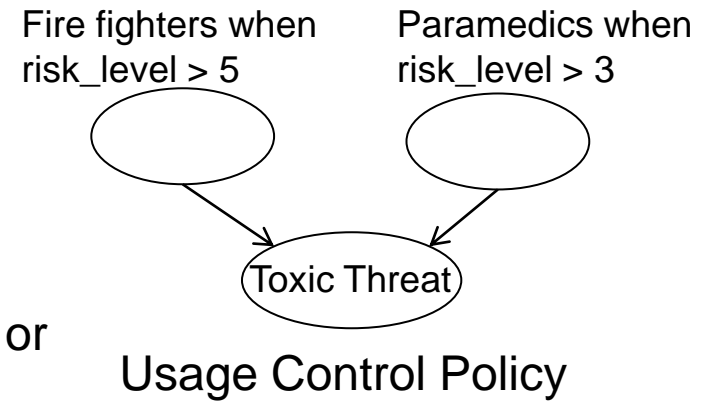


- Data disseminated opportunistically ... but existing security models assume recipients are known.
- Policies cannot be evaluated by remote (centralised authorities)
- **Authority over policy evaluation is devolved from MCC to rescuers and also distributed opportunistically, under policy constraints.**

# Data Categories



- Whenever new data is gathered, it is first associated with a *data category* that specifies:
  - a Usage Control policy;
  - an authority group (a set of users) that can access it;
- Sets of users are characterized by policies and...
- ...associated to a public/private key pair.
- Authority groups can be specified as a conjunction or disjunction of sets of users.
- Before distribution, data is encrypted for its authority group.



## Who can evaluate fire fighters and paramedics?

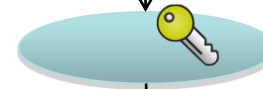
Data Category	Usage Control Policies	EAGs	Node Type
Toxic threat	accessType="read" ^ emergencyCode="red" signed by (FF_MCC, RC_MCC)	fire_fig param	loose
Terrorist threat	accessType="read"	pol_off	
Evacuation info	accessType="read" ^ spaceInclusion(gpsLocation,data.area)	...	...
Building plans	accessType="read" ^ emergencyCode="red" signed by (FF_MCC)	...	...
Facilities and services	...	...	...
Casualties	...	...	...

# PAES: Policy-Based Authority Evaluation Scheme

MCCs: Roots of authority



Who can evaluate policies  $n1...nk$ ? ....



...  
.

....

Who can say who can access the data?  
Who can evaluate policies  $11...1i$ ?



....



Who can access the data?  
Who can evaluate the Usage Control Policy?



....

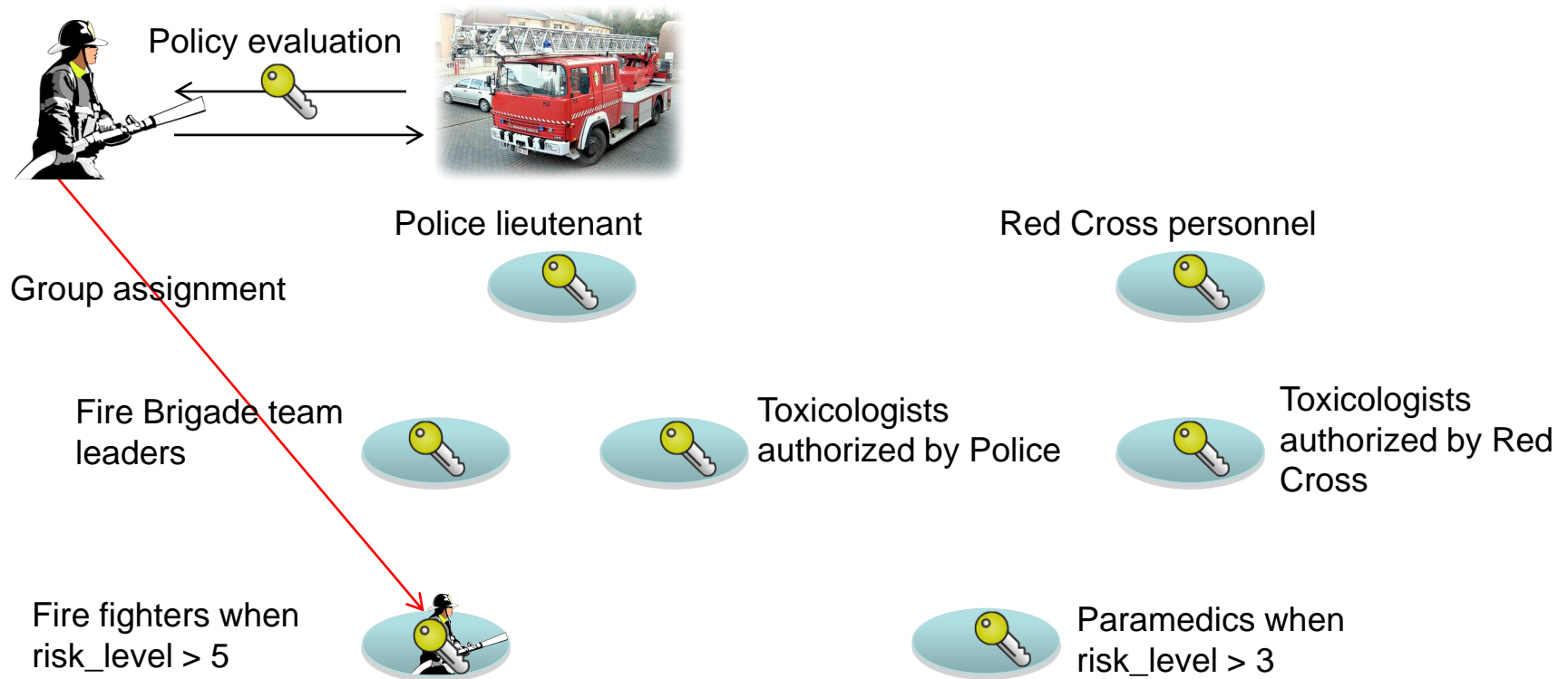


How can the data be used?



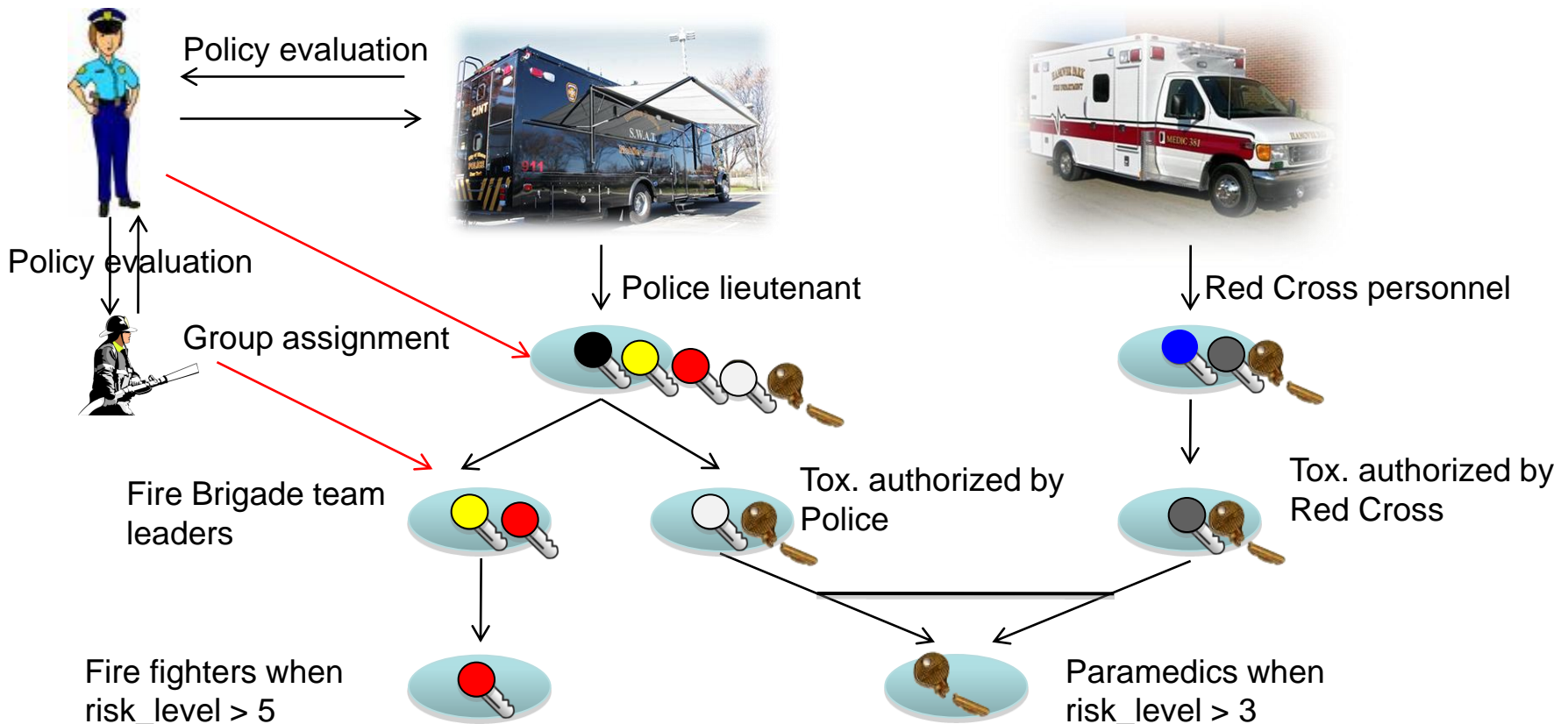
# Policies and Authorities

- Users can satisfy different policies when evaluated by a Trusted Authority, i.e. they can be assigned to different “authority groups”.
- Each authority group has rights over specific data categories.

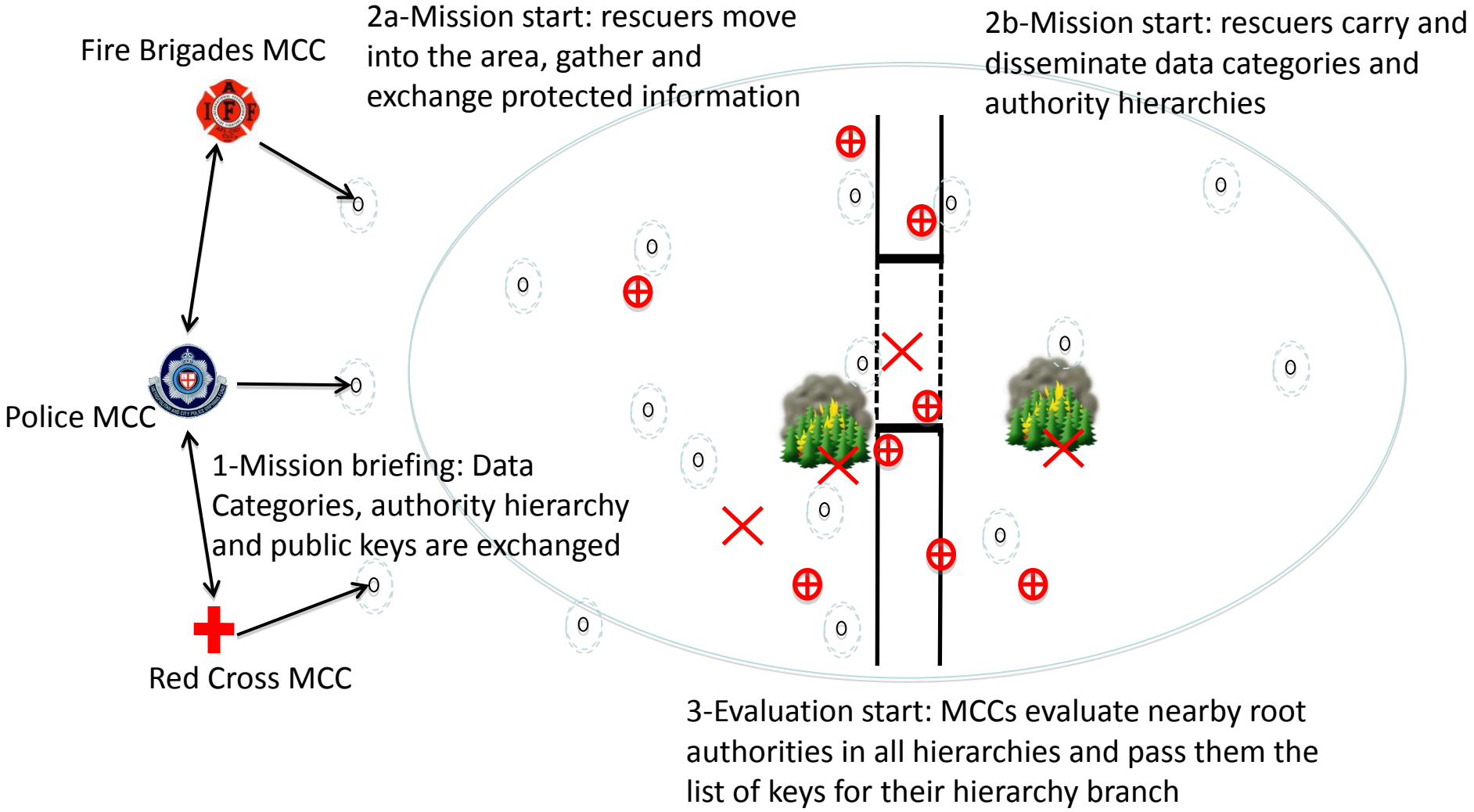


# A Hierarchy of Authorities

- Members of an authority group can act as evaluation authorities for underlying policies, i.e. they can substitute the organization's centralised trusted authority.
- Each organisation specifies its hierarchy including other organisations' rescuers

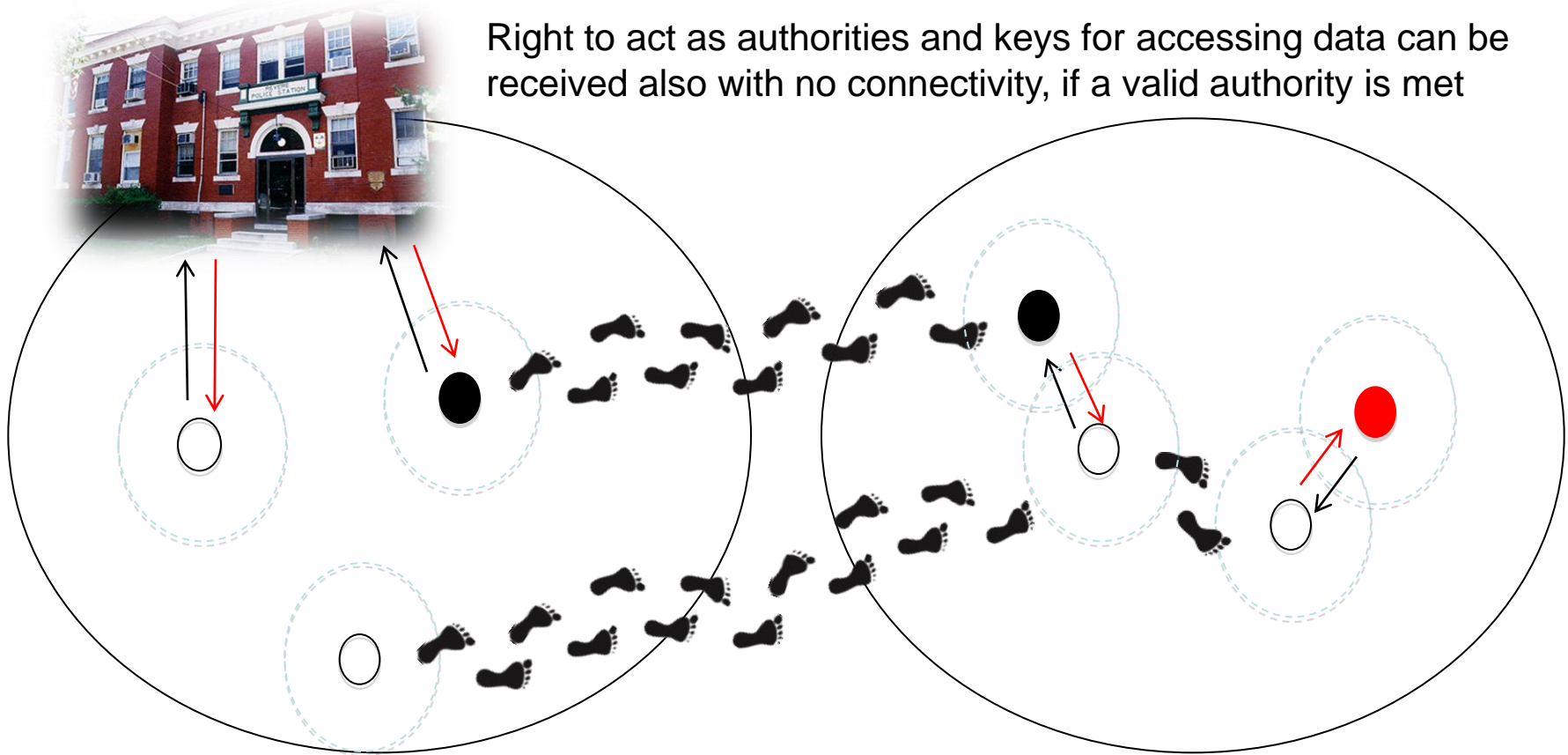


# Mission Set-up



# Opportunistic Evaluation

Right to act as authorities and keys for accessing data can be received also with no connectivity, if a valid authority is met



Connected area:  
traditional ERM policy evaluation  
infrastructure

Disconnected area:  
opportunistic policy evaluation  
whenever rescuers meet

# Is the overhead acceptable?



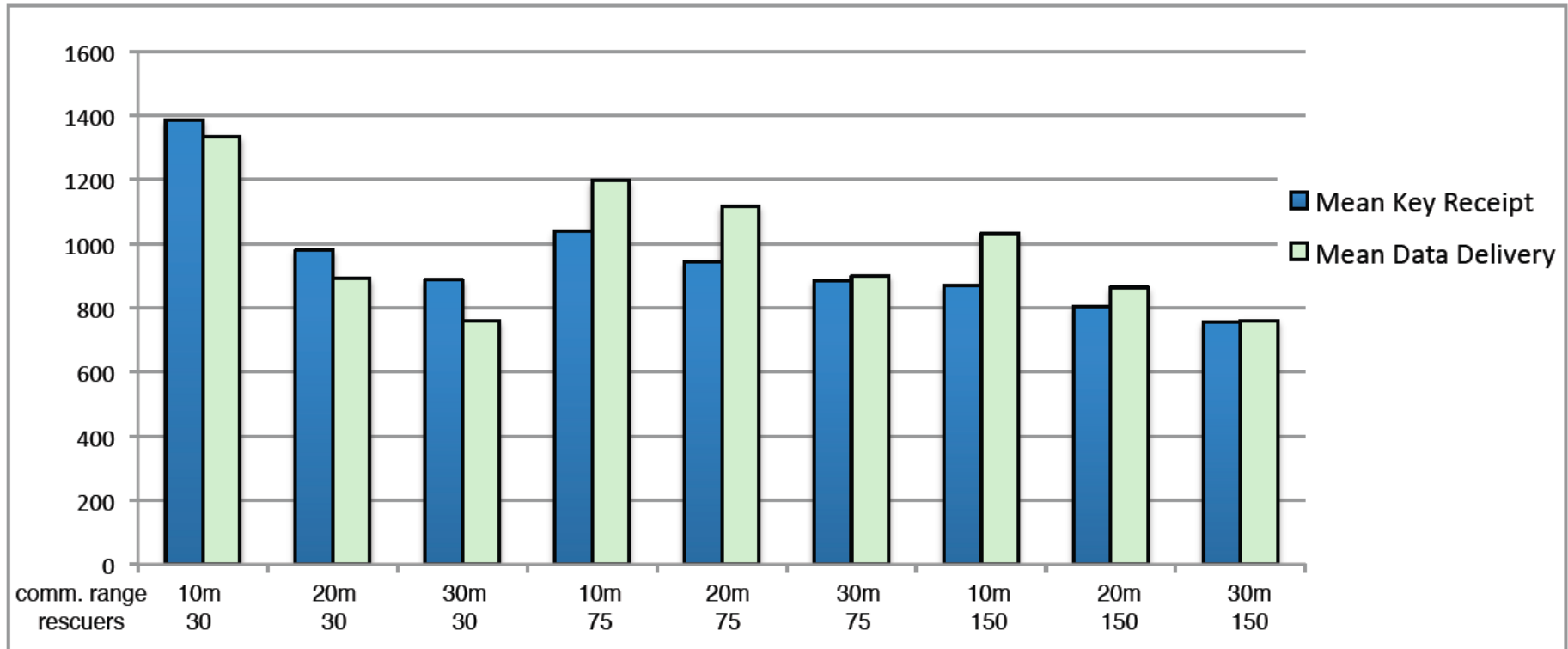
**Objective:** verify the protocol's overhead does not hinder data availability, i.e. either protected data is received after the key needed to access it or with a negligible delay.

**Simulations:** 30, 75 and 150 rescuers with varying communication ranges

- Fire Fighters move in team of 3 supervised by a Team Leader.
- Peers initially move from the respective Mobile Command Centres to the accident location.
- Random way point mobility model: search, stop (investigate, rescue, etc)
- Stop time 120 secs
- Speed: 1.1 m/s = human fast walking speed

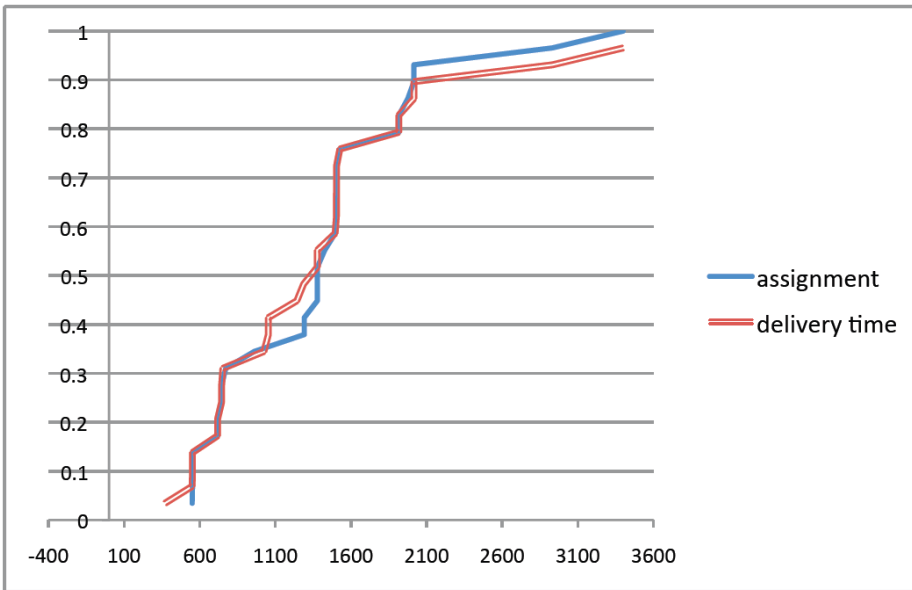
Random way point is not realistic but is the worst-case scenario. Forced routes like streets and obstacles would increase the number of meetings between peers.

# Results 1/2

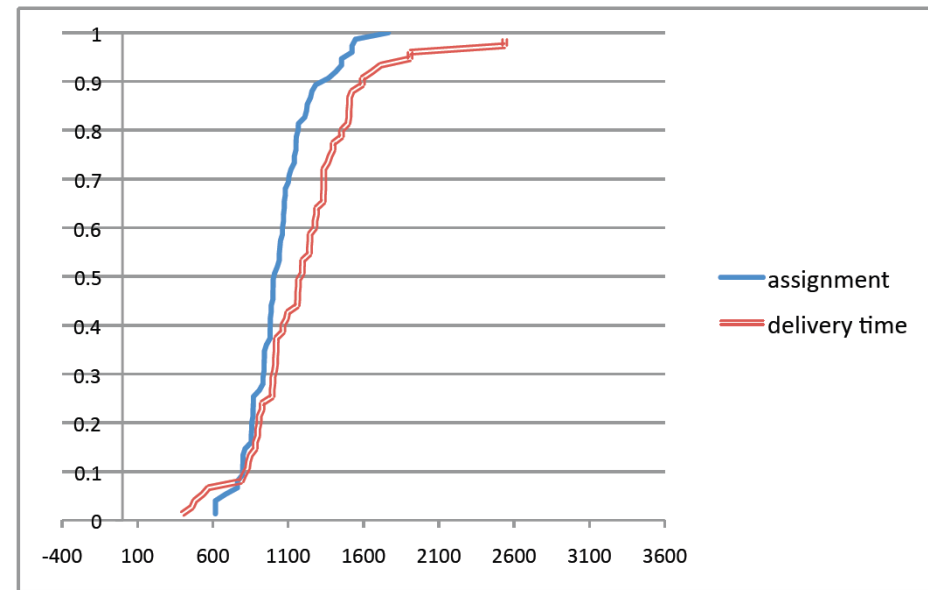


The more rescuers move in the area, the faster opportunistic evaluation is w.r.t. data dissemination

# Results 2/2



30 peers, 10m comm. range

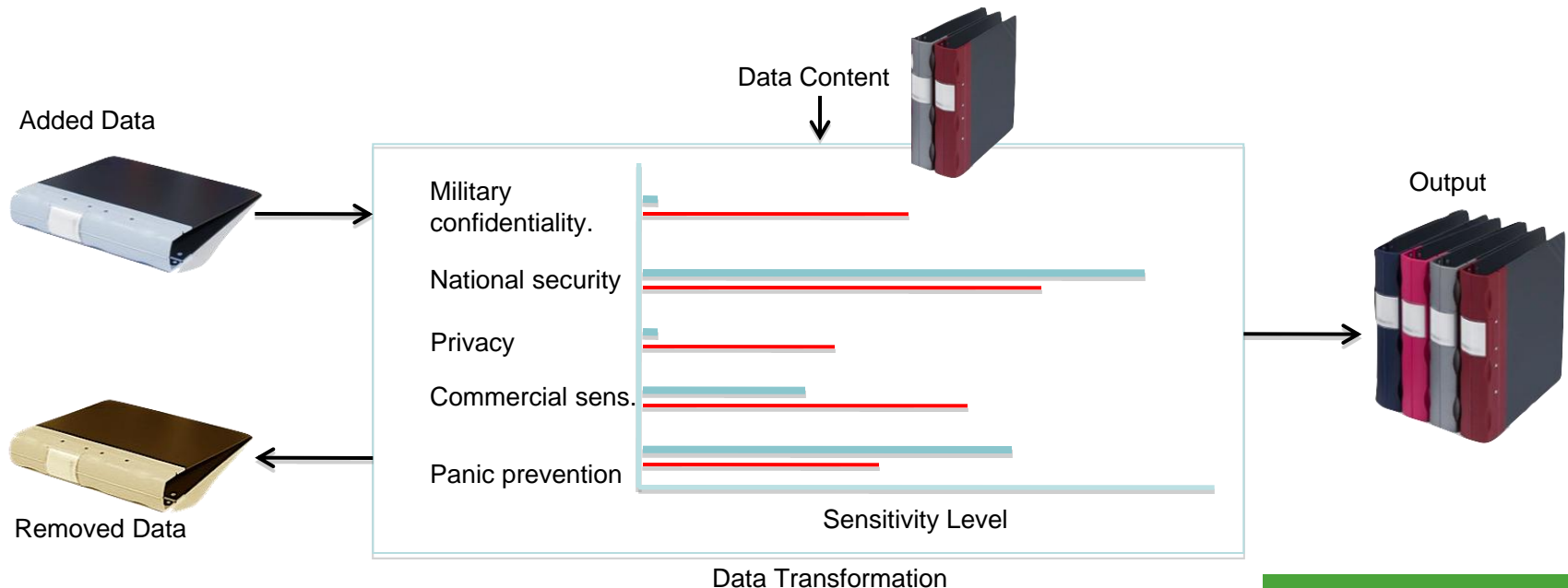


150 peers, 10m comm. range

- Key distribution is faster than data dissemination. Protocol does not compromise data availability and timeliness
  - Policies are evaluated starting from all MCCs, while data is disseminated starting from only one position
  - Group movements favour policy evaluation
- Simulation with simpler policy hierarchies show faster key distribution times

# Protecting Derived Data

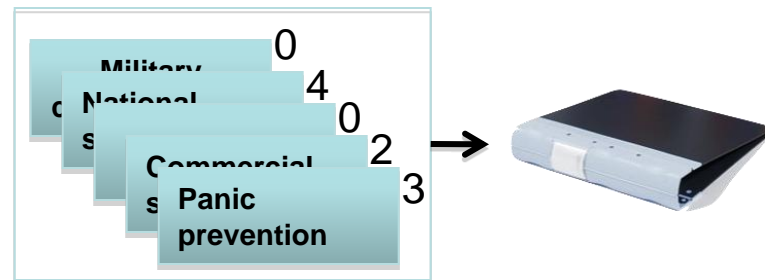
- Data must be protected for different reasons (purposes): confidentiality, integrity, privacy, commercial sensitivity, panic prevention – *protection domain*
- Each domain will have different *sensitivity levels* identifying how strictly must it be protected for that reason?;
- When data is transformed, sensitive content is removed or added for each security domain, increasing or decreasing the protection requirements for that domain.



# Data Labels and Tags

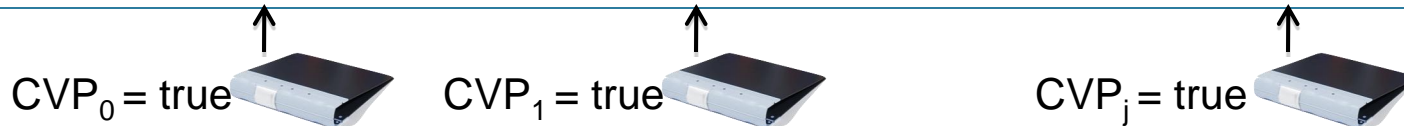
- Each protection domain is identified by a tag that specifies a range of discrete sensitivity levels.
- Data is associated with **labels** containing the sensitivity level for each tag.

$\{(\text{tag}_1, \text{value}_1), (\text{tag}_2, \text{value}_2) \dots (\text{tag}_n, \text{value}_n)\}$



- Each sensitivity level in a tag is associated with a *Content Verification Procedure (CVP)*, used to determine the initial levels for newly created data by analysing the data content e.g. facial recognition, xml analysis etc..

Tag: \*  $\rightarrow (0, \text{CVP}_0) \rightarrow (1, \text{CVP}_1) \rightarrow \dots \rightarrow (i, \text{CVP}_i) \rightarrow \dots \rightarrow (n, \text{CVP}_n)$



- \* level indicates that the security domain is not applicable to the data

# Example: Data and Tags



- XML documents containing:
  - Victims' information;
  - Information on the accident;
  - Information on the nearby hospitals (free beds and treatment facilities).
- CVPs as XPath expressions (possibly calling external java functions).
- Data Transformations as XQuery expressions (possibly calling external java functions).

**Privacy** \* → (0, true) → (1, NamesOrAddresses()) → (2, MedInfo ())

**confidentiality** \* → (0, true) → (1, Req(1)) → (2, Req(2)) → (3, Req(3) OR isImage())

**VideoPrivacy** \* → (0, true) → (1, FaceRecognition())

**Media** \* → (0, true) → (1, Victims())

Every XML element returned by a CVP is associated with the CVP's label:

```
<CVP id="MedInfo">
  <LABEL label="(privacy 2)"/>
  <XPATH>//CONDITIONS | //MEDICAL_HISTORY</XPATH>
</CVP>
<CVP id="NamesOrAddresses">
  <LABEL label="(privacy 1)"/>
  <XPATH>//VICTIM[NAME or SURNAME or ADDRESS]</XPATH>
</CVP>
<CVP id="FaceRecognition">
  <LABEL label="(videoPrivacy 1)"/>
  <XPATH>//child::*[custom:FaceRecognition(//child::*)]</XPATH>
</CVP>
<CVP id="Req(1,confidentiality)">
  <LABEL label="(confidentiality 1)"/>
  <XPATH>/child::*[custom:Req("1","confidentiality",/child::*)]</XPATH>
</CVP>
.....
```

# Sensitivity Levels are associated with nodes



```
<VICTIMS>
  <VICTIM status="alive" label="(privacy 1)">
    <NAME>Jane</NAME>
    ....
    <PHOTO label="(videoPrivacy 1)">
      <BLOB>.....</BLOB>
    </PHOTO>
    <CONDITIONS label="(privacy 2)">
      <SYMPTOMS>
        <SYMPTOM active="yes">
          <DESCRIPTION>head bleeding</DESCRIPTION>
          <FIRST-AID date="2000-01-12T12:33:00Z">bandaging</FIRST-
AID>
          <SUGGESTEDTHERAPY>apply
stitches</SUGGESTEDTHERAPY>
        </SYMPTOM>
      </SYMPTOMS>
    </CONDITIONS>
  </VICTIM>
</VICTIMS>
```

# And propagate down



```
<VICTIMS>
  <VICTIM status="alive" label="(privacy 1)">
    <NAME label="(privacy 1)">Jane</NAME>
    ....
    <PHOTO label="(privacy 1, videoPrivacy 1)">
      <BLOB label="(privacy 1, videoPrivacy 1)">.....</BLOB>
    </PHOTO>
    <CONDITIONS label="(privacy 2)">
      <SYMPTOMS label="(privacy 2)">
        <SYMPTOM active="yes" label="(privacy 2)">
          <DESCRIPTION label="(privacy 2)">head
bleeding</DESCRIPTION>
          <FIRST-AID date="2000-01-12T12:33:00Z" label="(privacy
2)">bandaging</FIRST-AID>
          <SUGGESTEDTHERAPY label="(privacy 2)">apply
stitches</SUGGESTEDTHERAPY>
        </SYMPTOM>
      </SYMPTOMS>
    </CONDITIONS>
  </VICTIM>
</VICTIMS>
```

- Each level is associated with a set of policies that protect the document
- Policies are ordered according to the protection they provide
- Example policies for the two levels of the privacy tag:

$$P_{p,1}(q, read) = \begin{cases} N & zone \neq hospital \vee zone \neq emr. area \\ Y & role \geq Volunteer \vee role \geq Policeman \\ NA & otherwise \end{cases}$$

$$P_{p,2}(q, read) = \begin{cases} N & (zone \neq hospital \vee zone \neq emr. area) \vee (aid exp. = false) \\ Y & role \geq Volunteer \vee role \geq Policeman \\ NA & otherwise \end{cases}$$

- Example policies for the three levels of the confidentiality tag:

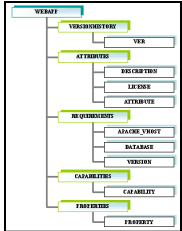
$$P_{c,1}(q, read) = \begin{cases} N & role \geq Policeman \wedge mission \neq emr1 \\ Y & role \geq Policeman \vee role > Volunteer \\ NA & otherwise \end{cases}$$

$$P_{c,2}(q, read) = \begin{cases} N & role \geq Policeman \wedge mission \neq emr1 \\ Y & role \geq Lieutenant \\ NA & otherwise \end{cases}$$

$$P_{c,3}(q, read) = \begin{cases} N & role \geq Policeman \wedge mission \neq emr1 \\ Y & role = Commander \\ NA & otherwise \end{cases}$$



- 1) An encrypted XML Document is received;
- 2) The publishing license and the recipient's credentials are sent to the publishing authority
- 3) The authority evaluates for each data label the corresponding policies:

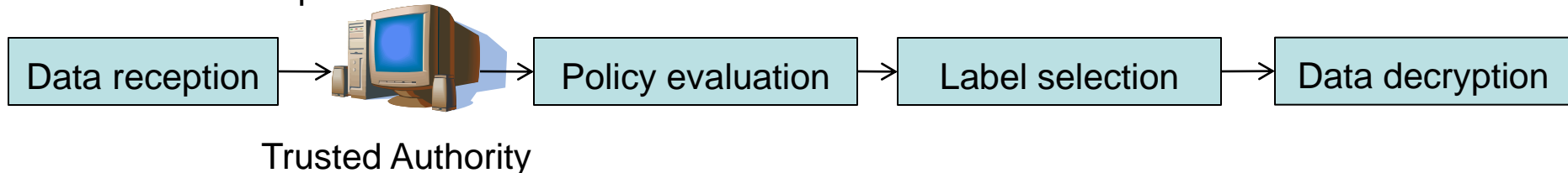


Label= {(privacy,1), (confidentiality,2)}

*Both policy p1 and c2 must be satisfied*

- 4) If all policies are satisfied, the key for the label is released to the recipient
- 5) The recipient can decrypt the parts of the document he is authorised to

## Access Request



- Data can be changed through a set of controlled *transformations*.
- Each transformation has a set of labels determining their effect on the data, ... , and thus allowing to determine the labels of the resulting data

## A transformation can:

- **Add sensitive information**
  - **Classification labels**  $L^c$  indicate the minimum sensitivity level of the output data's label.
- **Remove sensitive information**
  - **General declassification labels**  $L^g$  indicate the maximum sensitivity level of the output data after undergoing the transformation
- **Relatively change information sensitivity**
  - **Relative (de)classification labels**  $L^-$  indicate how the sensitive input data is changed with respect to its current level.
- **Change the information content in unpredictable ways**
  - **Decisional labels**  $L^t$  (values are booleans [true, false]) indicate whether the output sensitivity should be determined by CVPs

- Transformations are XQuery expressions that can generate complex XML documents.
- To label each XML element of the output differently, transformations are specified together with a set of XPath expressions. Each expression is associated with the four transformation labels and the input data that contributes to the creation of the returned XML elements.
- The output elements returned by an XPath expression are labelled as:

$$L_d(t) \leftarrow \begin{cases} \max(i) \mid CVP_{i,t} = true \text{ if } L^t(t) = true \\ L_d \leftarrow \sqcup ((L_{e_i} - L^-) \sqcap L^g) \sqcup L^c \quad otherwise \end{cases}$$

$$L_a - L_b(t) = \begin{cases} * & \text{if } L_a(t) = * \\ 0 & \text{if } L_a(t) \times L_b(t) < threshold \\ \lceil L_a(t) \times L_b(t) \rceil & otherwise \end{cases}$$

- i.e., the input data is first declassified and the upper bound of the declassified labels and classification label is returned as the derived data (output) label.
- The labelled output can then be published as new data.

# Example

Symptoms analysis service: checks if some victims suffer from either breathing problems or red eyes to verify whether there is a toxic contamination in the area.

```
<TOXIC_ANALYSIS date="{fn:current-dateTime()}"> {  
  if (every $victim in fn:doc("victims.xml")//VICTIM satisfies  
    (some $symptom in $victim/CONDITIONS/SYMPTOMS/SYMPTOM satisfies  
      $symptom/DESCRIPTION/text() = "breathing problems" or  
      $symptom/DESCRIPTION/text() = "red eyes"))  
  
  then "toxic contamination in the area!"  
  else "no contamination!"  
} </TOXIC_ANALYSIS>
```

```
<LABELS>  
  <LABEL clabel="(confidentiality 1)" glabel="(privacy 0)">  
    <XPATH>//TOXIC_ANALYSIS</XPATH>  
    <INPUTDATA file="victims" />  
  </LABEL>  
</LABELS>
```

- Hierarchy of policies permits devolution of authority for evaluating policies.
- This enables us to implement context aware usage control in environments with intermittent connectivity and opportunistic data disseminations.
- It facilitates the devolution of authority across organisations controlled by pre-defined policies.
- The delay introduced is acceptable and disappears with higher density of users.
  
- Data Labels are used to specify data protection requirements
- Classification and Declassification Labels are used to specify how transformations modify data and its protection requirements
- Tags need to be specified only once at system setup
- Utility increases with complex scenarios where several transformations are continuously applied in a cyclic way on data (e.g. research and testing environments)