CHECKSEM
SEMANTICINTELLIGENCERESEARCH

UCD DUBLIN

DFRWS
DIGITAL FORENSIC RESEARCH WORKSHOP

August 5, 2014

A Complete Formalized Knowledge Representation Model for Advanced Digital Forensics Timeline Analysis

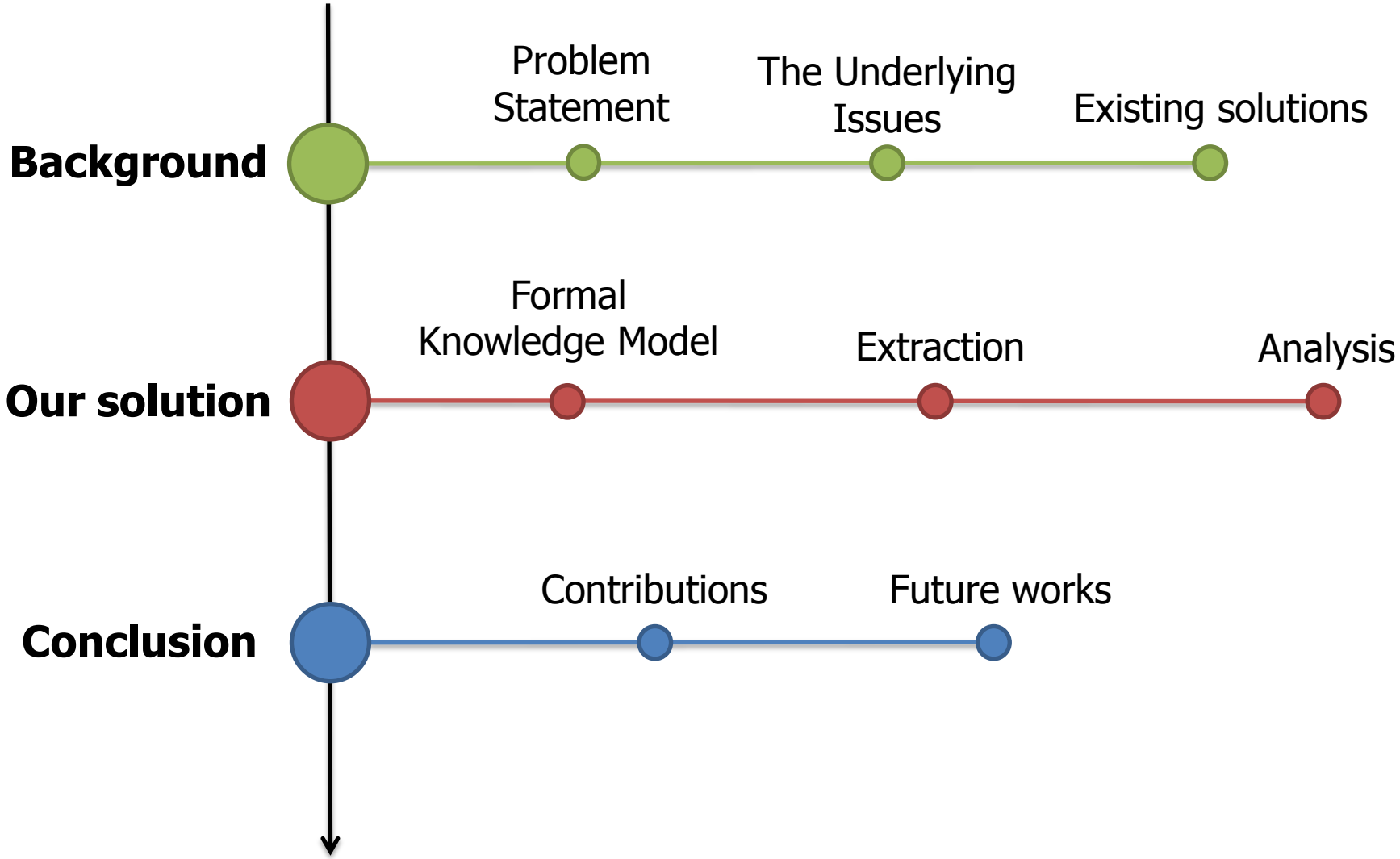
Yoan Chabot[a,b], Aurélie Bertaux[a], Christophe Nicolle[a] and M-Tahar Kechadi[b]

yoan.chabot@checksem.fr

[a] CheckSem Team, Laboratoire Le2i, Université de Bourgogne, Dijon, FRANCE
[b] School of Computer Science & Informatics, University College Dublin, IRELAND

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2i – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

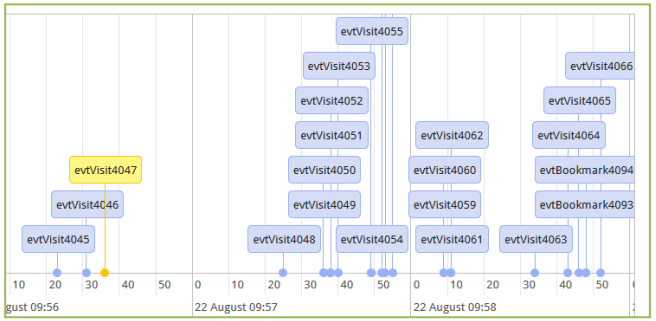
Outline



Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Event Reconstruction

Footprints



Timeline

Conclusions of the investigation

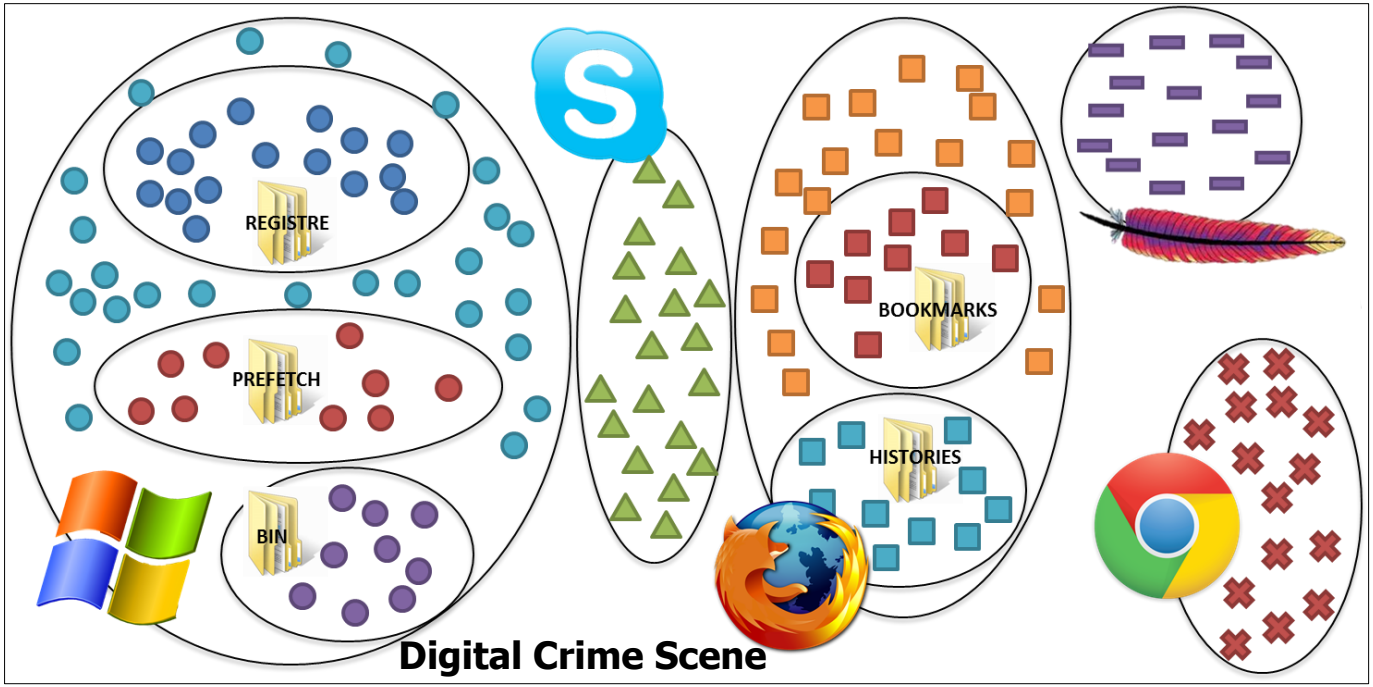
GOAL: Determine the circumstances of the incident

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

The Underlying Issues

Technical gaps

- Large amount of data
- Heterogeneity (Semantic, Format, Time)



Legal requirements

- Credibility
- Veracity
- Precision
- Reproducibility

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Existing Solutions

ECF, FORE, Finite state machine approach, Zeitline, Neural networks appraoch, CyberForensic TimeLab, etc.

log2timeline by Kristinn Gudjonsson

Super-timelines using a large number of sources

- Windows Event Logs
- Web Browsers Histories
- Apache logs
- PDF document metadata
- Firewall logs
- etc.

```
Yoan@Checksem-PC /cygdrive/j/Local Workspace/plaso
$ ./log2timeline ../output.dump ../Scenarios/scenario1/EnCase/scenario1.E01 > log.txt

Yoan@Checksem-PC /cygdrive/j/Local Workspace/plaso
$ ./psort -w ../timeline.txt ../output.dump > log.txt
```

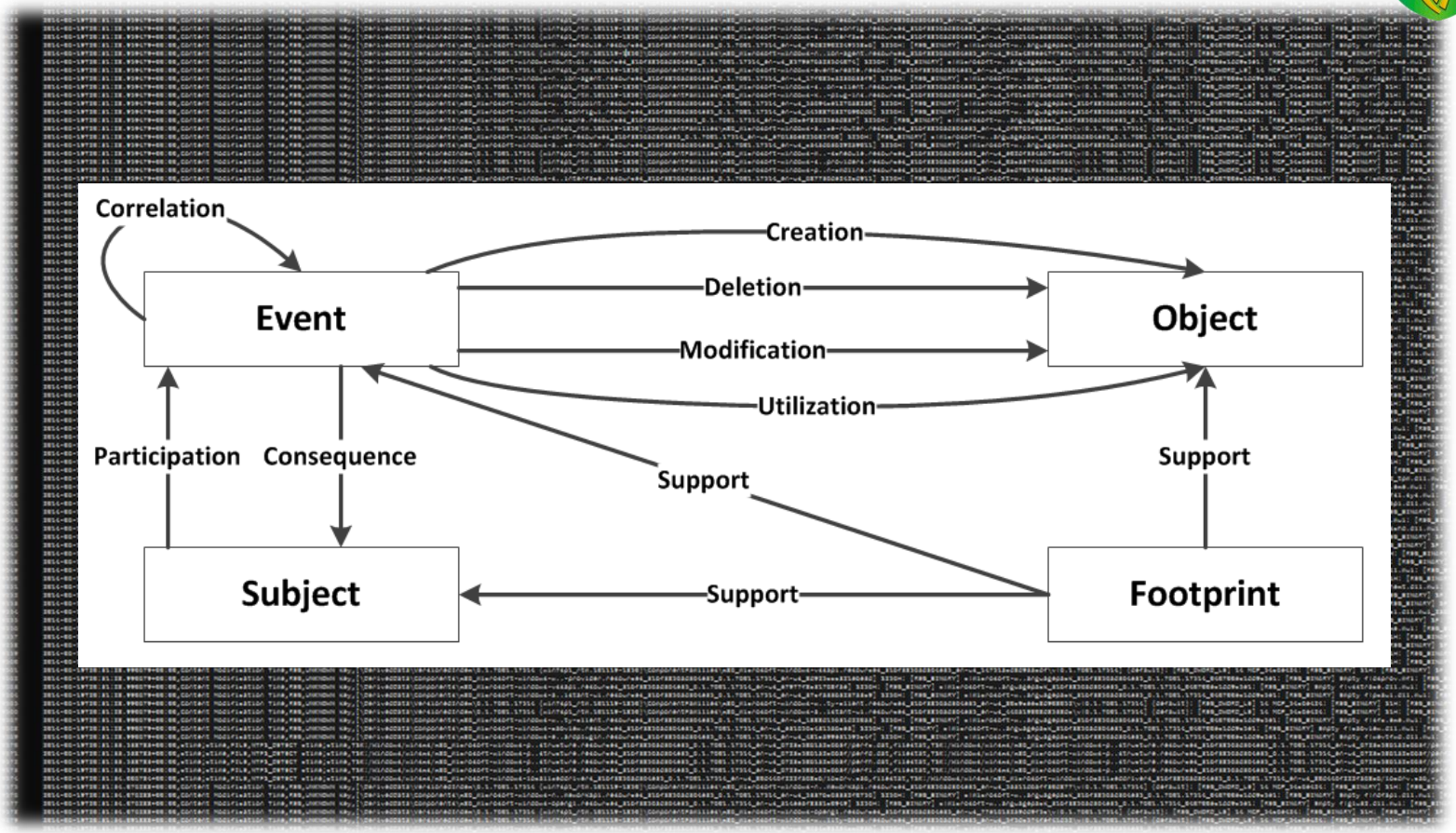
Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Knowledge Model

How to analyze this large amount of data?

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

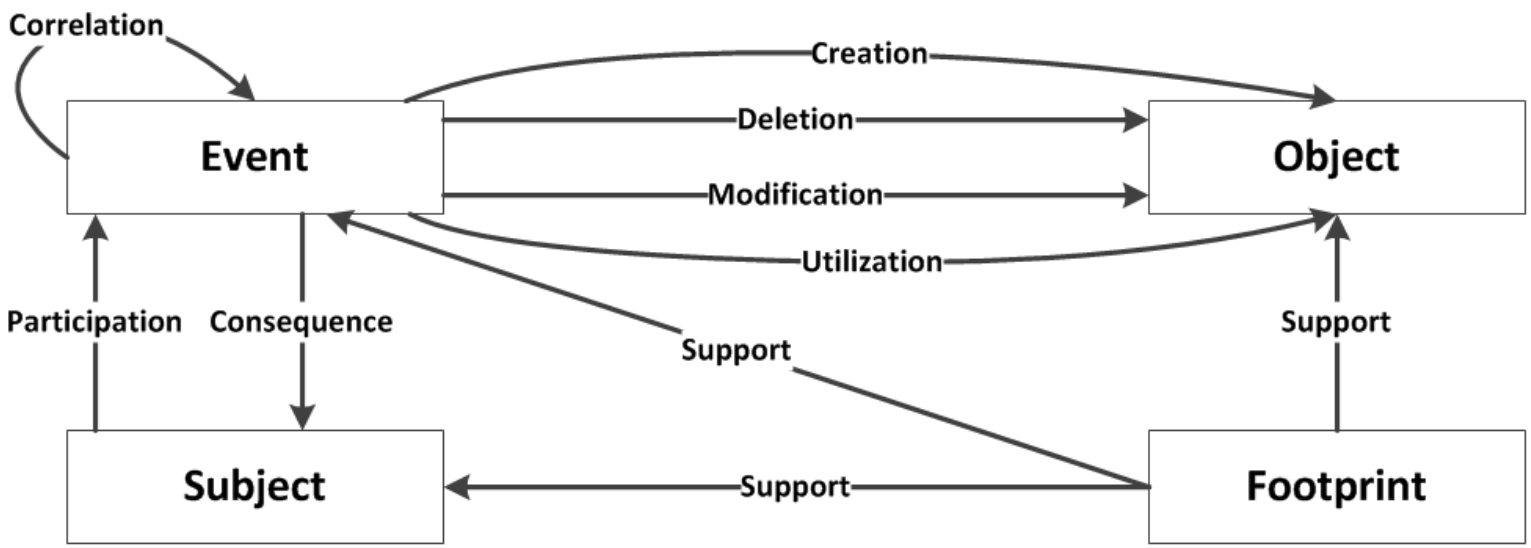
Knowledge Model



Introduce a semantically rich knowledge representation of events to enhance analysis capabilities

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Knowledge Model

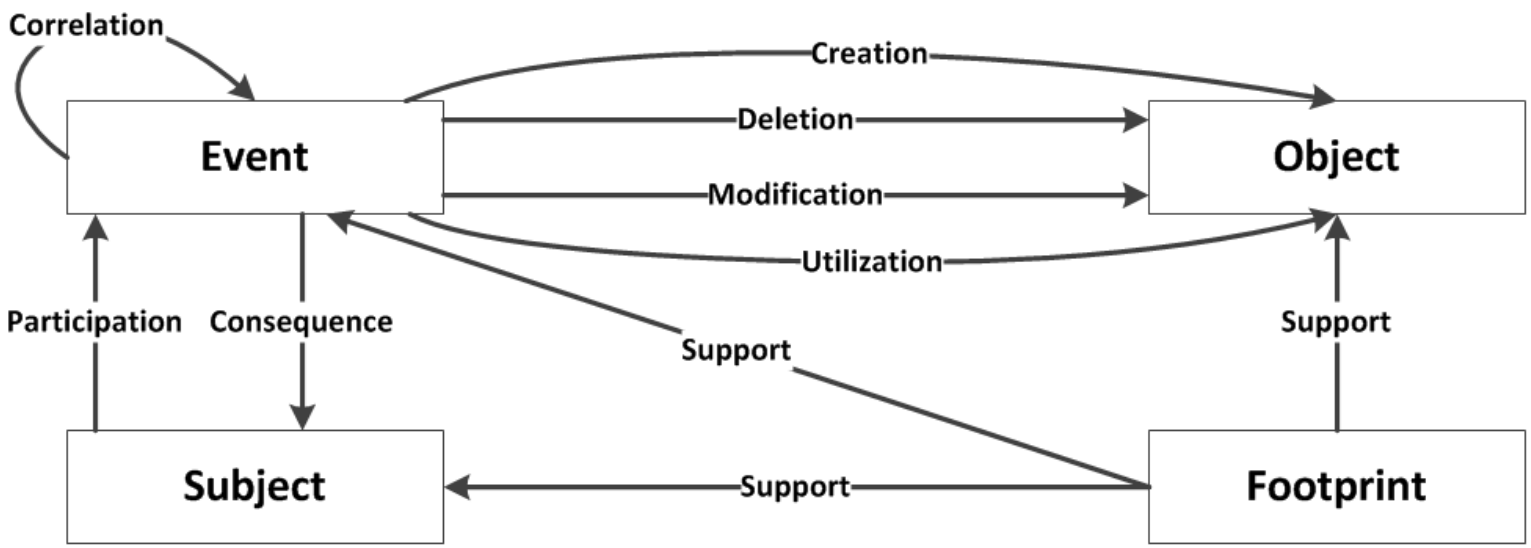


Entities

- $s \in S = \{a \in A_s \mid s\ \alpha_s\ a\}$
- $o \in O = \{a \in A_o \mid x\ \alpha_o\ a\}$
- $O \subseteq \wp(A_o)$
- $f \in F = \{f \in A_f \mid x\ \alpha_f\ a\}$
- $e \in E = \{t_{start}, t_{end}, l, S_e, O_e, E_e\}$
- $S_e = \{s \in S \mid e \in E, s\ \sigma_s\ e\}$
- $O_e = \{o \in O \mid e \in E, e\ \sigma_o\ o\}$
- $E_e = \{x \in E \mid e \in E, e\ \sigma_e\ x\}$

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Knowledge Model



Relationships

- $\sigma_s = \{participation, repercussion\}$
- $\sigma_o = \{creation, suppression, modification, utilization\}$
- $\sigma_e = \{correlation\}$
- $\sigma_f = \{support\}$
- $support(en \in \{E \times O \times S\}) = \{f \in F \mid f\ \sigma_f\ en\}$

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2i – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Operators

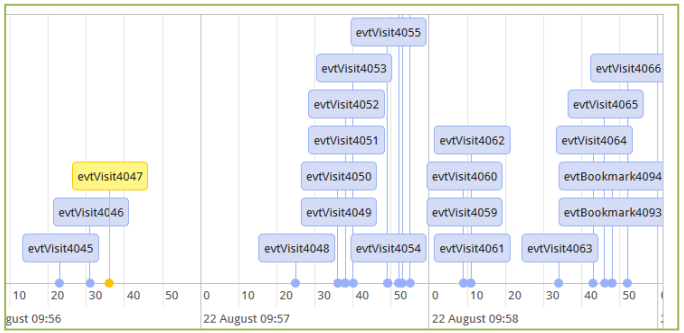
Footprints left on the crime scene

Extraction and Mapping Operators

Knowledge Base

Inference Operators

Analysis Operators



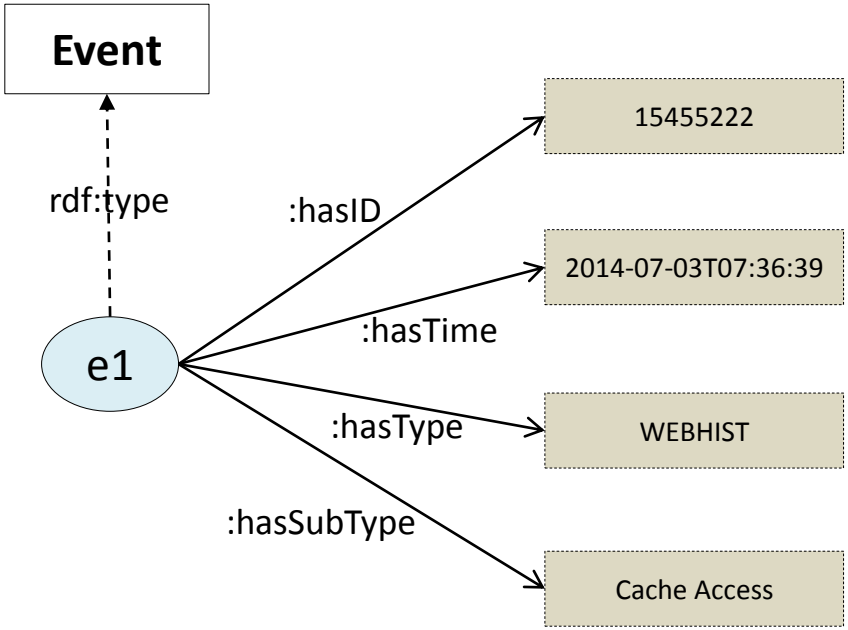
Enhanced Timeline

Conclusions

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

From Raw Data To Giant Graphs

2014-07-03T07:36:39.408000+00:00,Last Visited Time,WEBHIST,MSIE Cache File URL record,Location: Visited: Yoan@file:///C:/Users/Yoan/Pictures/dfrws12-039.jpg Number of hits: 2 Cached file size: 0,msiecf,TSK:/Users/Yoan/AppData/Local/Microsoft/Windows/History/History.IE5/index.dat,-,3,378480

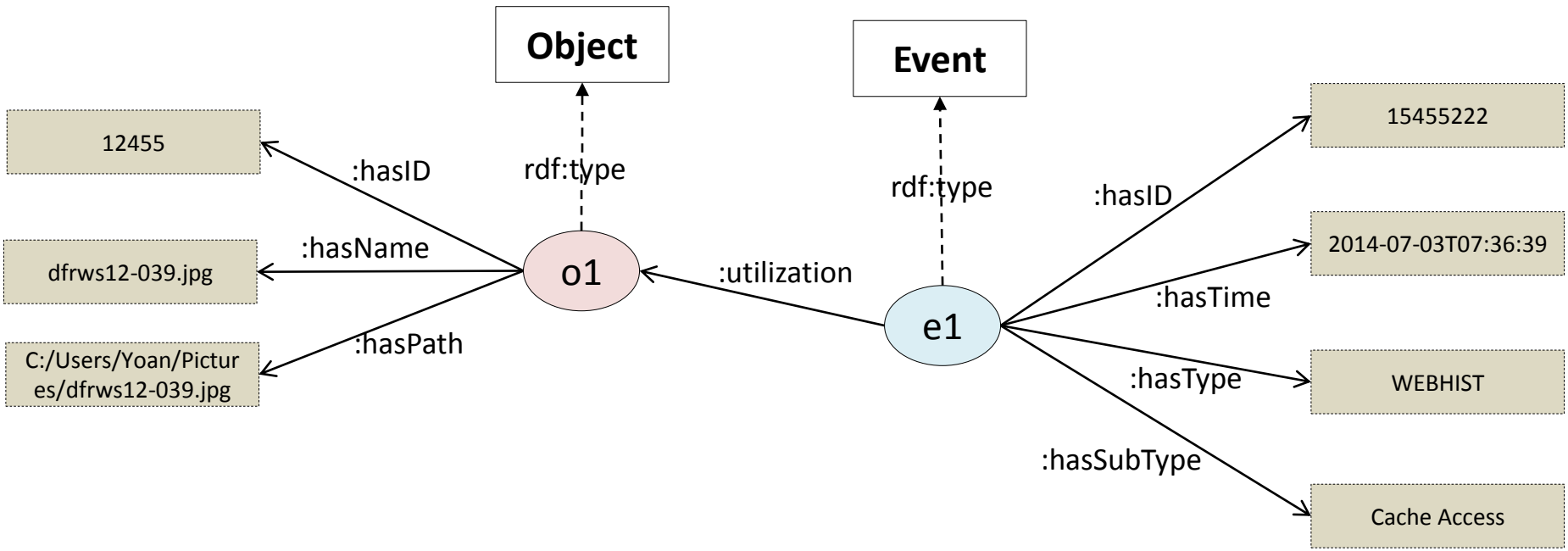


Mapping: 1st Example

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, Fran

From Raw Data To Giant Graphs

2014-07-03T07:36:39.408000+00:00,Last Visited Time,WEBHIST,MSIE Cache File URL record,Location: Visited: Yoan@file:///C:/Users/Yoan/Pictures/dfrws12-039.jpg Number of hits: 2 Cached file size: 0,msiecf,TSK:/Users/Yoan/AppData/Local/Microsoft/Windows/History/History.IE5/index.dat,-,3,378480

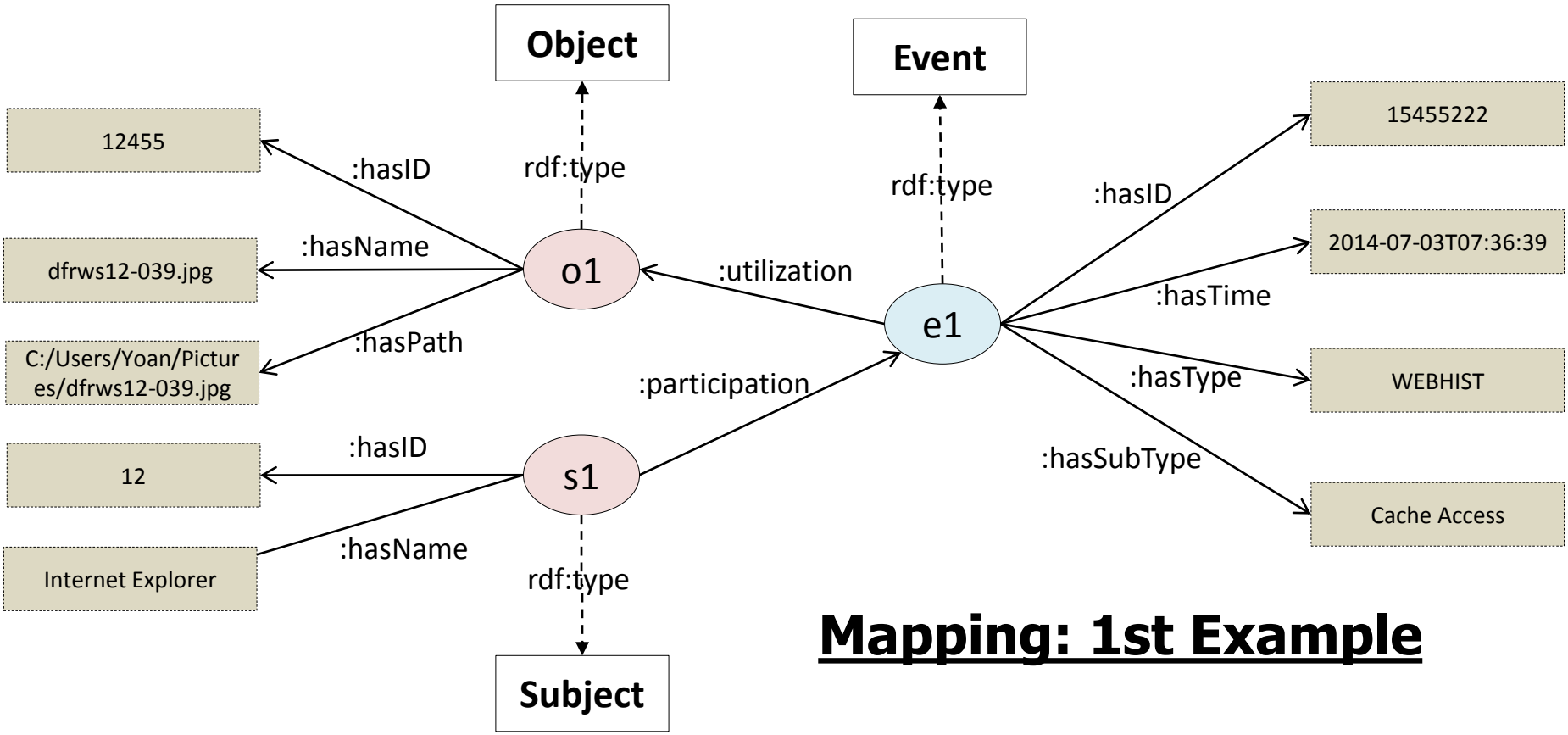


Mapping: 1st Example

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, Fran

From Raw Data To Giant Graphs

2014-07-03T07:36:39.408000+00:00,Last Visited Time,WEBHIST,MSIE Cache File URL record,Location: Visited: Yoan@file:///C:/Users/Yoan/Pictures/dfrws12-039.jpg Number of hits: 2 Cached file size: 0,msiecf,TSK:/Users/Yoan/AppData/Local/Microsoft/Windows/History/History.IE5/index.dat,-,3,378480

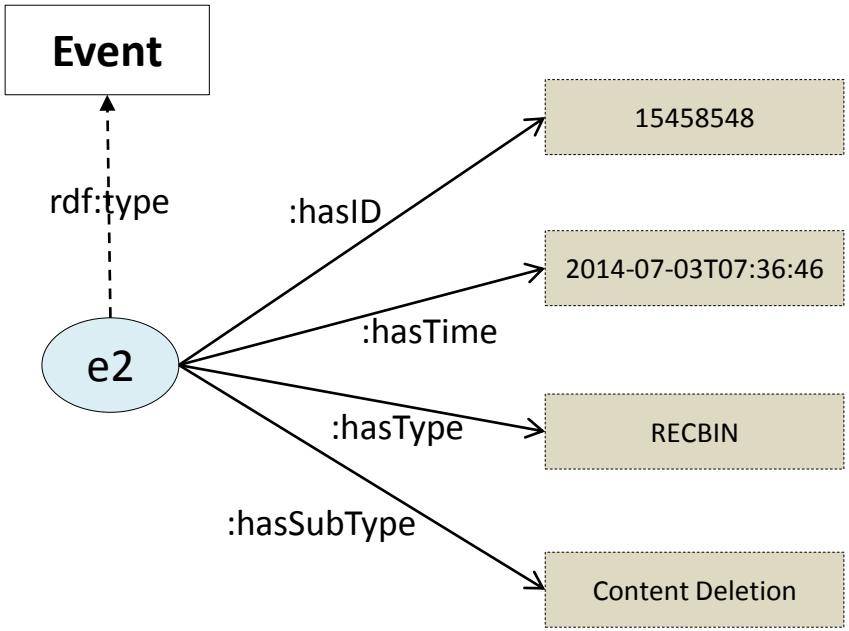


Mapping: 1st Example

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, Fran

From Raw Data To Giant Graphs

2014-07-03T07:36:46.662000+00:00,Content Deletion Time,RECBIN,Recycle Bin,C:\Users\Yoan\Pictures\dfrws12-039.jpg,recycle_bin,TSK:/$Recycle.Bin/S-1-5-21-3724914695-4089496160-3424763353-1000/$IAXNK4E.jpg,-,3,378521

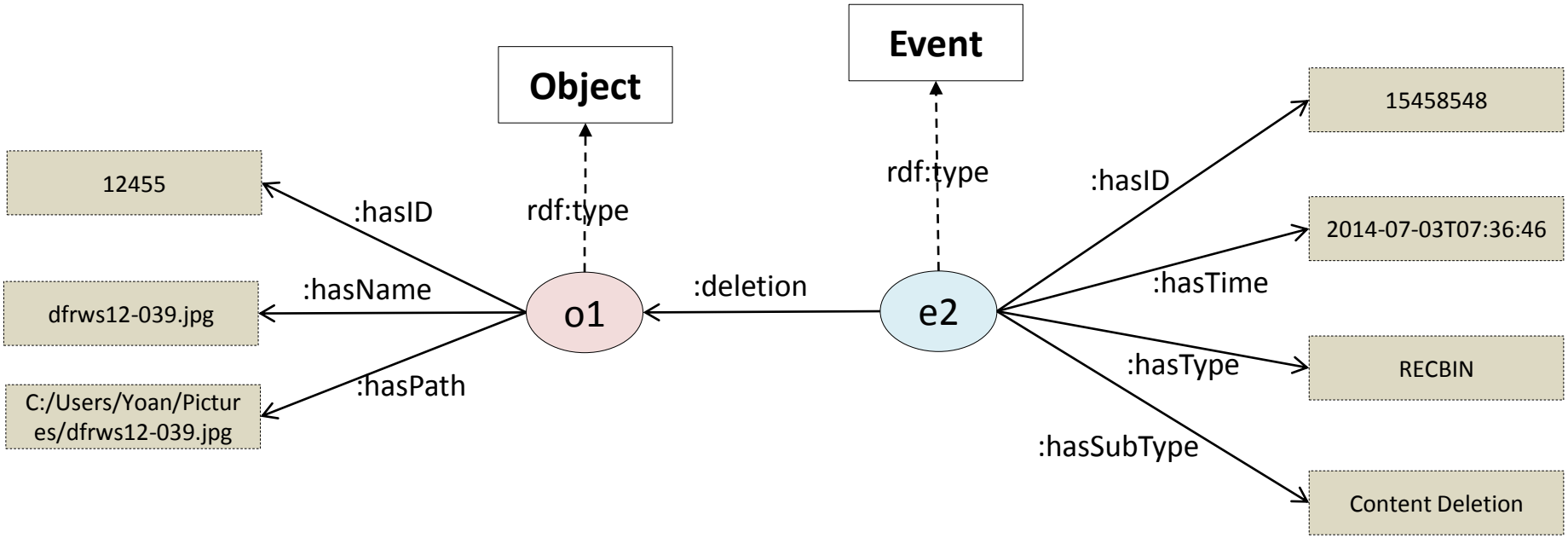


Mapping: 2nd Example

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informat
(LE2i – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex,

From Raw Data To Giant Graphs

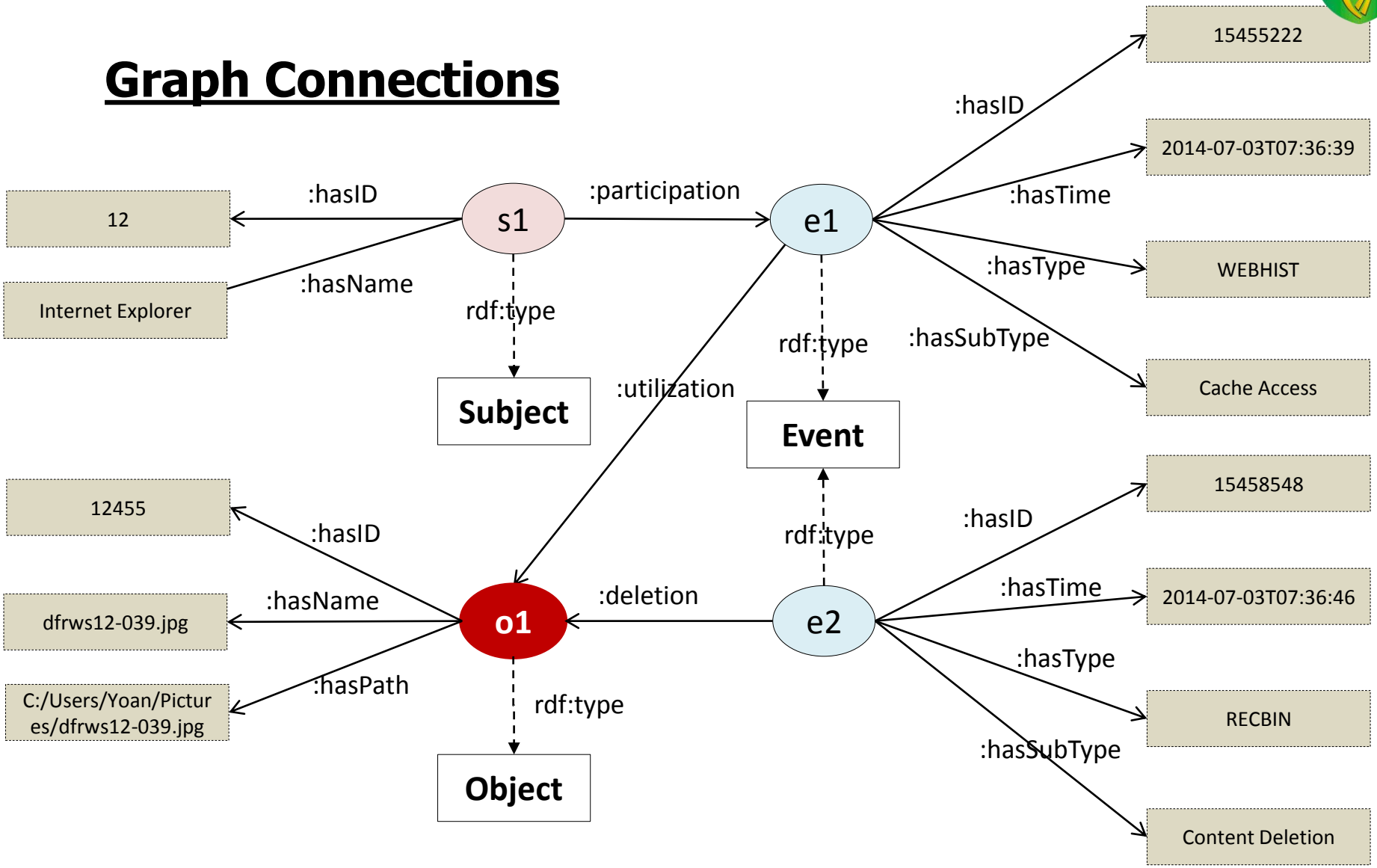
2014-07-03T07:36:46.662000+00:00,Content Deletion Time,RECBIN,Recycle Bin,C:\Users\Yoan\Pictures\dfrws12-039.jpg,recycle_bin,TSK:/$Recycle.Bin/S-1-5-21-3724914695-4089496160-3424763353-1000/$IAXNK4E.jpg,-,3,378521



Mapping: 2nd Example

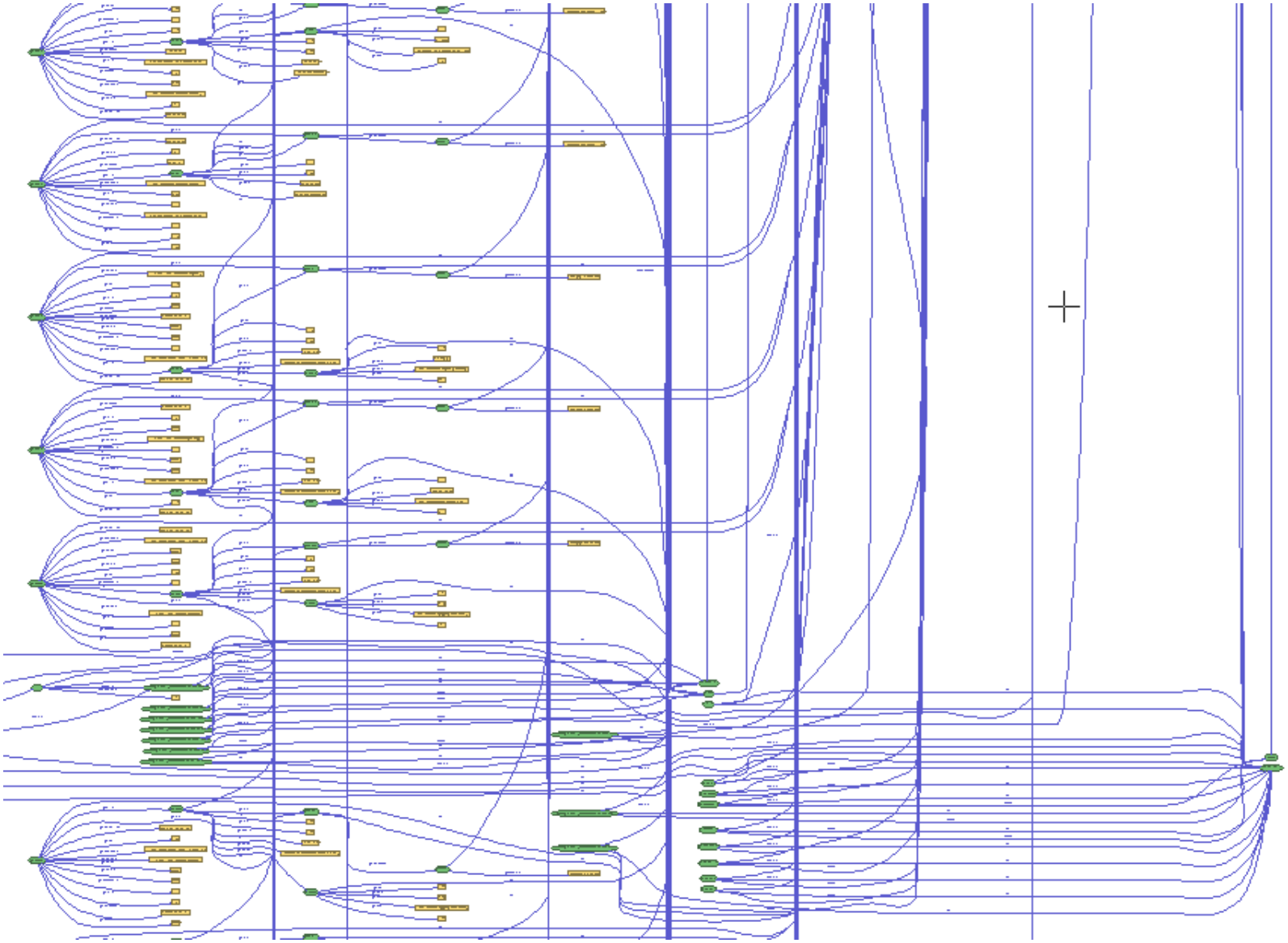
Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informat
(LE2i – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex,

Graph Connections



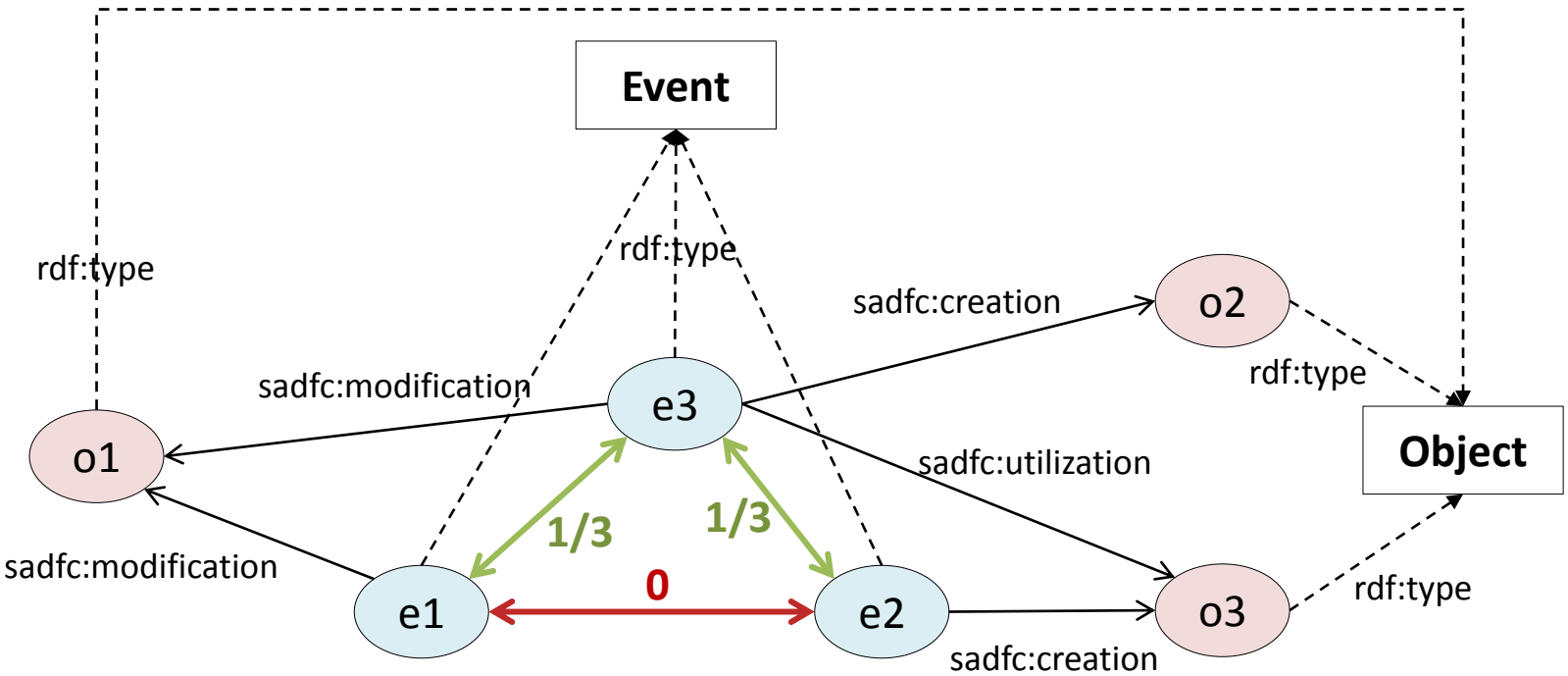
Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

From Raw Data To Giant Graphs



Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Automatic Analysis Operators

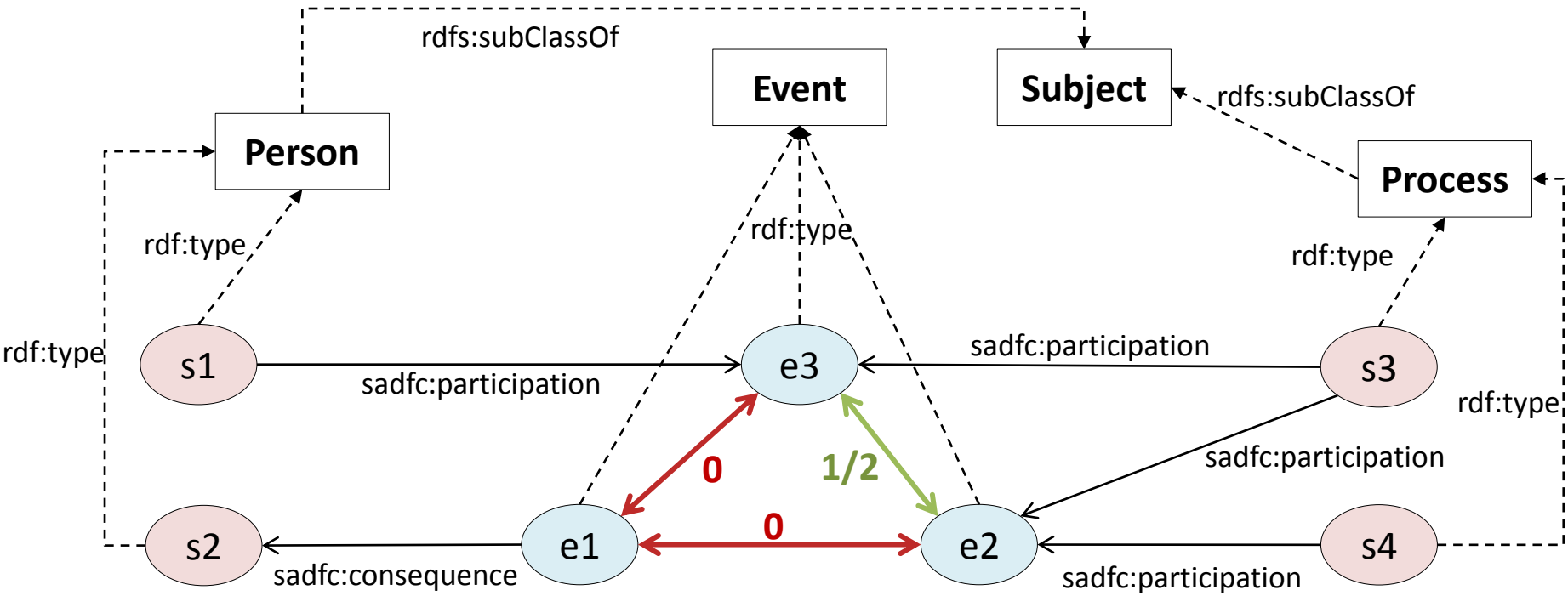


$$Correlation_O(e,x) = |O_e \cap O_x| / \max(|O_e|, |O_x|)$$

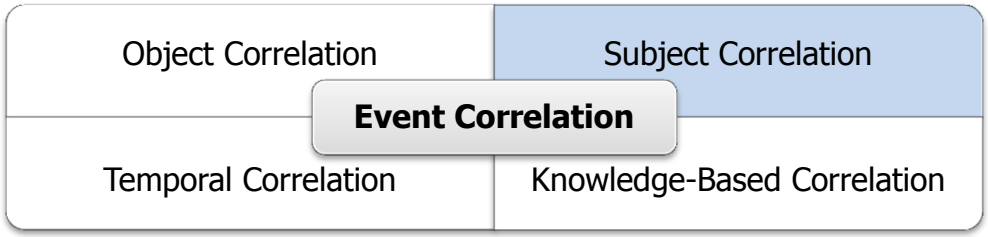
Object Correlation	Subject Correlation
Event Correlation	
Temporal Correlation	Knowledge-Based Correlation

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Automatic Analysis Operators

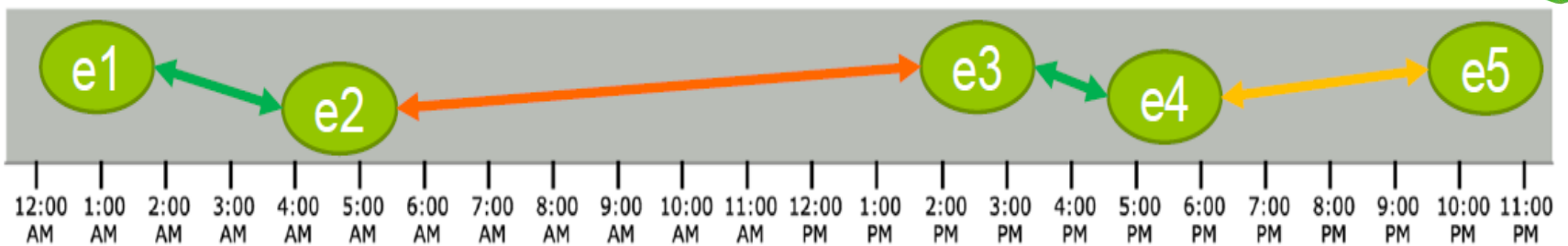


$$Correlation_S(e,x) = |S_e \cap S_x| / \max(|S_e|, |S_x|)$$



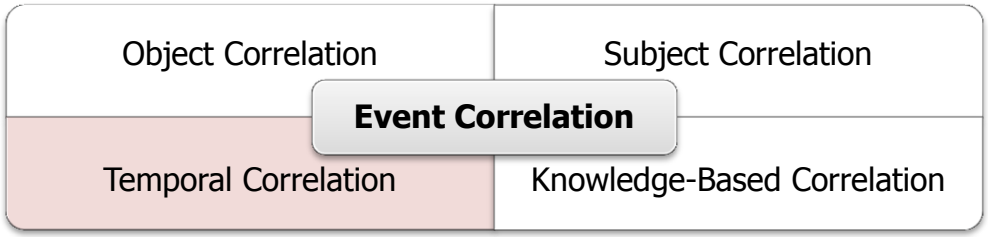
Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Automatic Analysis Operators



$$Correlation_T(e,x) = \alpha \times starts(e,x) + \alpha \times equals(e,x) + meets(e,x) + overlaps(e,x) + during(e,x) + finishes(e,x) + before(e,x)$$

Functions	Constraints	Example
before(X,Y)	$x_{t_{end}} < y_{t_{start}}$	X Y
equal(X,Y)	$x_{t_{start}} = y_{t_{start}}$ && $x_{t_{end}} = y_{t_{end}}$	X Y
meets(X,Y)	$x_{t_{end}} = y_{t_{start}}$	X Y
overlaps(X,Y)	$x_{t_{start}} < y_{t_{start}}$ && $x_{t_{end}} > y_{t_{start}}$	X Y
during(X,Y)	$x_{t_{start}} > y_{t_{start}}$ && $x_{t_{end}} < y_{t_{end}}$	X Y
starts(X,Y)	$x_{t_{start}} = y_{t_{start}}$	X Y
finishes(X,Y)	$x_{t_{end}} = y_{t_{end}}$	X Y



Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

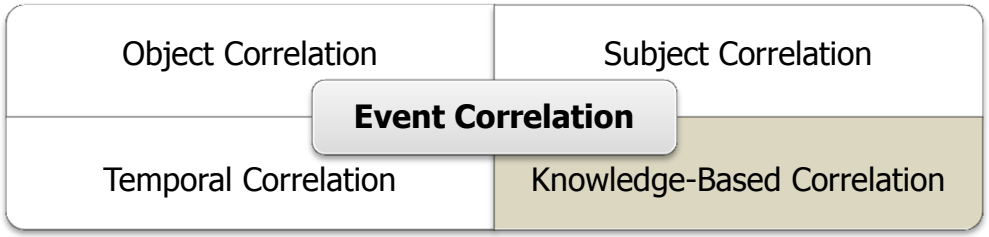
Automatic Analysis Operators

2014-06-20T13:57:16.544000+00:00 | Creation Time | WEBHIST | Firefox History | Bookmark URL CheckSem - Semantic Intelligence Research (http://checksem.u-bourgogne.fr/www/)| sqlite | TSK:/Users/Yoan/AppData/Roaming/Mozilla/Firefox/Profiles/94zxtt2a.default/places.sqlite | - | 3 | 373176 |

2014-06-20T13:57:21.474000+00:00 | Page Visited | WEBHIST | Firefox History | http://checksem.u-bourgogne.fr/www/ (CheckSem - Semantic Intelligence Research Host: checksem.u-bourgogne.fr visited from: http://checksem.u-bourgogne.fr/www/ (checksem.u-bourgogne.fr) Transition: BOOKMARK | sqlite | TSK:/Users/Yoan/AppData/Roaming/Mozilla/Firefox/Profiles/94zxtt2a.default/places.sqlite | - | 3 | 373182 |

$$Correlation_{KBR}(e,x) = \sum_{r=1}^{n} rule_r(e,x)$$

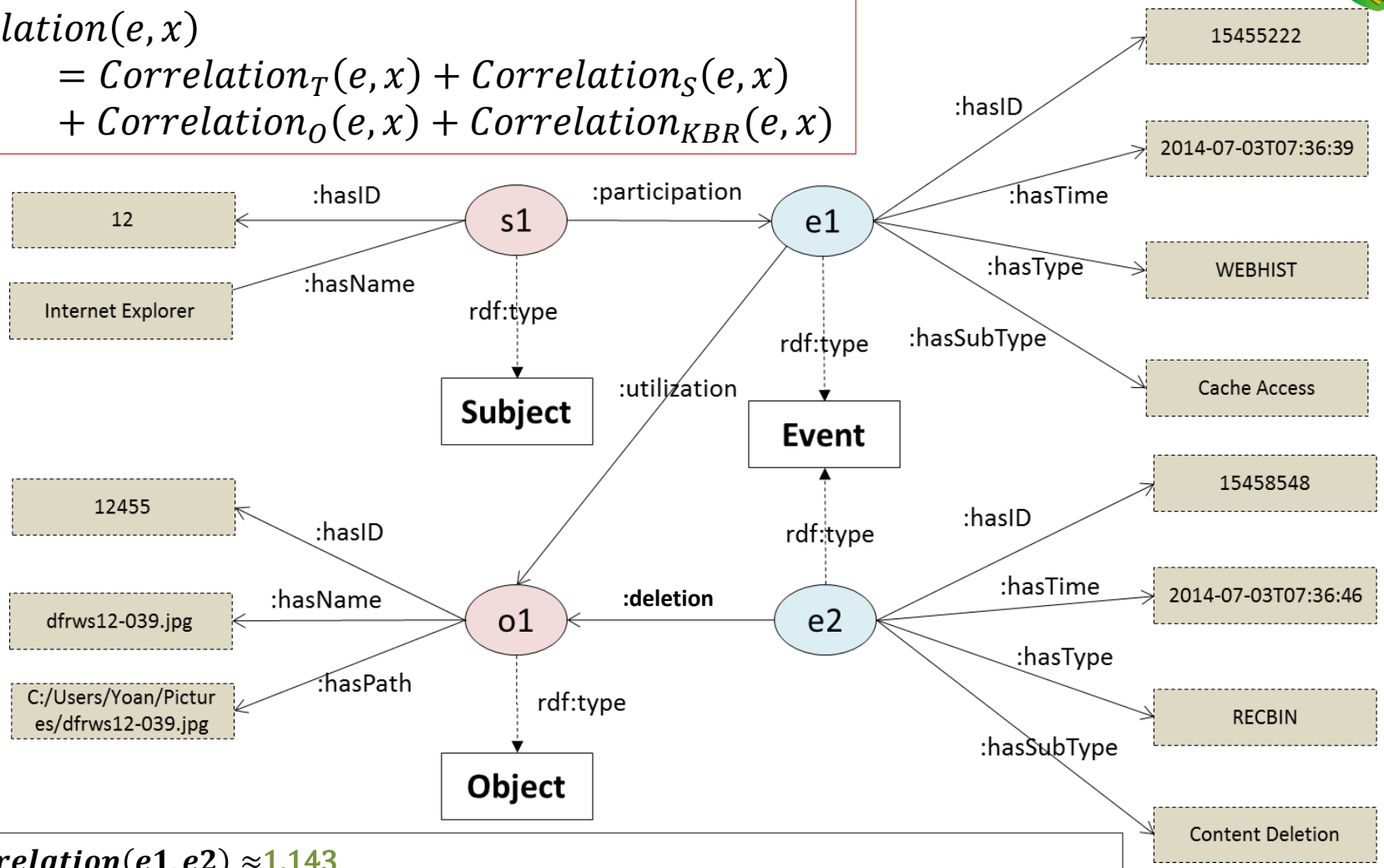
With $rule_r(e,x) = 1$ if the rule is satisfied and 0 otherwise



Yoan Chabot – yoan.chabot@checksem.fr - Laboratoire Electronique Informatique et Im
(LE2I – UMR CNRS 5158) IUT Dijon-Au gne, BP 47870, 21078 Dijon Cedex, France

Automatic Analysis Operators

$$Correlation(e,x) = Correlation_T(e,x) + Correlation_S(e,x) + Correlation_O(e,x) + Correlation_{KBR}(e,x)$$



$\boldsymbol{Correlation(e1,e2)} \approx$ 1,143

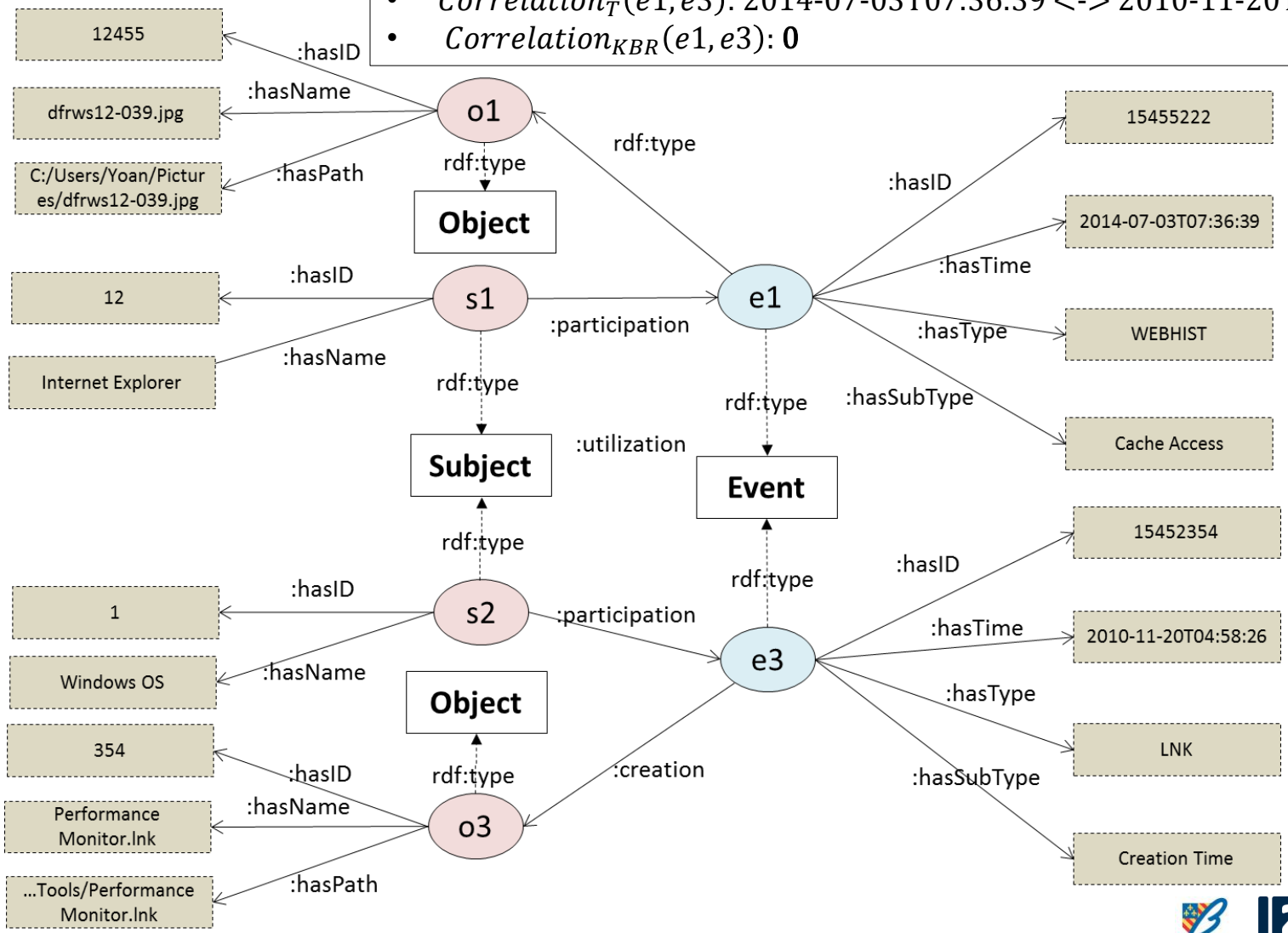
- $Correlation_O(e1,e2)$: *o1* ➔ ***1/1 = 1***
- $Correlation_S(e1,e2)$: $\emptyset$ ➔ **0/1 = 0**
- $Correlation_T(e1,e2)$: 2014-07-03T07:36:39 <-> 2014-07-03T07:36:46 ➔ ≈**0,143**
- $Correlation_{KBR}(e1,e2)$: **0**

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Automatic Analysis Operators

$\boldsymbol{Correlation(e1, e3)} \approx \mathbf{0}$

- $Correlation_O(e1, e3)$: $\emptyset$ ➔ $\mathbf{0/1 = 0}$
- $Correlation_S(e1, e3)$: $\emptyset$ ➔ $\mathbf{0/1 = 0}$
- $Correlation_T(e1, e3)$: 2014-07-03T07:36:39 <-> 2010-11-20T04:58:26 ➔ $\approx \mathbf{0}$
- $Correlation_{KBR}(e1, e3)$: $\mathbf{0}$



Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatiqu
(LE2i – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Contributions

Data volume

- Automatic Operators
- Scalable Technologies

Heterogeneity

- Unified model of knowledge representation
- Extractors dedicated to each source

Credibility

- Based on a formal knowledge representation

Reproducibility

- Storing information about provenance

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

Future Works

New Analysis operators

New Semantic dimensions

Mechanisms for knowledge checking and reproducibility

Yoan Chabot – yoan.chabot@checksem.fr - Equipe de projet Checksem – Laboratoire Electronique Informatique et Image (LE2I – UMR CNRS 5158) IUT Dijon-Auxerre – Université de Bourgogne, BP 47870, 21078 Dijon Cedex, France

DFRWS
DIGITAL FORENSIC RESEARCH WORKSHOP

August 5, 2014

A Complete Formalized Knowledge Representation Model for Advanced Digital Forensics Timeline Analysis

Yoan Chabot[a,b], Aurélie Bertaux[a], Christophe Nicolle[a] and M-Tahar Kechadi[b]

yoan.chabot@checksem.fr

[a] CheckSem Team, Laboratoire Le2i, Université de Bourgogne, Dijon, FRANCE
[b] School of Computer Science & Informatics, University College Dublin, IRELAND