Dynamic Correlation of Digital Forensics Reports

Christoph Beckmeyer

Aachen University of Applied Sciences

christoph.beckmeyer@alumni.fh-aachen.de

IMF 2014

8th International Conference on It Security Incident Management and IT Forensics

Presentation Overview

Problem description

Proposed solution

Demo

Limitations & Future Work

DIRECT - Project History

Project started 2012

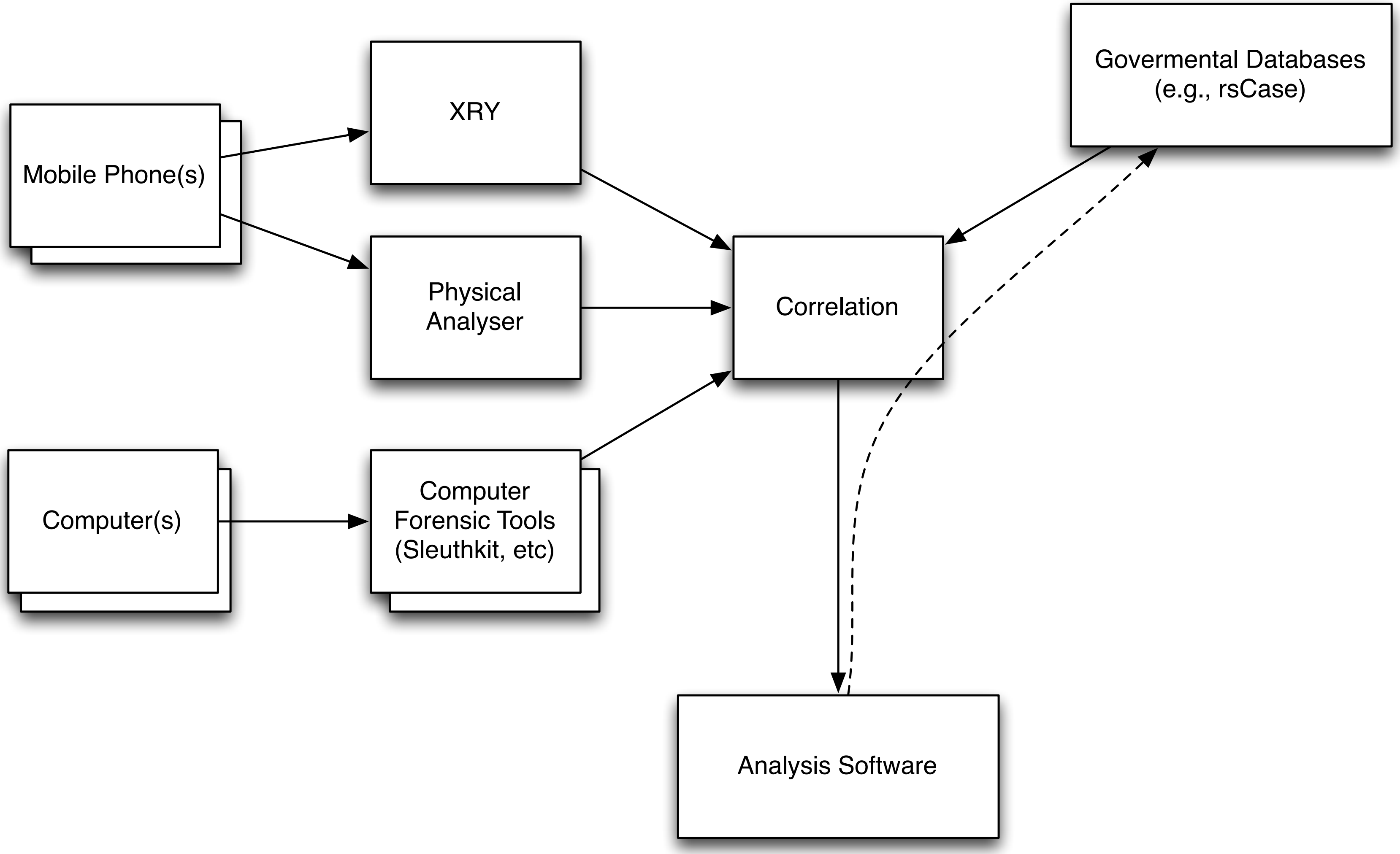
by Martin Pfeiffer as his Bachelor thesis.

Vision of a Research Tool.

Prototypical Database Design for Correlation.

Importer for Physical Analyzer 2.

Correlation in Digital Forensics


Mobile Phone(s)
XRY
Physical Analyser
Computer(s)
Computer Forensic Tools (Sleuthkit, etc)
Correlation
Govermental Databases (e.g., rsCase)
Analysis Software

Correlation in Digital Forensics

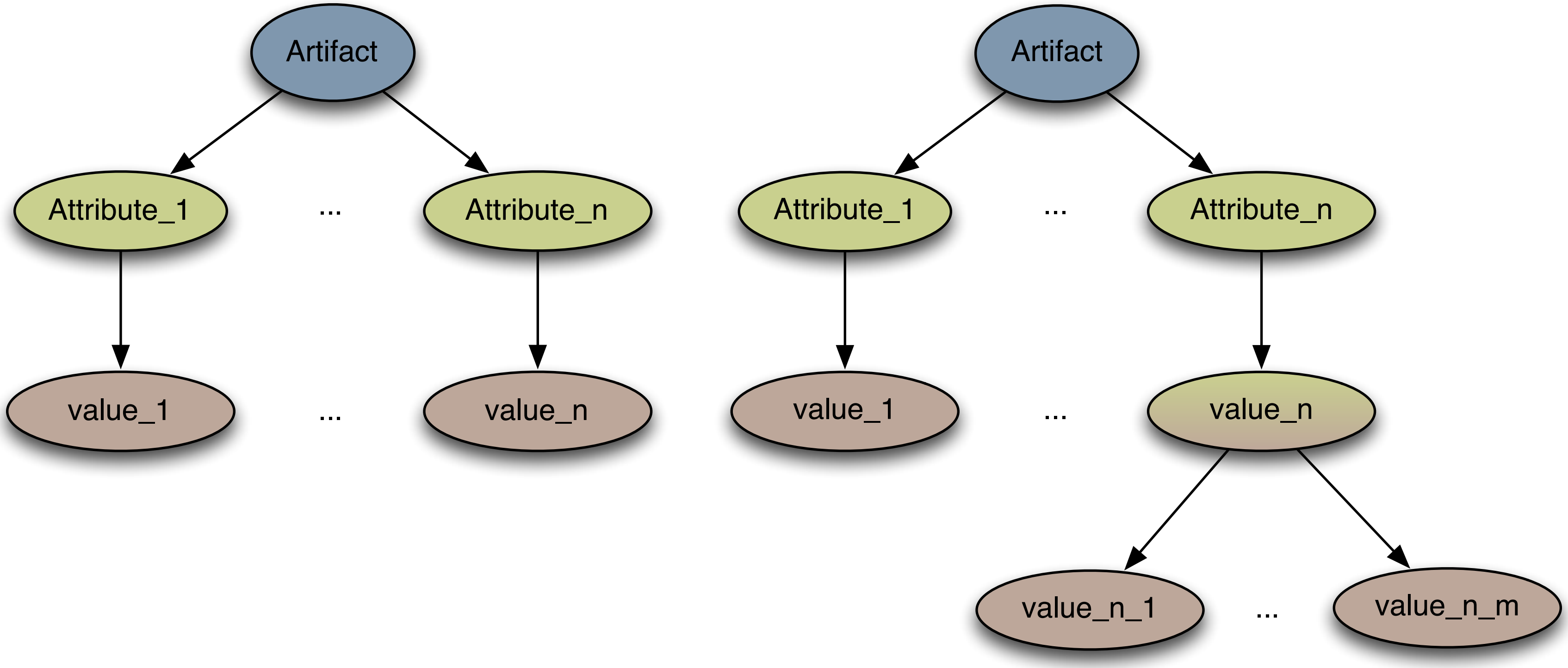
1st Level Correlation:

Joining of sets of artifacts with same semantic from different sources.

2nd Level Correlation:

Establishing semantic relationships between individual artifacts.

Artifacts


Artifact
Attribute_1
...
Attribute_n
value_1
...
value_n
Artifact
Attribute_1
...
Attribute_n
value_1
...
value_n
value_n_1
...
value_n_m

Correlation in Digital Forensics

Use cases for correlation:

Case evidence spread across devices.

Extraction tools have different capabilities.

Organizations use different tools.

Correlation as a Problem

Many existing file formats

New file formats emerge

Existing formats change (CDR)

Across formats: different syntax for same semantics

Practitioners' current Solutions

Excel

Often manual labor.

Custom Development

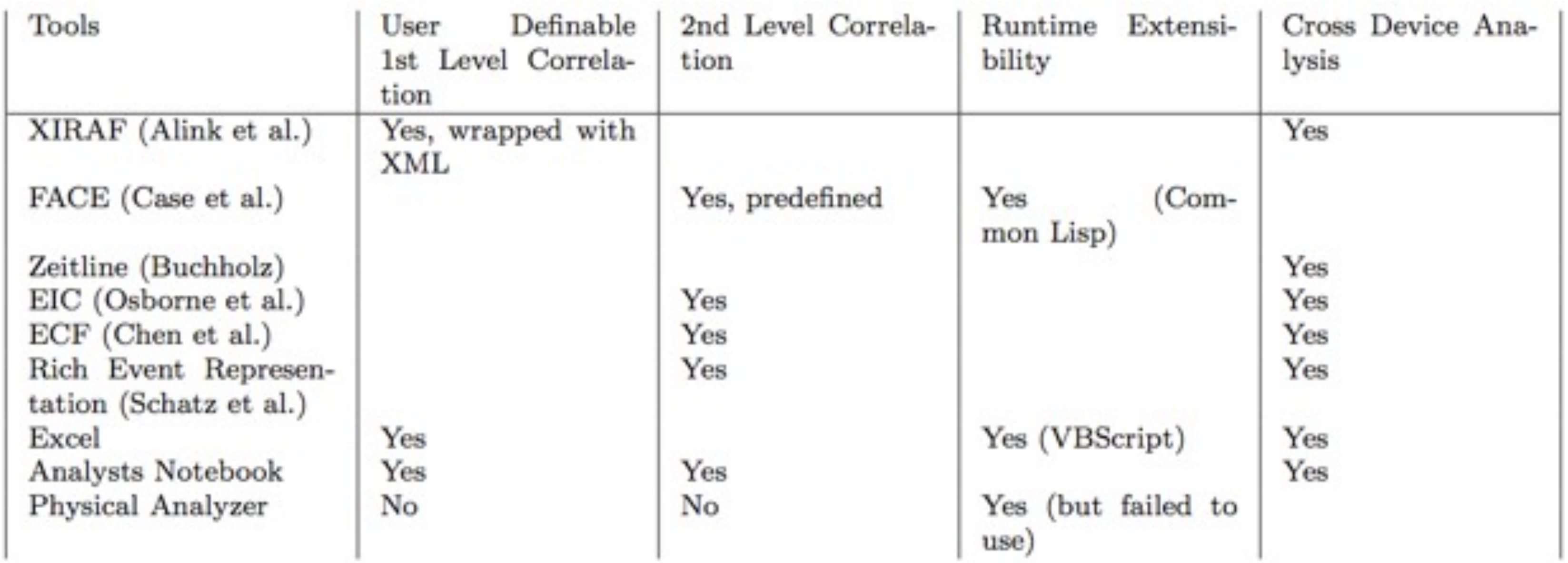
Longer development cycle.

Commercial Analysis Tools

Expensive.

Current Solutions to Correlation

Reviewed tools:



Tools	User Definable 1st Level Correlation	2nd Level Correlation	Runtime Extensibility	Cross Device Analysis
XIRAF (Alink et al.)	Yes, wrapped with XML			Yes
FACE (Case et al.)		Yes, predefined	Yes (Common Lisp)	
Zeitline (Buchholz)				Yes
EIC (Osborne et al.)		Yes		Yes
ECF (Chen et al.)		Yes		Yes
Rich Event Representation (Schatz et al.)		Yes		Yes
Excel	Yes		Yes (VBScript)	Yes
Analysts Notebook	Yes	Yes		Yes
Physical Analyzer	No	No	Yes (but failed to use)	

Current Solutions to Correlation

Previous Lessons from DIRECT:

Embedding into existing tool failed.

Modeling domain into RDBMS was inflexible.

Conclusions from current solutions:

Practitioners are in need for flexible tool to:

- Quickly correlate by themselves. (Cut development cycle)

Current tools often correlate into static model.

- Hinders fast adaptation to change in inputs.

Flexible correlation is possible with commercial tools.

- But they might not support your use case. (e.g., comparison of inputs)

Proposed Solution

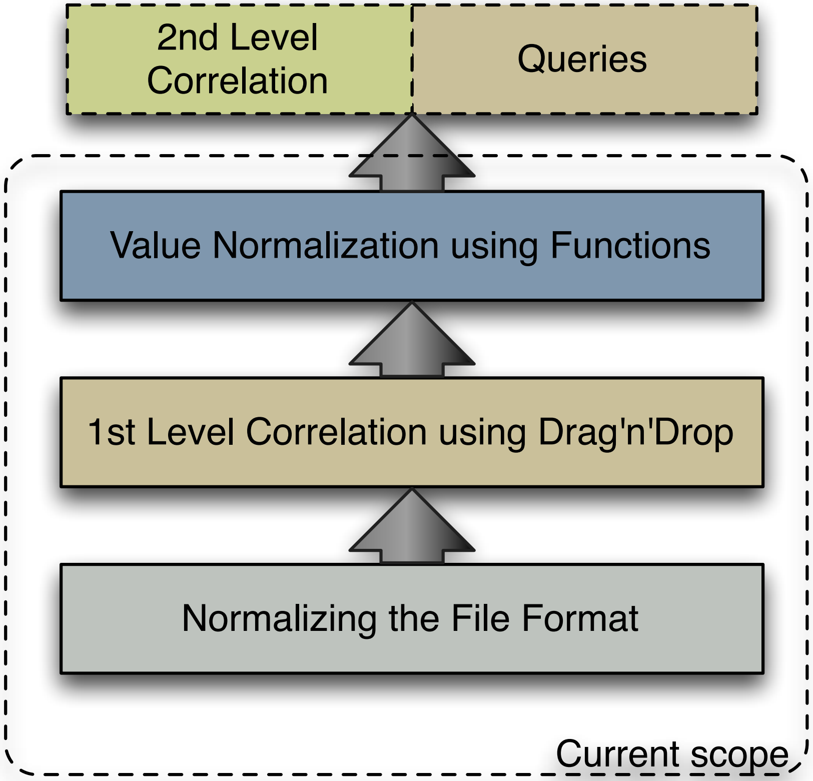
Introduction of Abstractions

1. Normalizing the file format.

2. Correlation using Drag'n Drop.

3. Normalization by attaching functions.

Introducing Abstractions


2nd Level Correlation
Queries
Value Normalization using Functions
1st Level Correlation using Drag'n'Drop
Normalizing the File Format
Current scope

1. Normalizing the File Format

```
<model type="Call" id="82a…."
            deleted_state="Intact">
        <field name="Name" type="String">
          <empty />
        </field>
        <field name="Type"
                  type="CallType">
          <value type="CallType"><![CDATA[Outgoing]]></value>
        </field>
        <field name="TimeStamp" type="TimeStamp">
          <value type="TimeStamp">2005-09-13T09:11:34+02:00</value>
        </field>
</model>
```

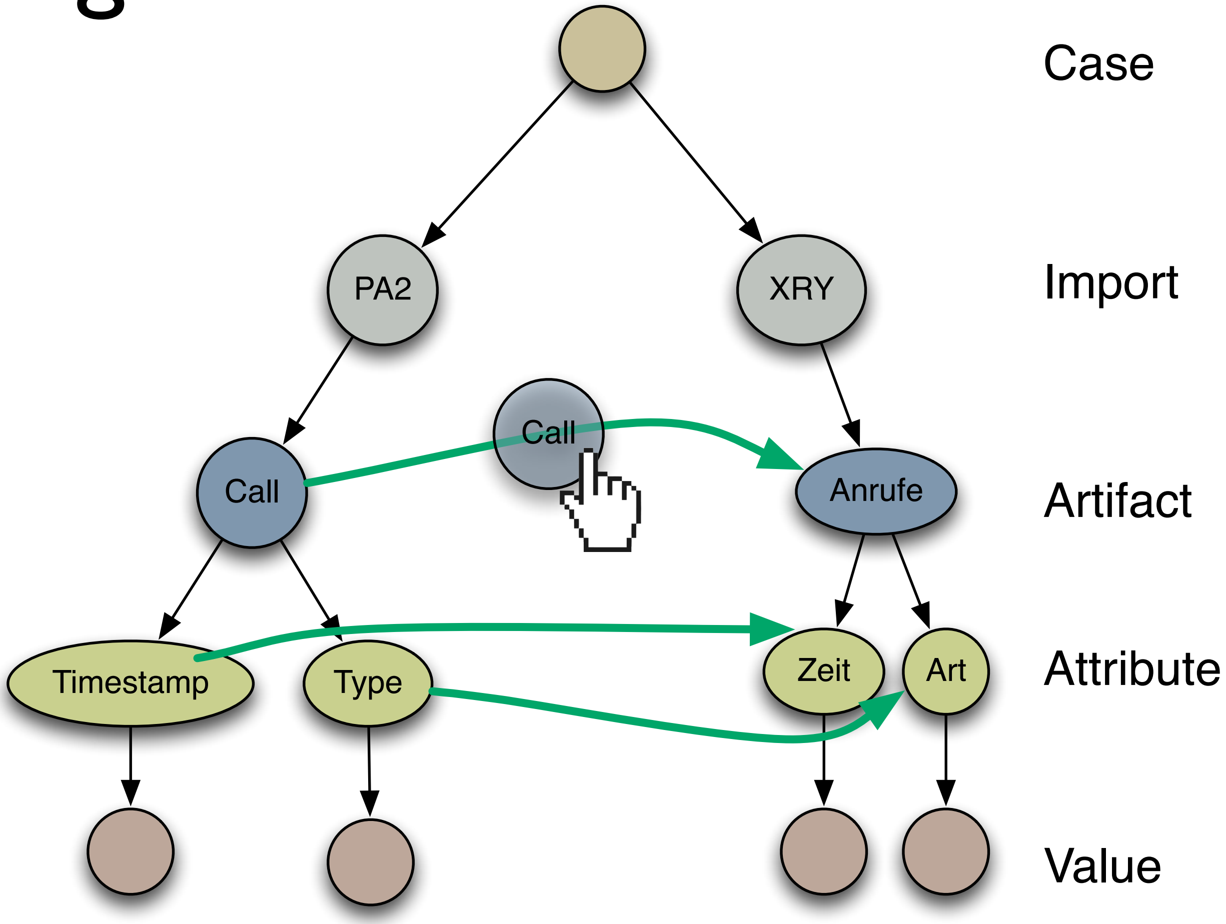
PA2

```
  <view name="Anrufe">
<item>
      <field name="Art"
            value="Entgegengenommen"
            class="STATUS"/>
      <field name="Zeit"
             value="13.10.2013 11:37:30 (Gerät)"
             class="TIME"/>
    </item>
```

XRY

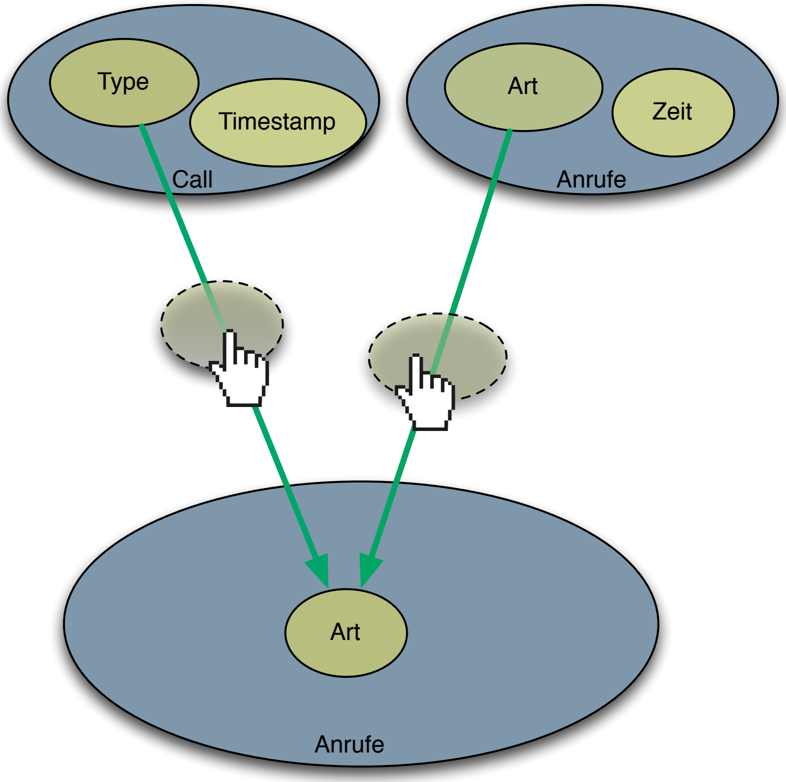
2. Correlation using Drag'n Drop

Merging artifacts

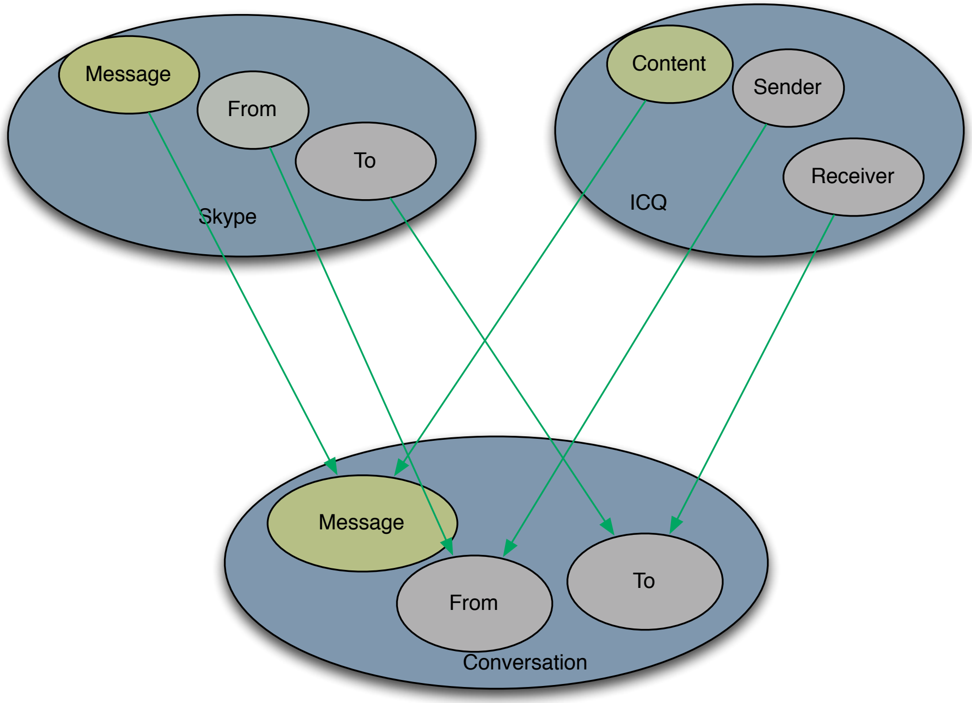


2. Correlation using Drag’n Drop

Modeling the domain

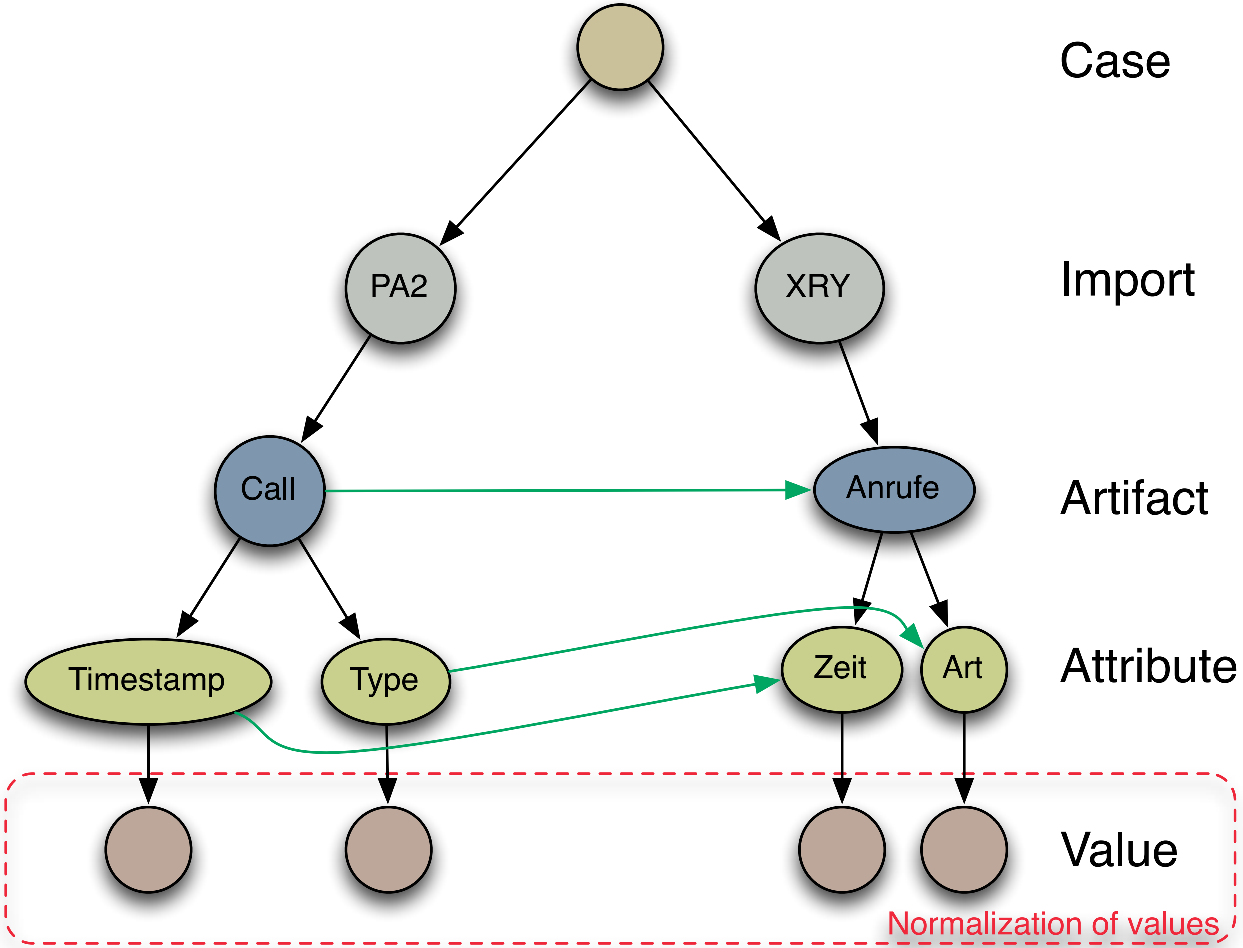


Fusion of XRY and Physical Analyzer



Model conversations

3. Value Normalization



3. Value Normalization

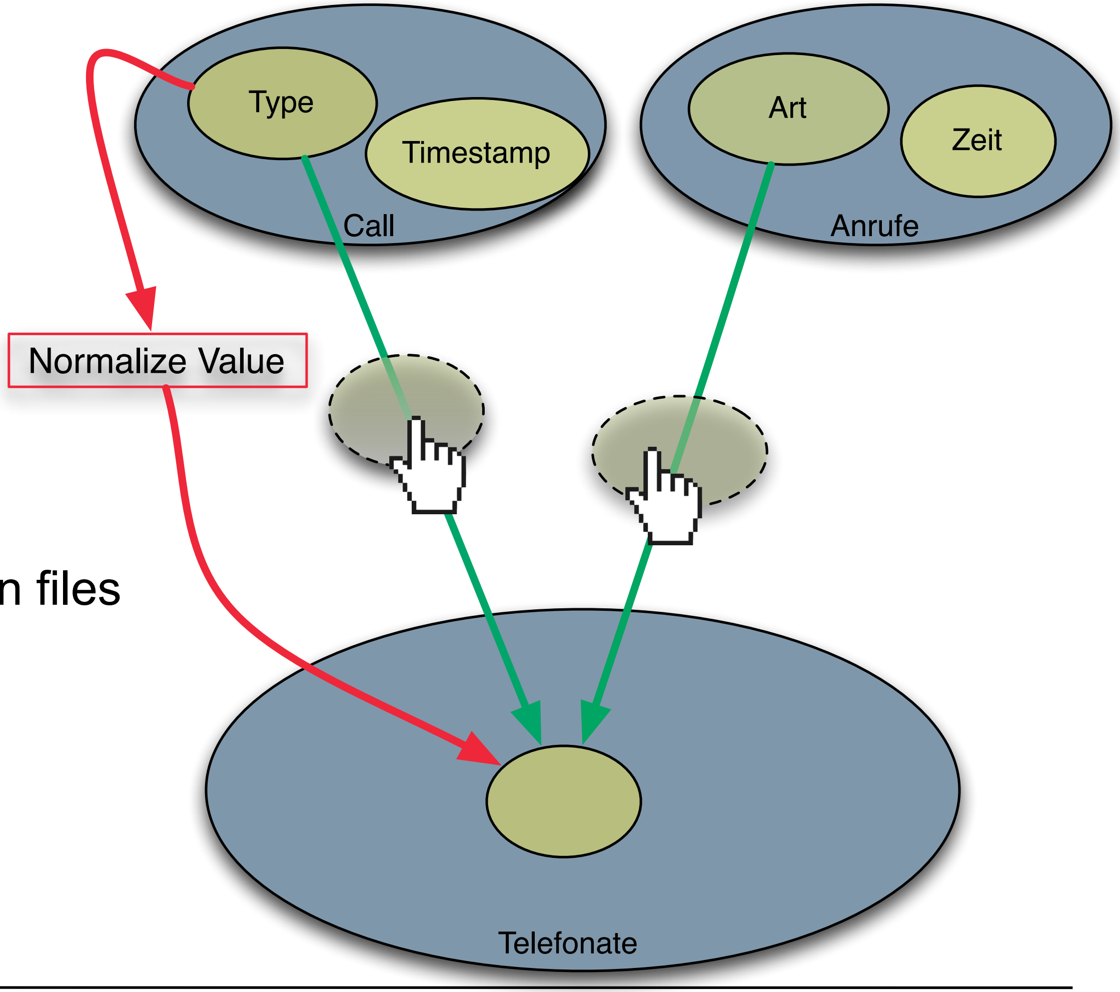
Attachable to Attributes

Examples:

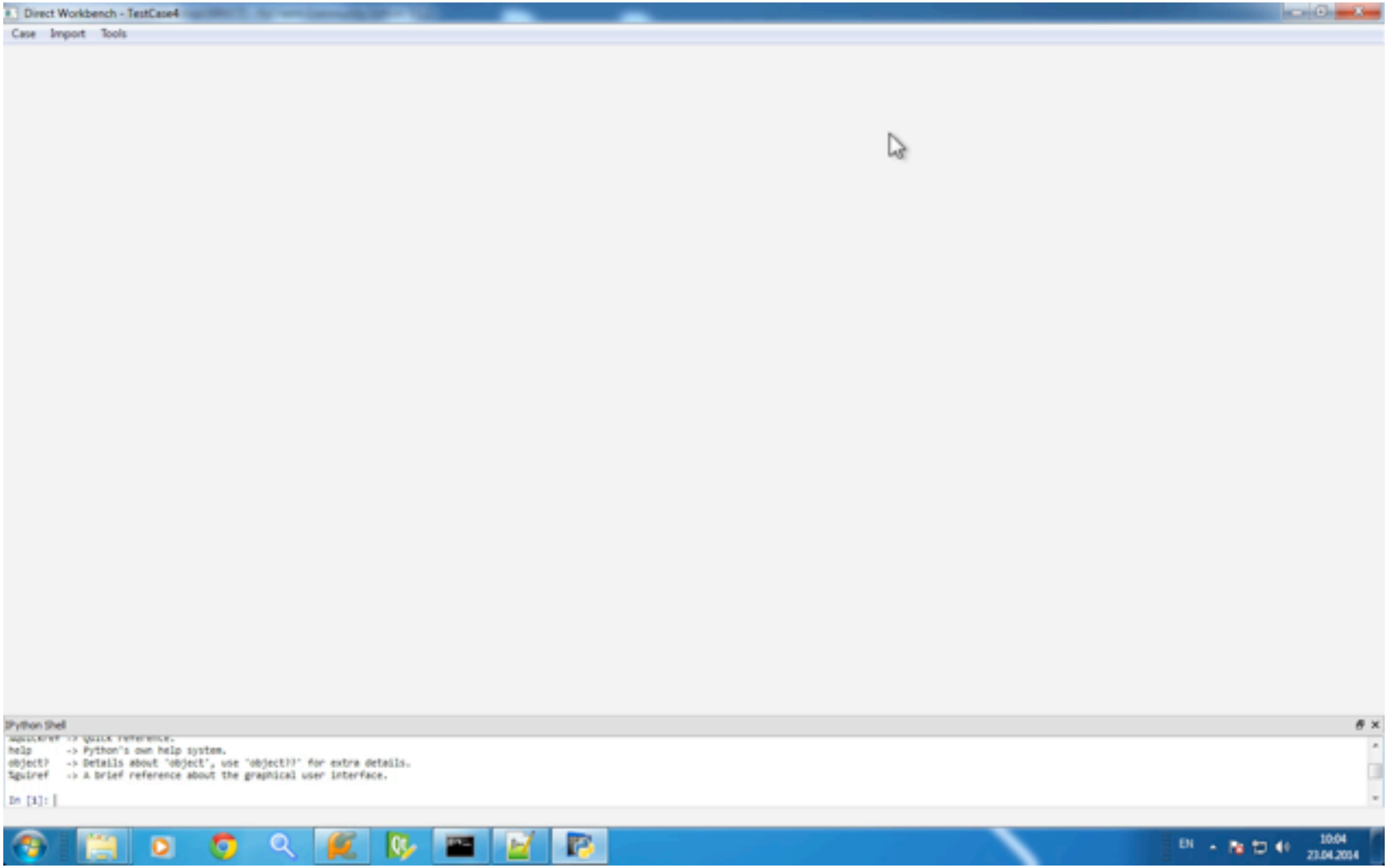
Date formats

Phone numbers

Matching against known files



Demo



```
%quickref -> Quick reference.
help      -> Python's own help system.
object?   -> Details about 'object', use 'object??' for extra details.
%guiref   -> A brief reference about the graphical user interface.

In [1]: |
```

Limitations

Better Graphical User Interface.

More functions.

Tracing to the original file.

Queries on the correlation result.

Future Work

Automatic Matching

On artifact / attribute names

On attributes values

2nd Level Correlation

Semantic Network

Deduction of relations

Thank you

Dynamic Correlation with DIRECT

1. Normalizing the file format.
2. Drag'n Drop correlation for artifacts & attributes.
3. Function library for value normalization.

Christoph Beckmeyer
chrisb@alumni.fh-aachen.de
www.it-forensik.fh-aachen.de/projekte/direct