# *UniQue:* An Approach for Unified and Efficient Querying of Heterogeneous Web Data Sources

## Building mashup applications is *challenging*

➤ *Heterogeneous data formats*
➤ *Heterogeneous query techniques*
➤ *Inefficient use of network resources*

## The *UniQue* approach

The unified querying (UniQue) approach and a proxy-based implementation provides a uniform and declarative interface for querying heterogeneous data sources across the Web. Besides hiding the differences between the underlying data formats and query techniques, UniQue heavily embraces open W3C standards to minimize the learning effort required by web mashup developers.
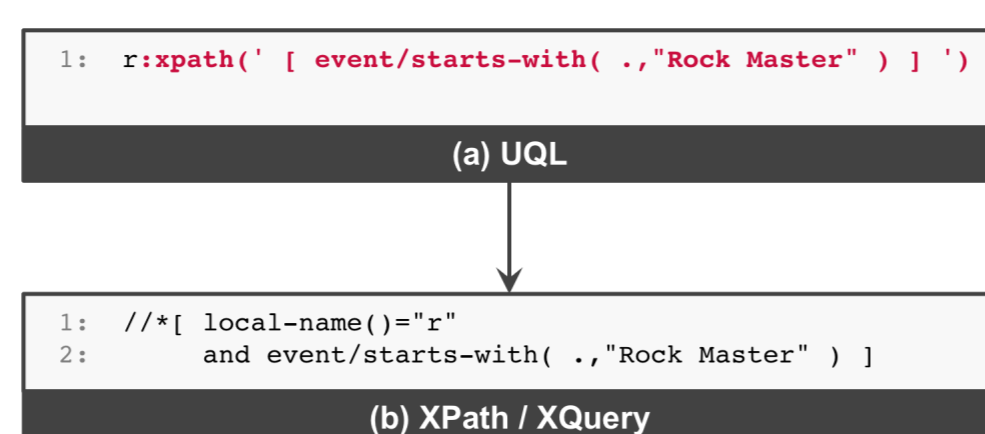
### One data format



```
 1: {
 2:     "meta": {
 3:         "code": 200
 4:     },
 5:     "data": {
 6:         "username": "chris_sharma",
 7:         ...
 8:         "counts": {
 9:             "media": 243,
10:             "followed_by": 119135,
11:             "follows": 425
12:         },
13:         "id": "264724726"
14:     }
15: }
```
(a) JSON

```
 1: rank,event,date
 2: 18,Rock Master (L) – Arco (ITA),16.07.2010
 3: 8,IFSC World Cup (B) – Vail (USA),06.06.2008
 4: 2,UIAA World Cup (B) – Lecco (ITA),25.06.2003
 5: 12,UIAA World Cup (L) – Aprica (ITA),18.10.2002
 6: 2,UIAA World Cup (B) – Rovereto (ITA),14.09.2002
 7: 49,UIAA World Cup (B) – Munich (GER),22.07.2001
 8: 3,UIAA World Cup (B) – GAP (FRA),30.06.2001
 9: 3,UIAA World Cup (B) – Rovereto (ITA),16.09.2000
10: 8,Rock Master (L) – Arco (ITA),10.09.1999
11: 6,Rock Master (B) – Arco (ITA),10.09.1999
12: 5,UIAA World Cup (L) – Birmingham (GBR),29.11.1997
13: 2,UIAA World Cup (L) – Imst (AUT),22.11.1997
14: 1,UIAA World Cup (L) – Kranj (SLO),09.11.1997
15: ...
```
(a) CSV

```
 1: <json object="true">
 2:     <meta object="true">
 3:         <code type="number">200</code>
 4:     </meta>
 5:     <data object="true">
 6:         <username type="string">chris_sharma</username>
 7:         ...
 8:         <counts object="true">
 9:             <media type="number">243</media>
10:             <followed_by type="number">119135</followed_by>
11:             <follows type="number">425</follows>
12:         </counts>
13:         <id type="string">264724726</id>
14:     </data>
15: </json>
```
(b) XML

```
 1: <csv separator="," quote="&quot;">
 2:     <h>
 3:         <rank type="number">rank</rank>
 4:         <event type="string">event</event>
 5:         <date type="string">date</date>
 6:     </h>
 7:     <r>
 8:         <rank>18</rank>
 9:         <event>Rock Master (L) – Arco (ITA)</event>
10:         <date>16.07.2010</date>
11:     </r>
12:     ...
13: </csv>
```
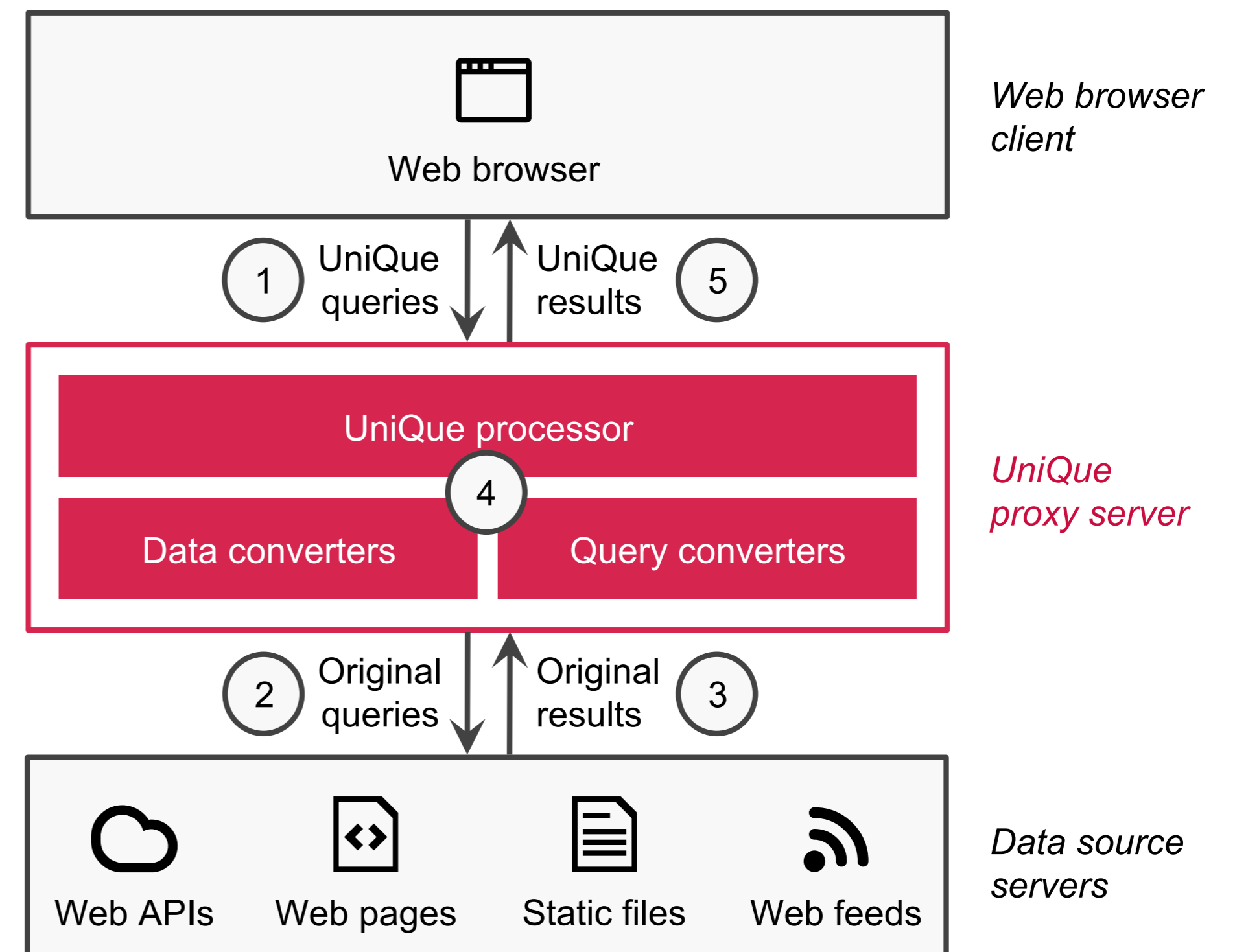(b) XML

Examples of mapping data between (a) JSON / CSV and (b) XML.

### Extendable query language

Example of extending the expressiveness of CSS Selectors with XPath. The query is translated from (a) UQL to (b) XPath / XQuery. Our :xpath() extension in UQL is highlighted in red.
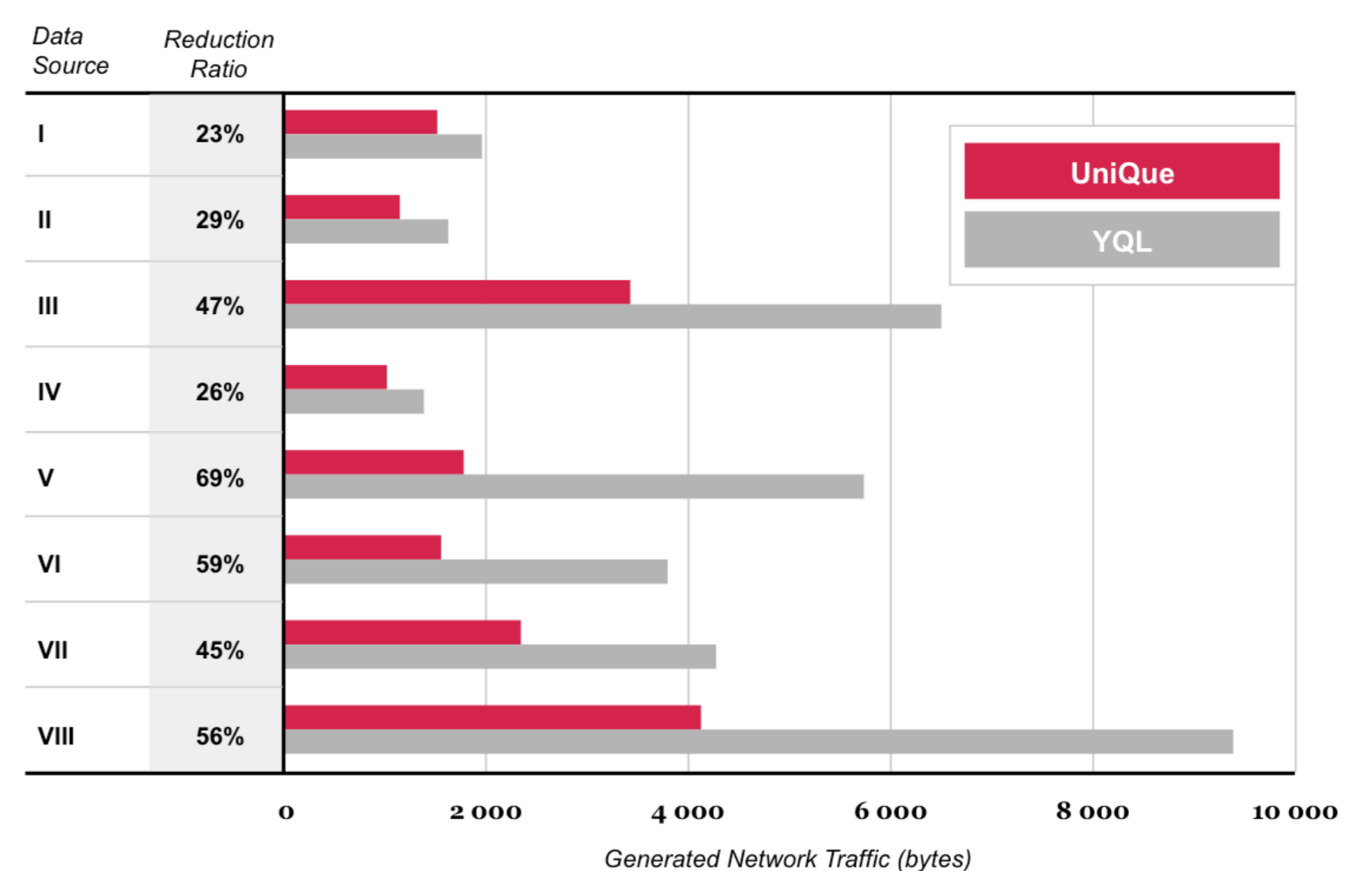
```
 1: r:xpath(' [ event/starts-with( .,"Rock Master" ) ] ')
```
(a) UQL

```
 1: //*[ local-name()="r"
 2:      and event/starts-with( .,"Rock Master" ) ]
```
(b) XPath / XQuery

## The *UniQue* implementation



UniQue proxy between a web browser and web data sources.

## Results



Comparison of generated network traffic per data source between UniQue and YQL.

### More information
https://mediatech.aalto.fi/publications/webservices/unique/

**Aalto University**
**School of Science**

**Department of Computer Science**

**Markku Laine**, M. Sc.
**Jari Kleimola**, D. Sc.
**Petri Vuorimaa**, D. Sc., Prof.