

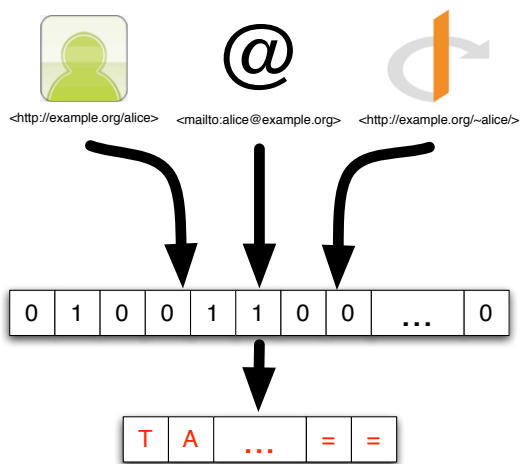


Overview

SPARQL is the standard query language for structured data on the Web. Querying a SPARQL server about known resources is a common task, but supporting "identity reasoning" with such queries can be verbose, often to the point of being impractical. A SPARQL extension may be used to express such queries in a compact form using Bloom filters.

Creating a Filter

The identities of resources of interest are encoded into a Bloom filter. The query is constrained by using a Base64 encoding of the Bloom filter in a SPARQL extension function, allowing queries over many times more resources than with standard SPARQL.



```
SELECT ?person ?phone
WHERE {
  ?person a foaf:Person ;
  foaf:phone ?phone .
  FILTER(
    :bloom(
      ?person, "TA...=="
    )
  )
}
```

Processing Results

Query results are returned with an identity map, allowing local data to be joined with the appropriate results.

?person	?phone
http://example.org/alice	tel:+1-123-555-0000
_:alice	tel:1234
_:eve	tel:+1-123-555-9999

Identity Map
"!foaf:openid <http://eve.name/>" → _:eve
"!foaf:mbox <mailto:alice@example.org>" → _:alice

?person	?phone
http://example.org/alice	tel:+1-123-555-0000
http://example.org/alice	tel:1234

("Any node locally identified by the foaf:mbox property <mailto:alice@example.org> may be joined with the _:alice node in the query results.")

Status and Conclusion

This technique yields 2 orders of magnitude improvement on query size, and opens new possibilities for query answering on a Web that is growing faster than existing approaches can handle.