

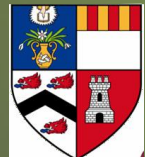
Stream Reasoning For Linked Data

M. Balduini, J-P Calbimonte, O. Corcho,
D. Dell'Aglio, E. Della Valle, and J.Z. Pan

<http://streamreasoning.org/sr4ld2013>



ISWC 2013
Sydney, Australia



RDF stream processing models

- As you know, “*RDF is a standard model for data interchange on the Web*” (<http://www.w3.org/RDF/>)
 $\langle \text{sub}_1 \text{ pred}_1 \text{ obj}_1 \rangle$
 $\langle \text{sub}_2 \text{ pred}_2 \text{ obj}_2 \rangle$
- We want to extend RDF to model data streams
- A data stream is an (infinite) ordered sequence of **data items**
- A data item is a self-consumable informative unit

- With **data item** we can refer to:

1. A **triple**

`<:alice :meets :bob>`

2. A **graph**

`<:alice :participates :meeting>`

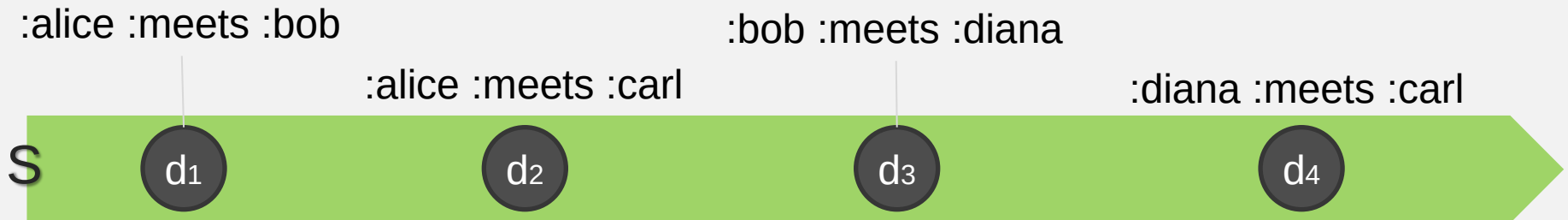
`<:bob :participates :meeting>`

`<:meeting :haslocation :room>`

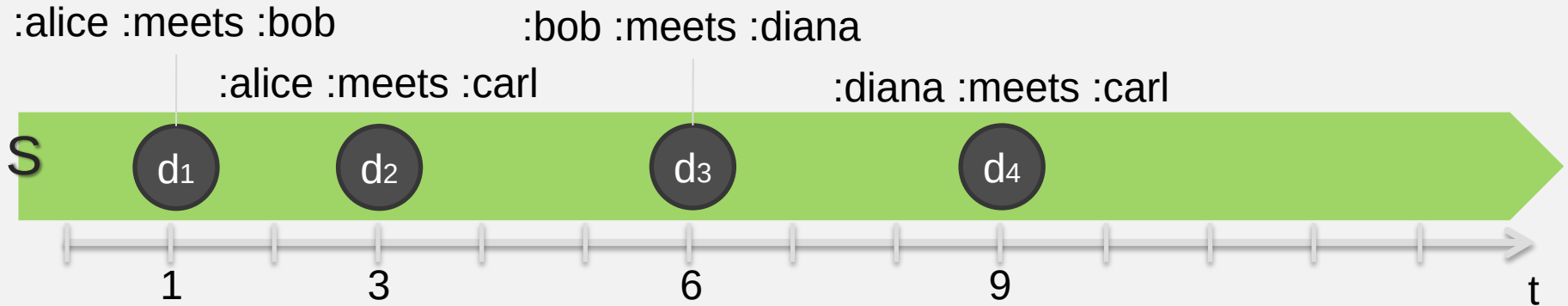
} `:graph1`

- Do we need to associate the time (as metadata) to data items?
 - It depends on what we want to achieve (see next!)
- If yes, how to take into account the time?
 - Time should not (but could) be part of the schema
 - Time should not be accessible through the query language
 - Time as object would require a lot of reification
- How to extend the RDF model to take into account the time?

- A timestamp is a temporal identifier associated to a data item
- The **application time** is a set of one or more timestamps associated to the data item
- Two data items can have the same application time
 - Contemporaneity
- Who does assign the application time to a data item?
 - The one that generates the data stream!

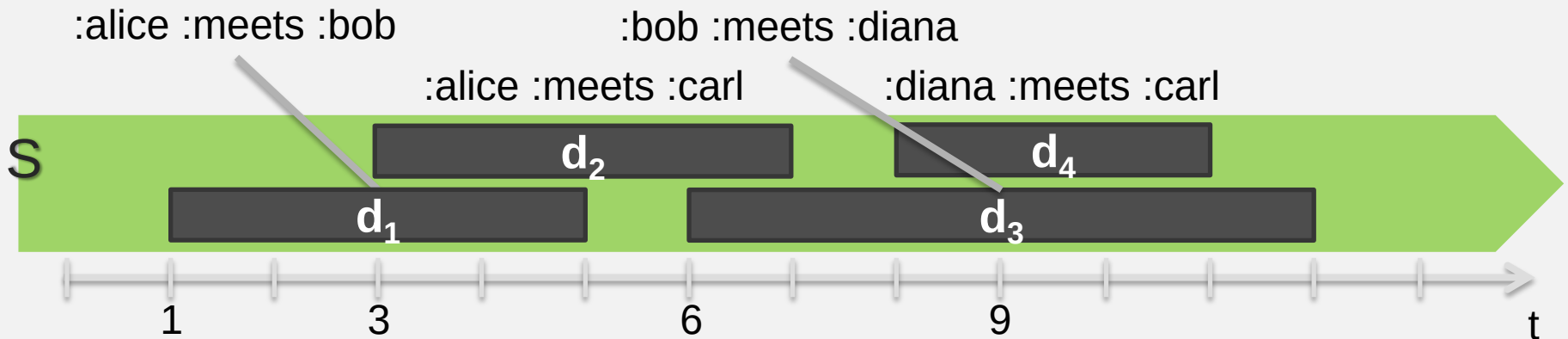


- A RDF stream without timestamp is an ordered sequence of data items
- The order can be exploited to perform queries
 - Does Alice meet Bob before Carl?
 - Who does Carl meet first?



- One timestamp: the time on which the data item occurs
- We can start to compose queries taking into account the time
 - How many people has Alice met in the last 5m?
 - Does Diana meet Bob and then Carl within 5m?

Application time: two timestamps



- Two timestamps: the time range on which the data item is valid (from, to]
- It is possible to write even more complex constraints:
 - Which are the meetings the last less than 5m?
 - Which are the meetings with conflicts?

	Triple	Graph
No timestamp	Instans	
One timestamp	C-SPARQL CQELS SPARQLstream	SLD
Two timestamps	EP-SPARQL/Etalis	