

# Disentangling the notion of SPARQL datasets

Daniel Hernández  
dhernand@dcc.uchile.cl

Claudio Gutierrez  
cguierr@dcc.uchile.cl

Center for Semantic Web Research  
Universidad de Chile

Alberto Mendelzon Workshop  
Lima, 2015

# Frequently questions

Can blank nodes be shared across named graphs?

What dataset is used if there is no dataset description?

What should be done if an IRI occurs twice or more in a dataset description?

What should be done if there are repeated clauses in a dataset description?

Can be graphs be named with blank nodes?

Must an IRI that occurs in a dataset description be dereferenced if it also names a graph in the default dataset of the endpoint?

**A dataset is**

$\{G_0, (u_1, G_1), \dots, (u_n, G_n)\}$

# HTTP

(  ,  $a$  )

(  ,  $e$  )

(  ,  $d$  )

(  ,  $f$  )

GET  $x$  is answered with a  ,  $\delta(x)$  is a 

# HTTP

( , *a*)

( , *c*)

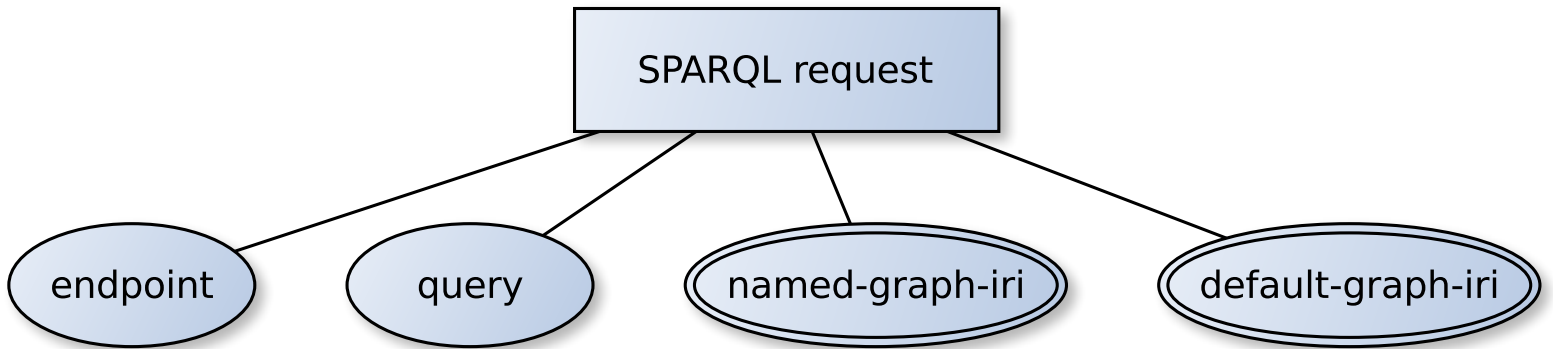
( , *b*)

( , *e*)

( , *f*)

( , *d*)

GET x?query=q



GET `s?query=q&default-graph-iri=u`

# Dataset description syntax

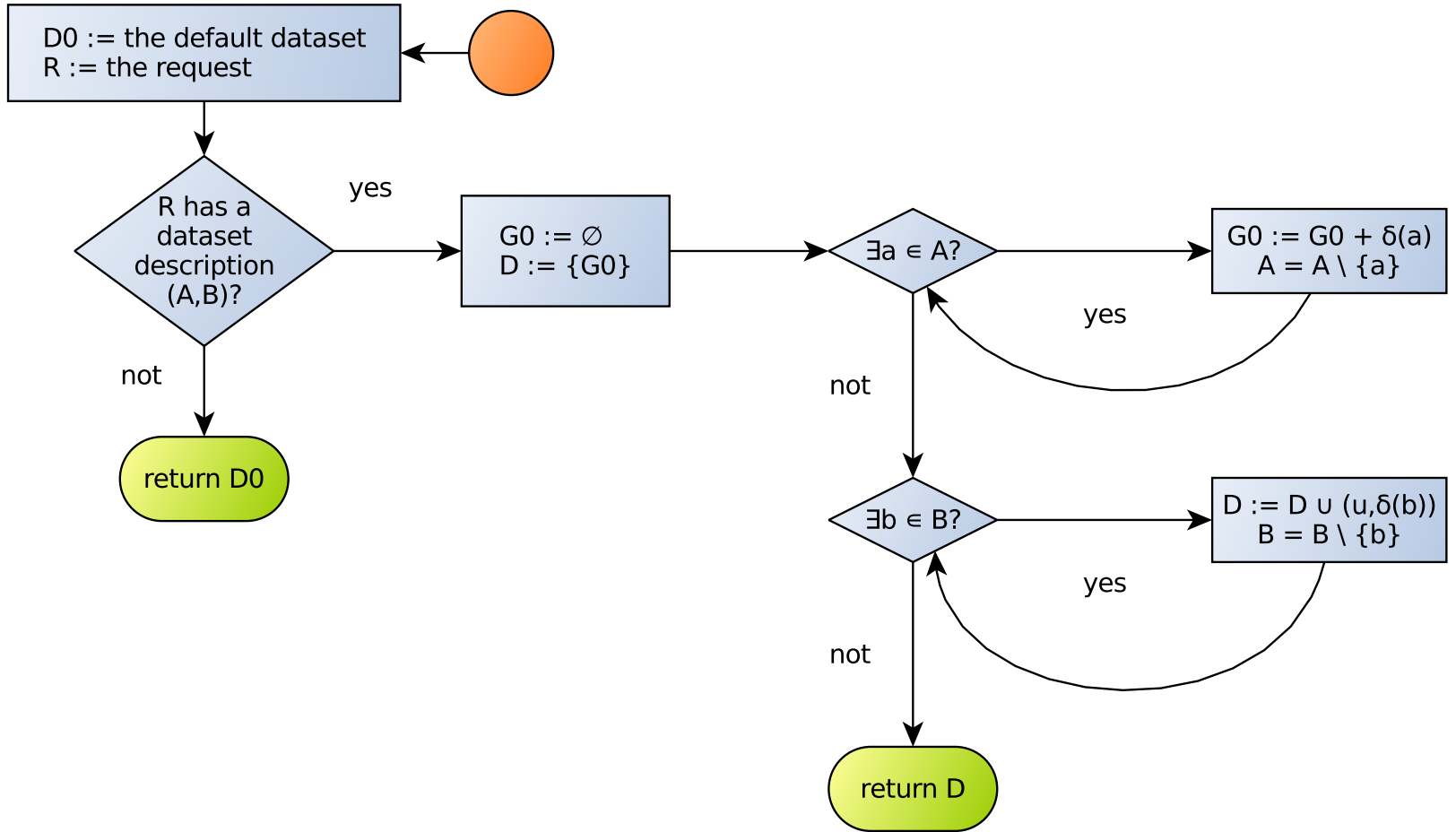
1. Query string parameters
2. FROM a FROM NAMED b

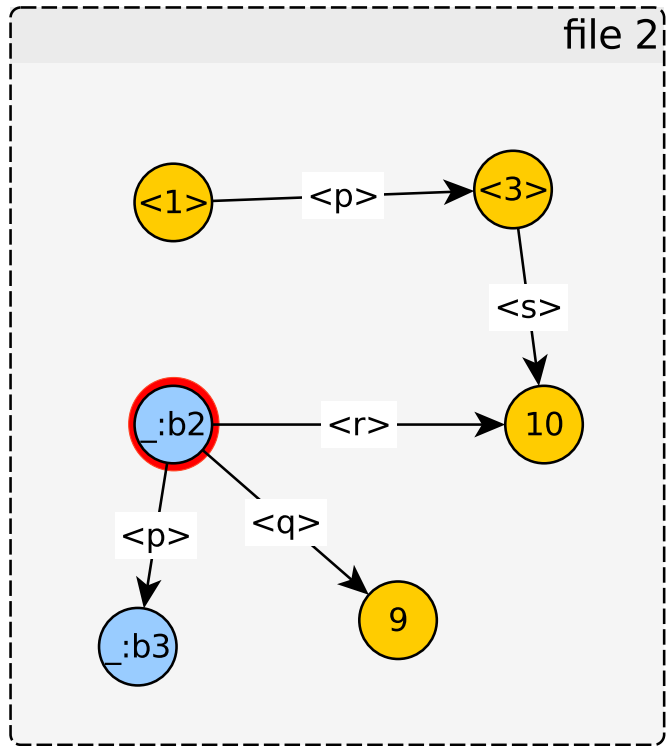
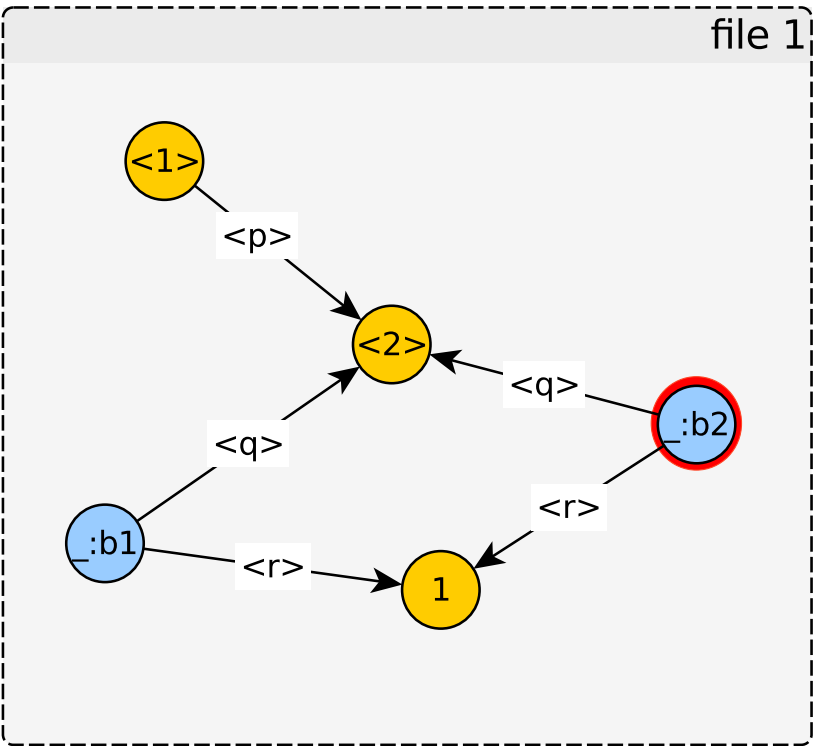


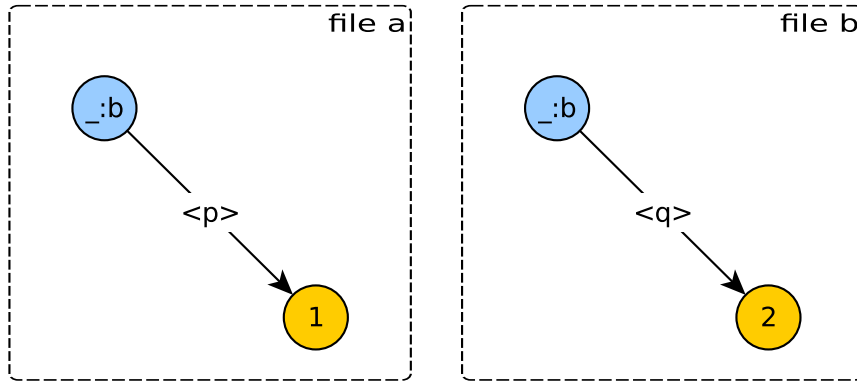


**What dataset is used  
to answer a  
SPARQL request?**

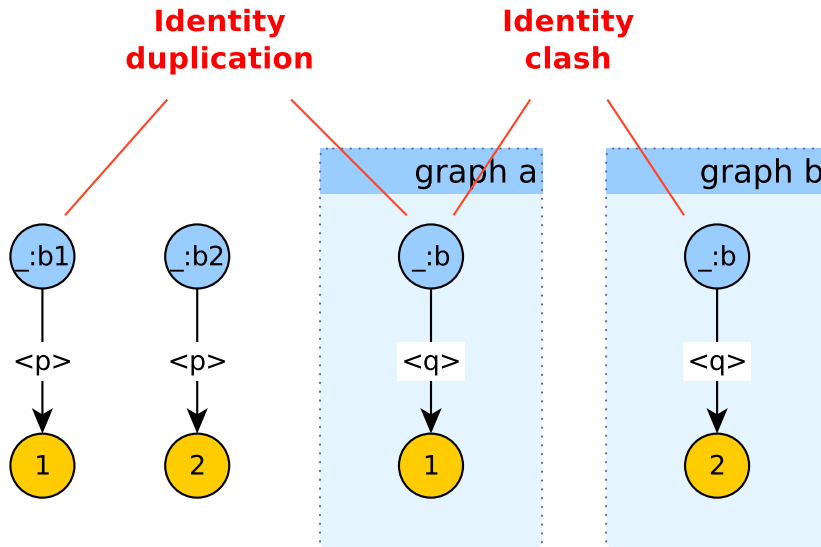
(dataset description≠dataset)





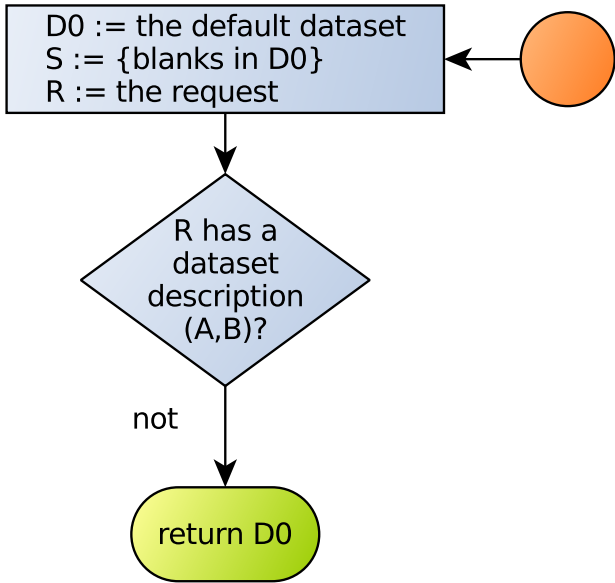


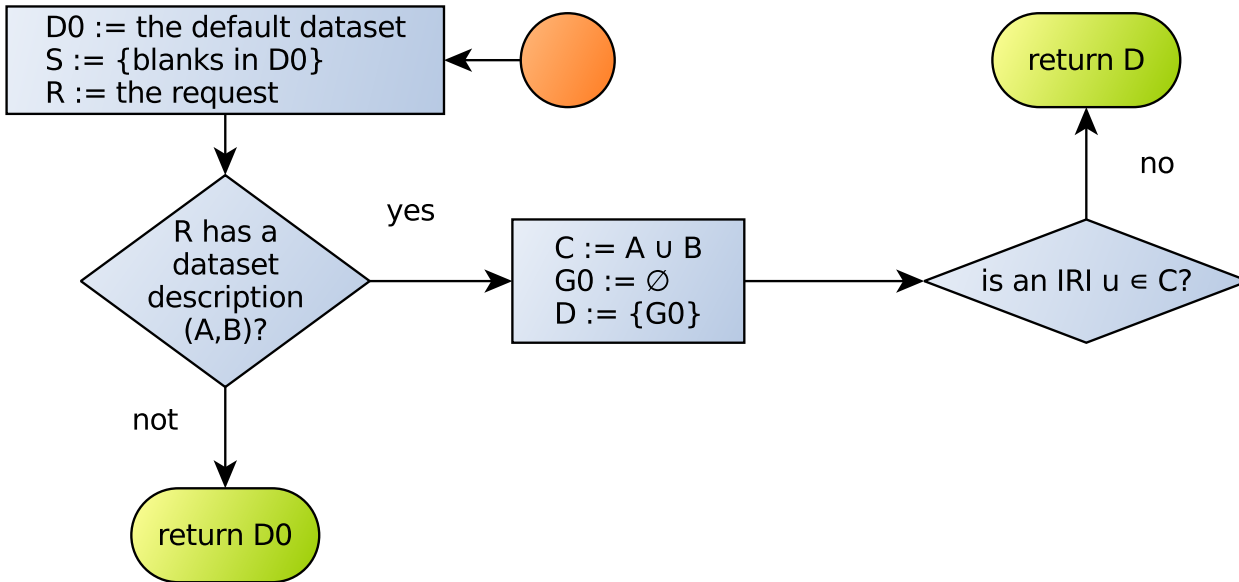
FROM a FROM b FROM NAMED a FROM NAMED b

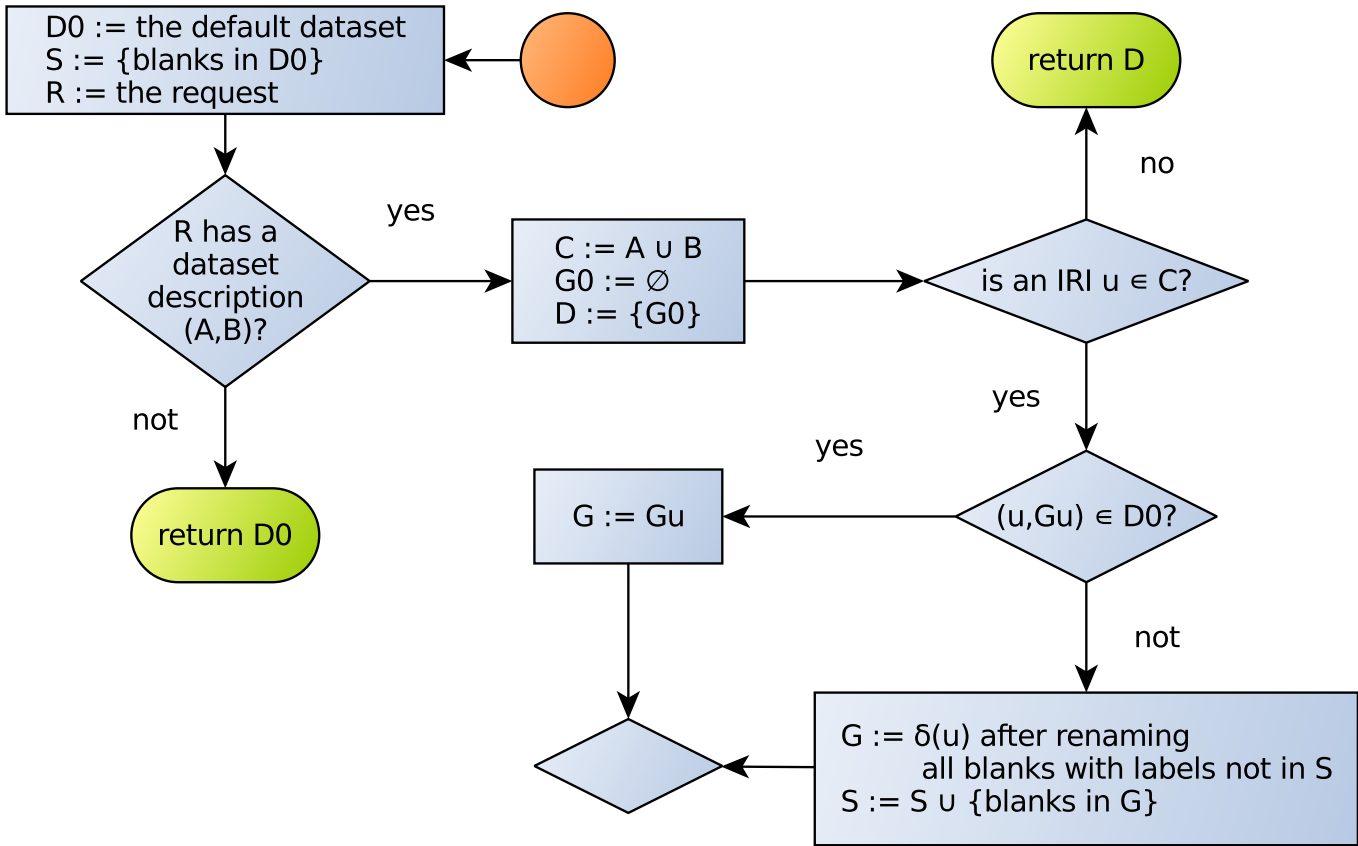


# Desiderata for a semantics of the evaluation of a dataset description

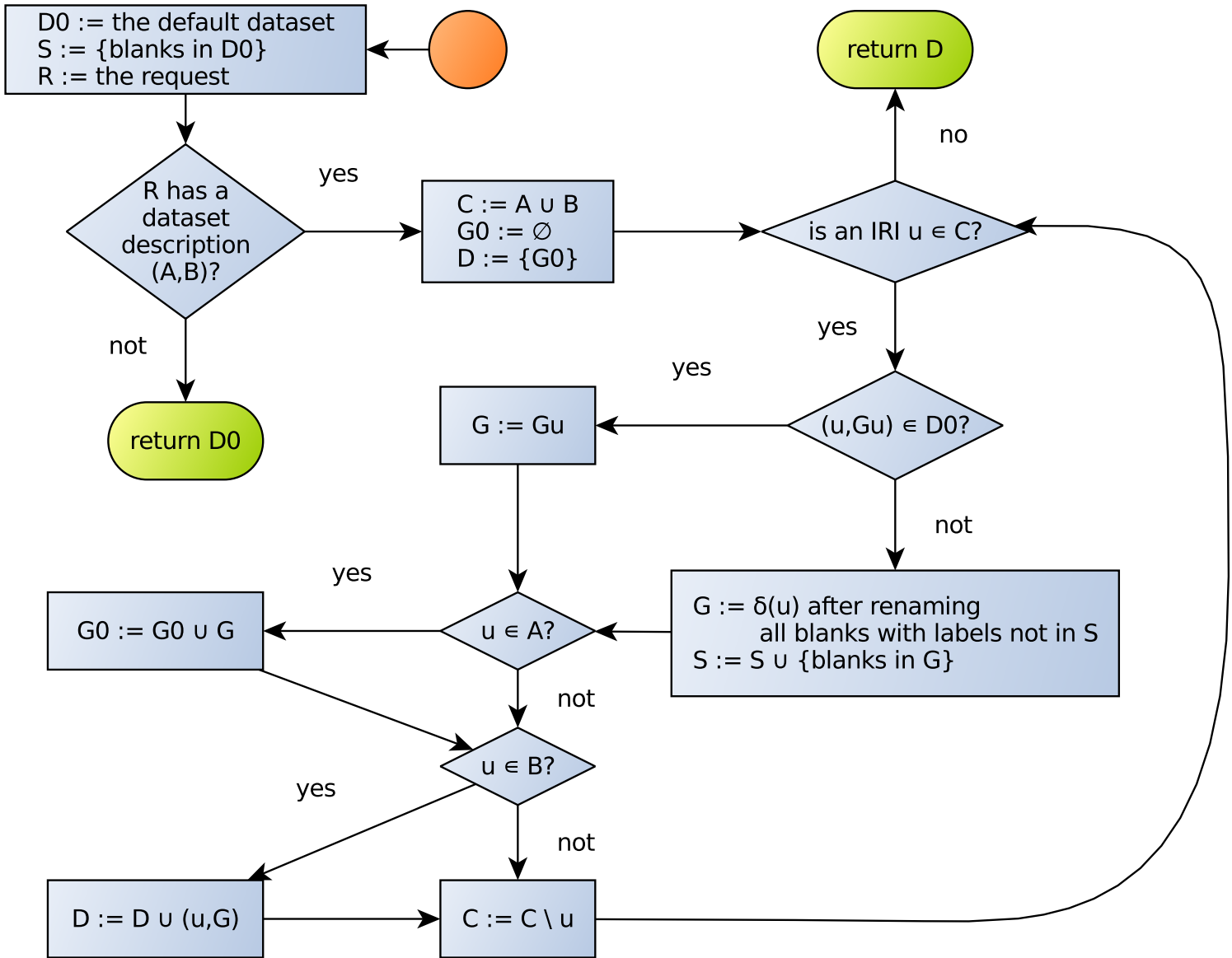
1. A dataset description should be composed by sets of IRIs instead of lists.
2. Files should be retrieved once.
3. The named graph in the default dataset should not be dereferenced.
4. The default dataset should define a scope.
5. Retrieved files should define a scope.
6. Avoid duplications and clashes of blank nodes.











**Thank you!**