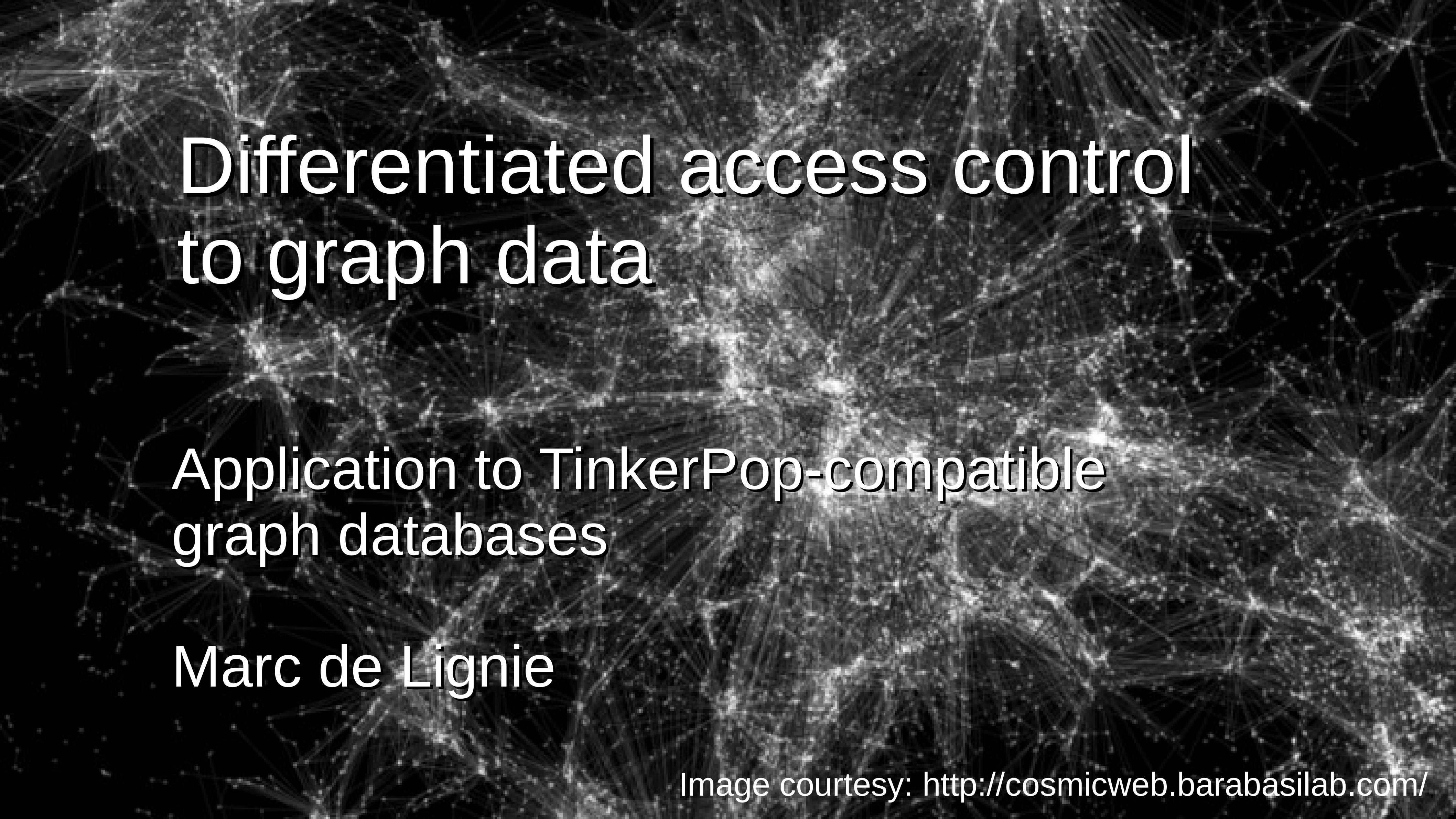


Differentiated access control to graph data

Application to TinkerPop-compatible graph databases

Marc de Lignie

Image courtesy: http://cosmicweb.barabasilab.com/

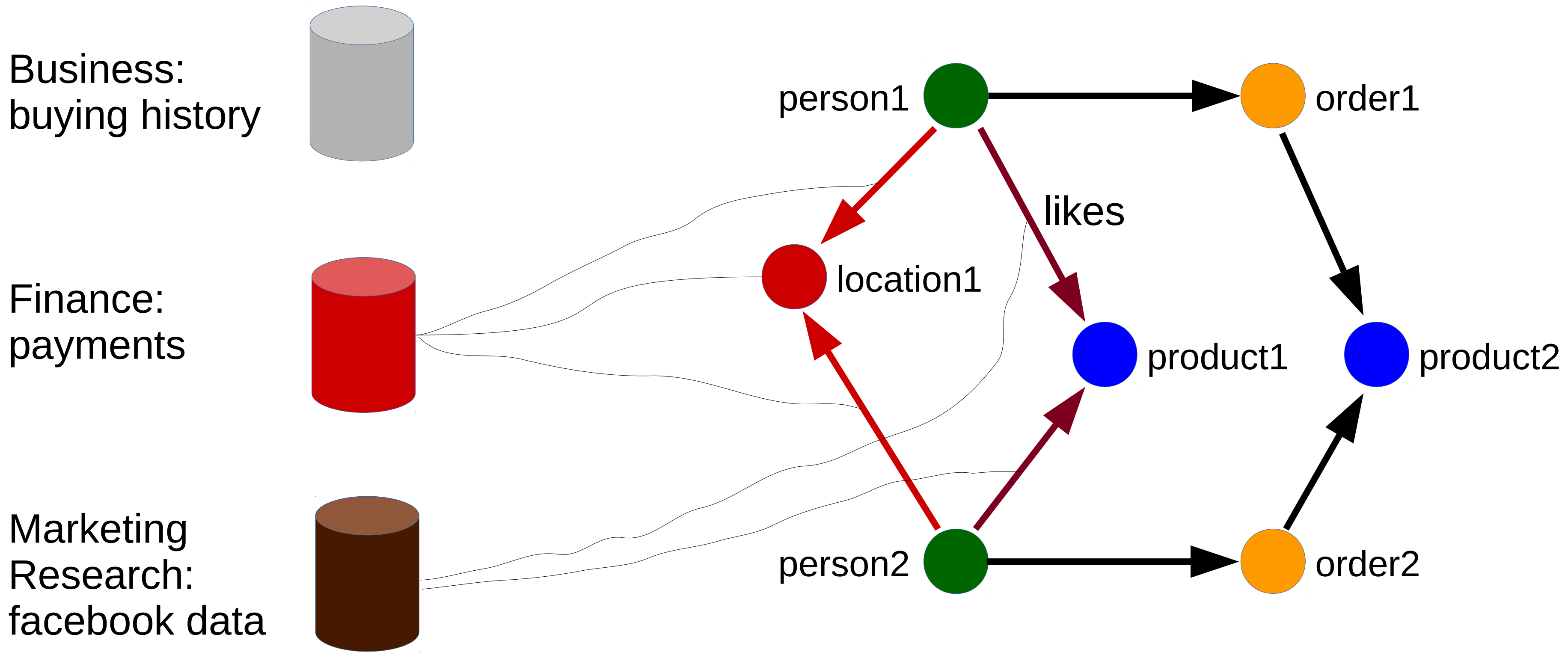
About me

1. self-taught data scientist, starting from a PhD in physics
2. interested in graph analytics and data fusion
3. employed at a Dutch government agency
4. contributor to Apache TinkerPop
5. active in JanusGraph community
6. http://yaaics.blogspot.com

Differentiated access control to graph data

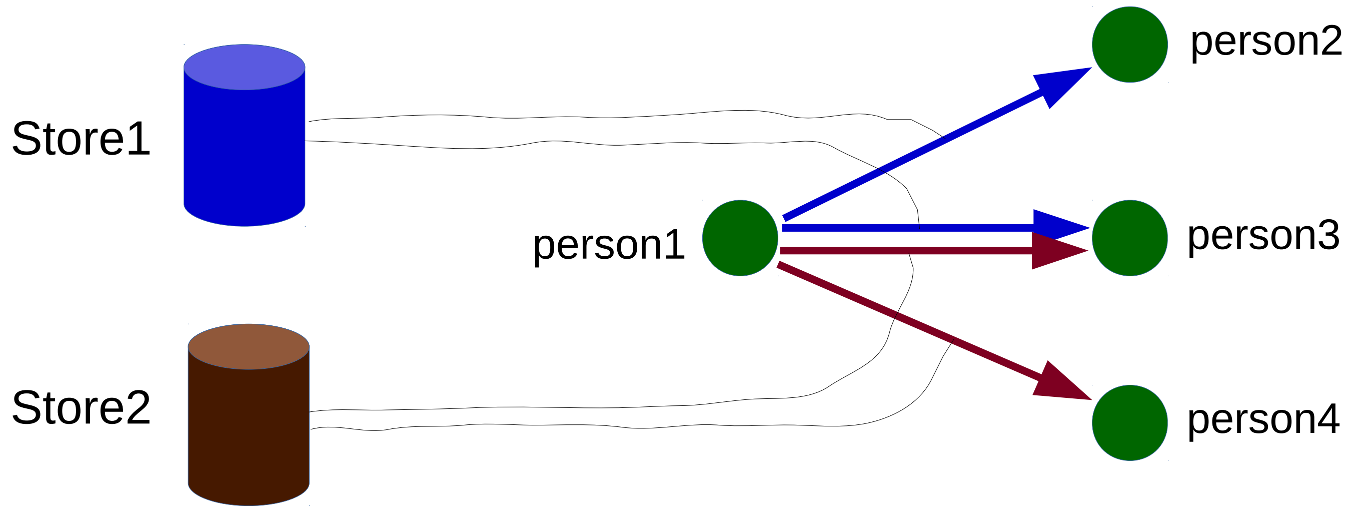
1. Exploration
2. Directions
3. Application to TinkerPop/JanusGraph
   << notebook demo>>
4. Wrap-up

Exploration: N data sources into 1 graph



(This) business department may not be allowed to use exact location and facebook data for recommendations

Exploration: unauthorized edges



Some users may not be allowed to traverse edges from Store2

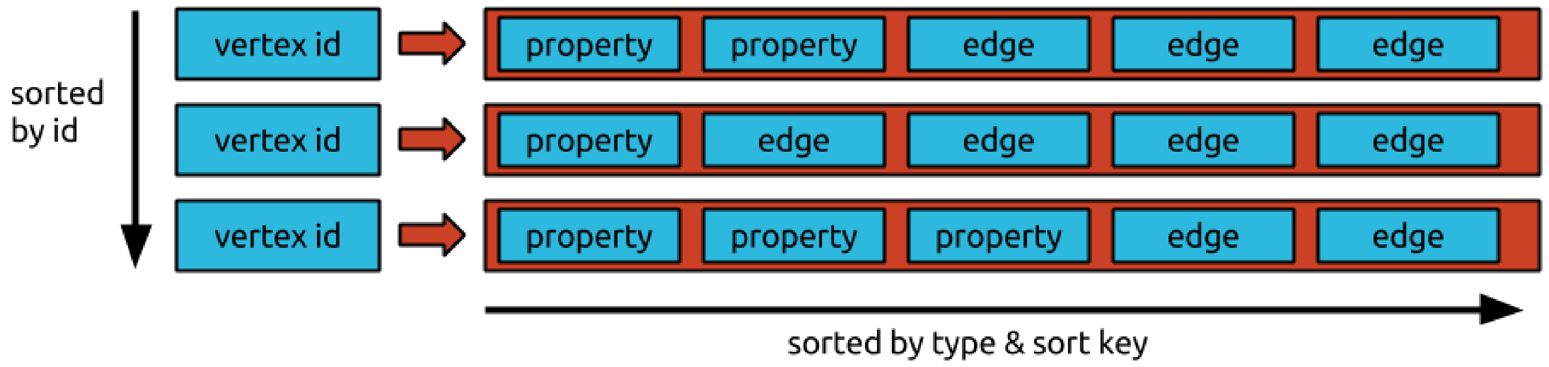
Differentiated access control to graph data

1. Exploration
2. Directions
   - separate graph stored per user group
   - datastore with cell-level security
   - filtering while traversing the graph
3. Application to TinkerPop/JanusGraph
   << notebook demo>>
4. Wrap-up

Directions: separate graph stored per user group

Criterion	one graph for all	graph per user group
#management processes	+ limited	○ scales with #groups
available (cache) memory	+ exclusive	○ divided between groups
CPU efficiency	○ authorization processing	○ support additional I/O
network I/O efficiency	+ data shared	○ no sharing
disk I/O efficiency	+ data shared	○ no sharing
resilience wrt corruption	○ everyone or no one	+ just one graph
scalability #user groups	+ not needed	○ limited

Directions: datastore with cell-level security



- need cell-level security to have the data store honor user authorizations
- cell-level user authorizations not implemented
  in current JanusGraph and Neo4j data formats

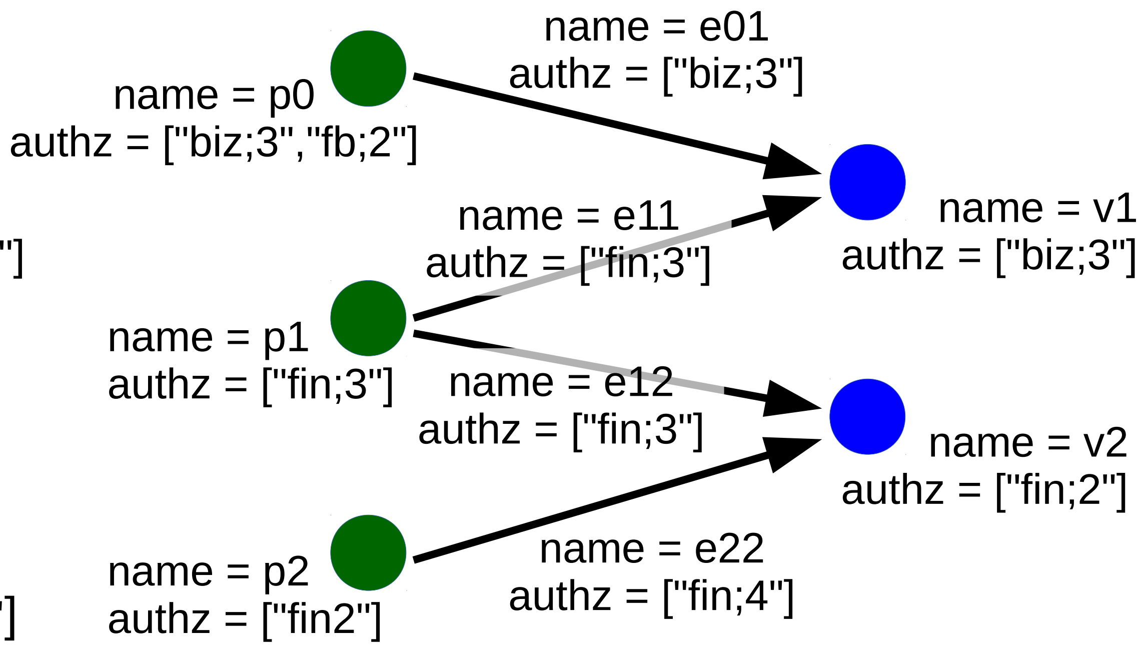
  https://docs.janusgraph.org/0.3.1/data-model.html
  http://key-value-stories.blogspot.com/2015/02/neo4j-architecture.html

Directions: filtering while traversing the graph [1/2]

user 1
authz = ["biz;1","biz;2","biz;3","fb;1"]

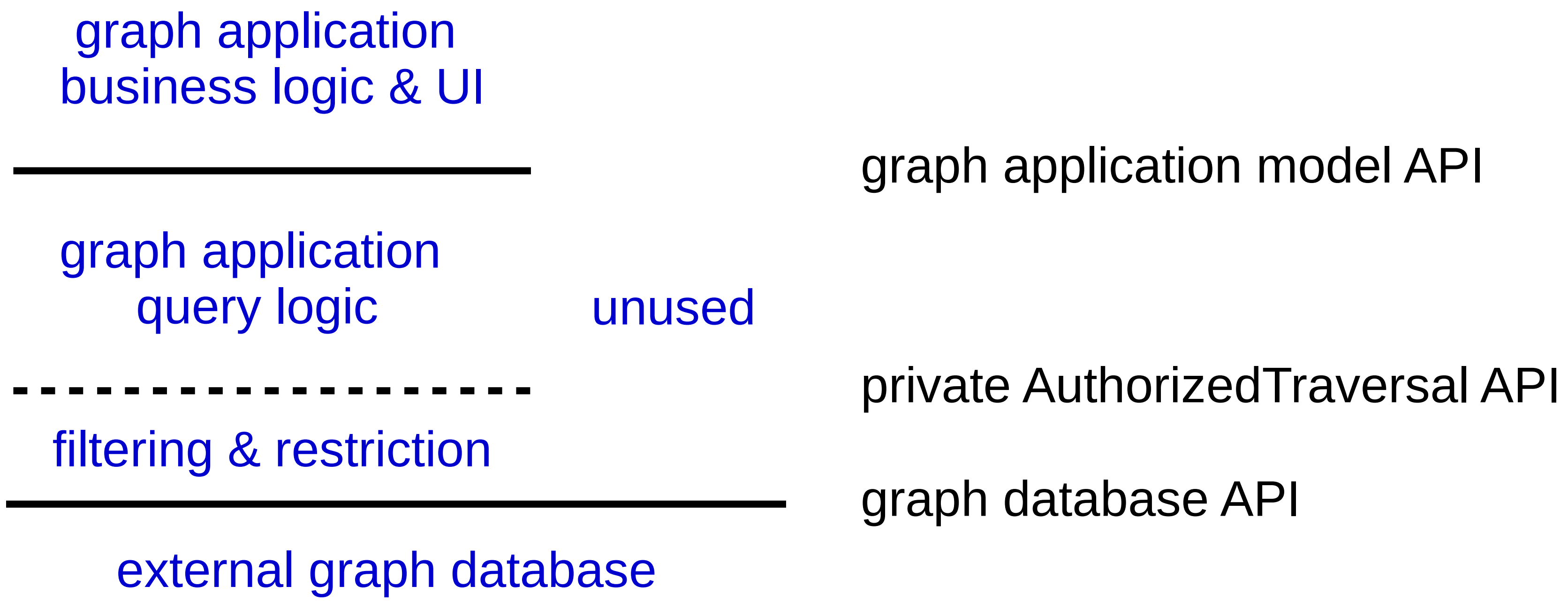
user 2
authz = ["biz;1","fin;1","fin;2","fin;3"]

Authorizations assigned to users



Authorization options for element access

Directions: filtering while traversing the graph [2/2]

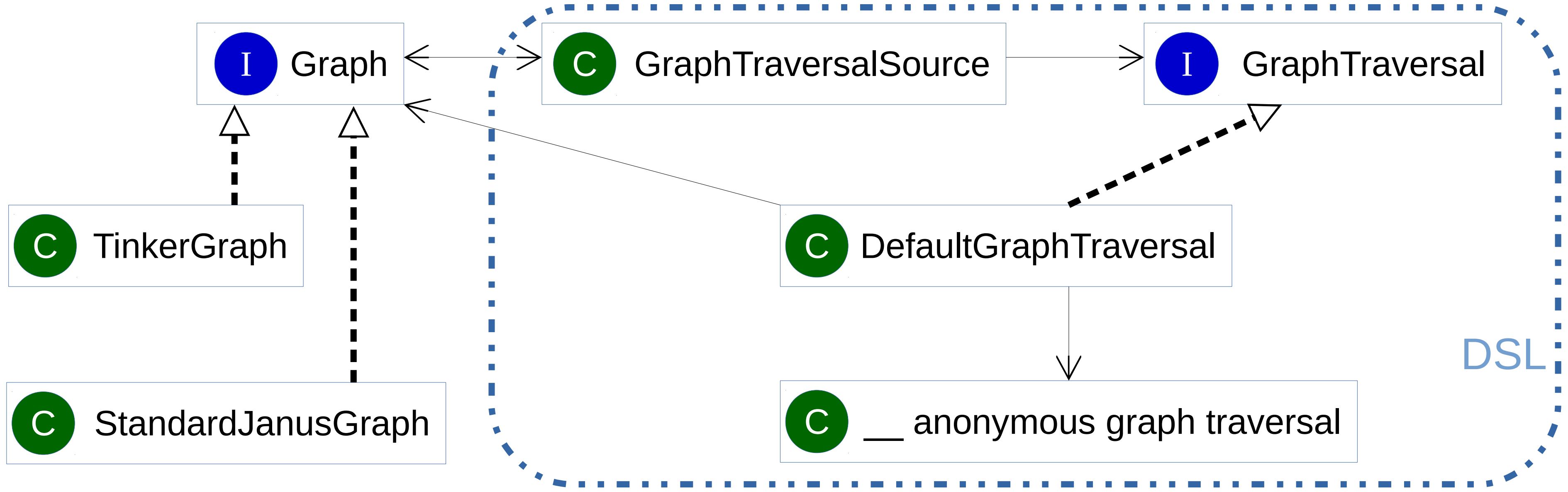

graph application business logic & UI
graph application model API
graph application query logic
unused
private AuthorizedTraversal API
filtering & restriction
graph database API
external graph database

Correctly honoring user authorizations as a separate concern

Differentiated access control to graph data

1. Exploration
2. Directions
3. Application to TinkerPop/JanusGraph
   << notebook demo>>
4. Wrap-up

Application to TinkerPop: java-gremlin DSL



AuthorizedTraversalSource extends GraphTraversalSource:

- a java-gremlin DSL on top of the TinkerPop APIs
- restricts the TinkerPop APIs to authorized data access
  (this needs a few instances of stack inspection, which is fragile)

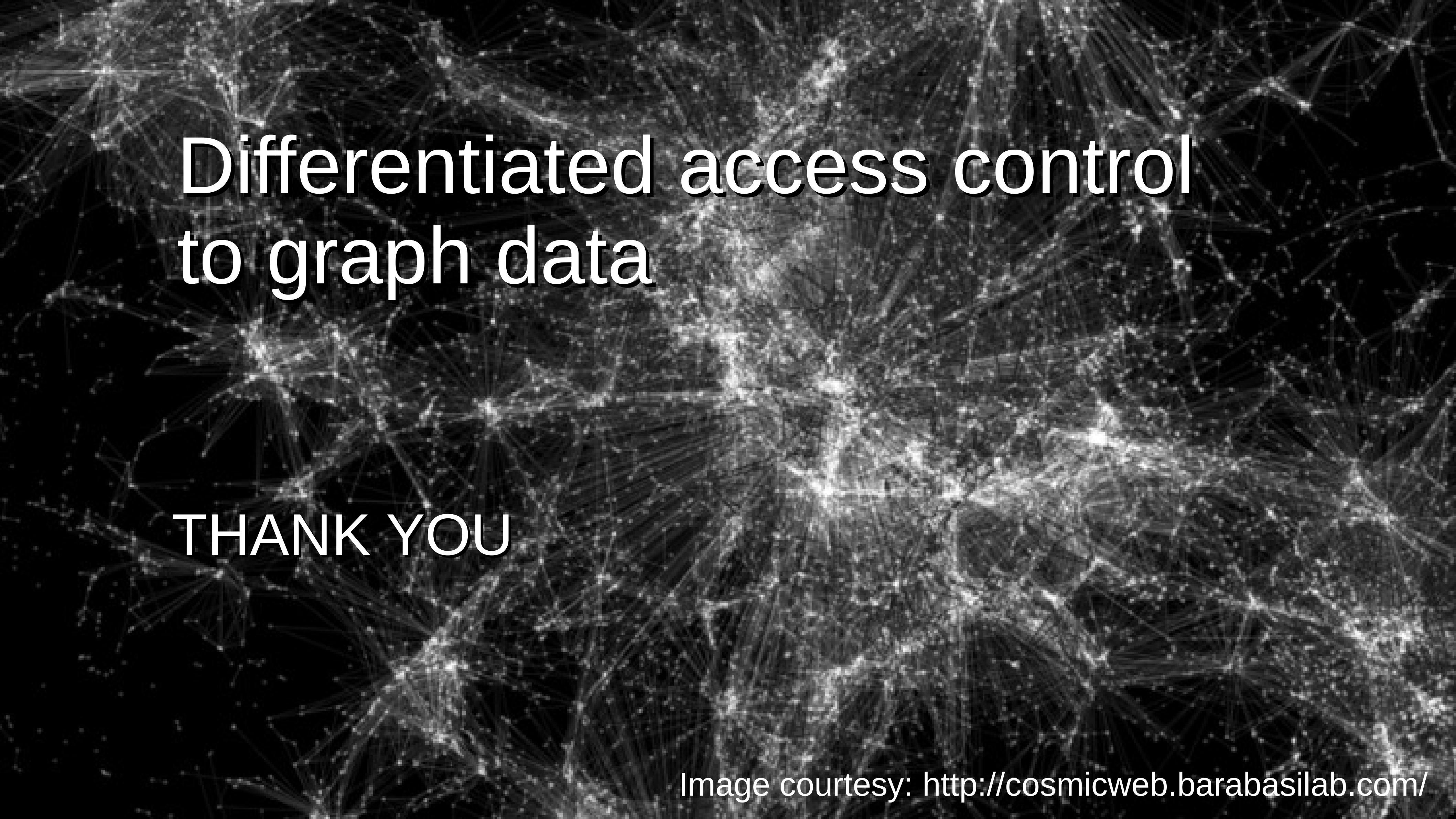
Application to TinkerPop: notebook demo

```
userAuthz = ["biz;1", "biz;2", "biz;3"]

graph.traversal().
    V().has("authz", within(userAuthz)).has("name", "Mathilde").
    outE("likes").has("authz", within(userAuthz)).
    inV().has("authz", within(userAuthz)).
    outE("lives").has("authz", within(userAuthz)).
    inV().has("authz", within(userAuthz)).has("city", "Brussels")

graph.traversal(AuthorizedTraversalSource.class).
    withAuthorization(userAuthz).
    V().has("name", "Jane").
    out("likes").
    out("lives").has("city", "Brussels")
```

https://github.com/vtslab/janusgraph/tree/fosdem2019/fosdem2019

Wrap-up

1. Right visibility of sensitive graph data to different user groups is not easy to achieve
2. Separate graphs per user group result in penalties for performance and maintenance
3. Cell-level security is not part of data format of current graph databases
4. Filtering while traversing the graph is feasible – if fragile – provided that it is done within the context of a secure endpoint



Differentiated access control to graph data

THANK YOU

Image courtesy: http://cosmicweb.barabasilab.com/