

Information Flows

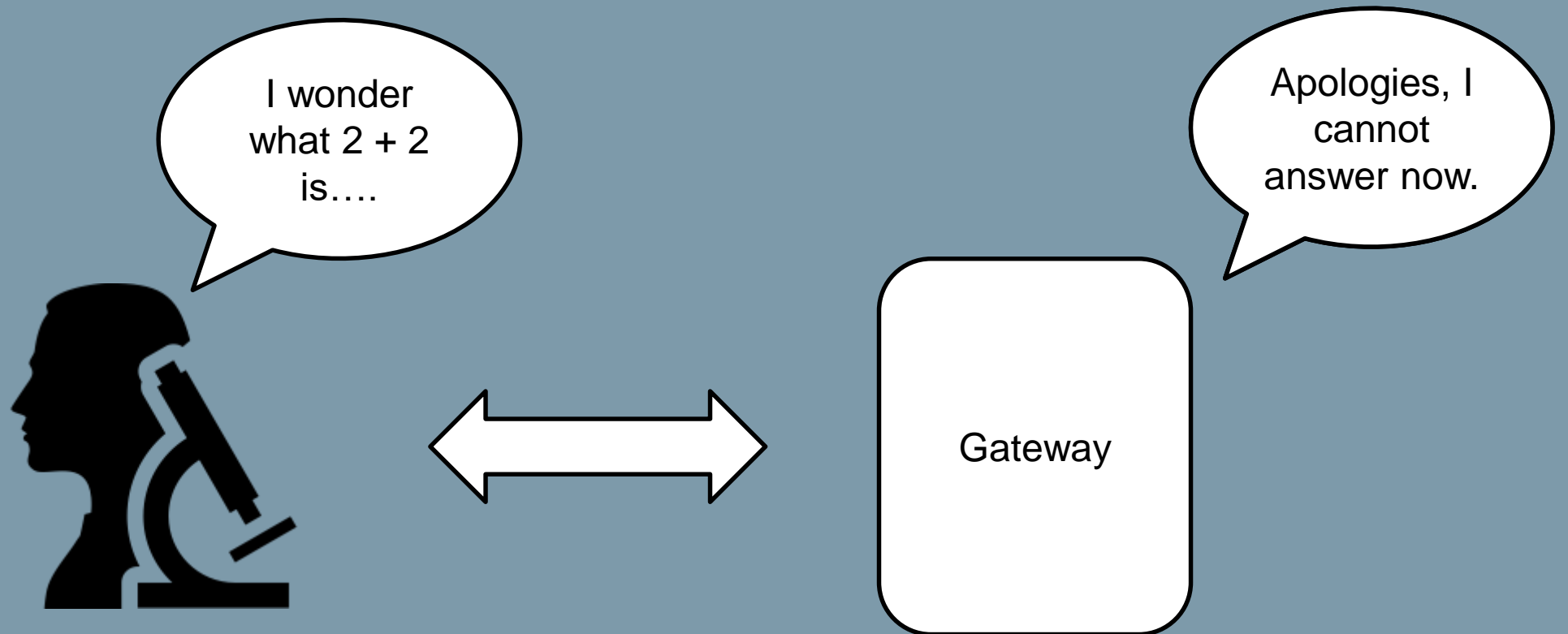
Gabriele Pierantoni, Dermot Frost, Sandra Gesing, Silvia Olabarriaga,
Mahdi Jaghoori, Gabor Tertyansky and Junaid Arshad



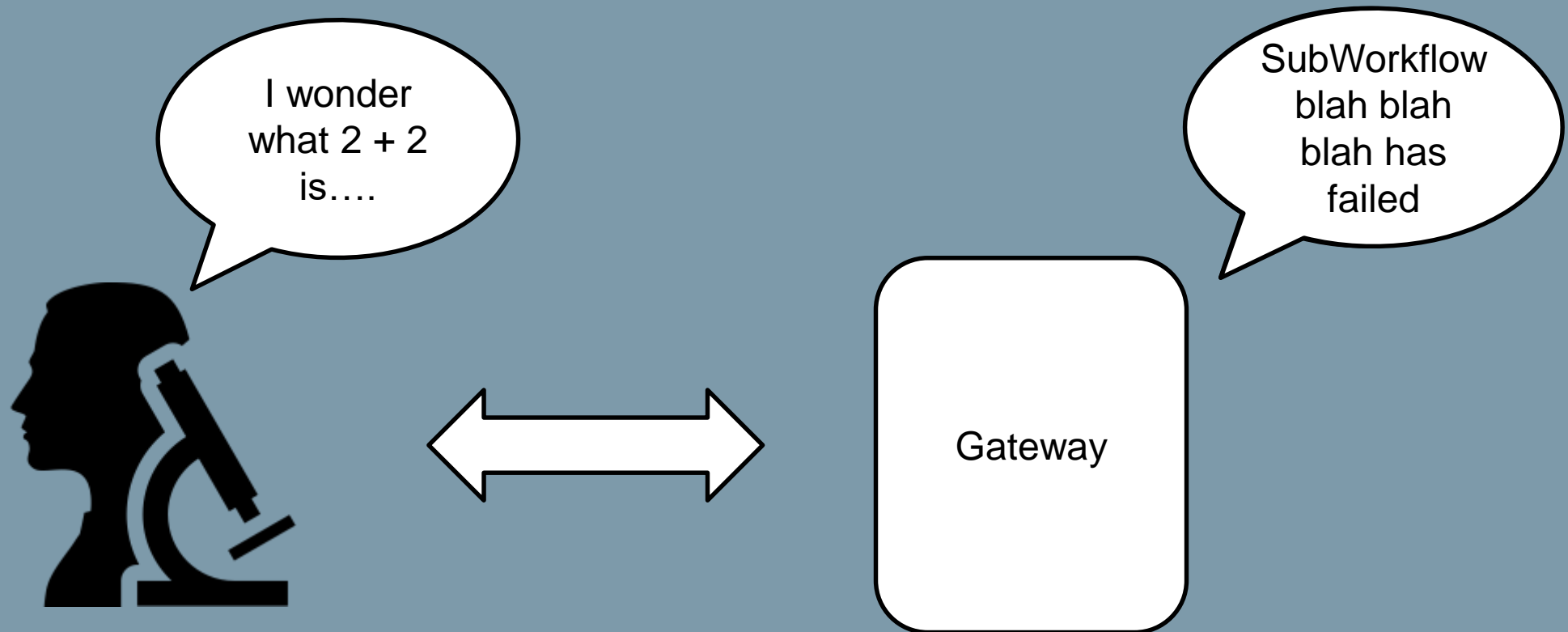
Information Flows

- This paper is part of an ongoing discussion between TCD, UOW, AMD and Notre Dame.
- An initial paper on user profiles: “***Scientific Workflow Management – For Whom?***” in 2014 IEEE 10th International Conference on e-Science
- A lightning talk in Budapest (2015)
- A presentation in Rome (2016) – Proceedings published on IWSG16
- A PeerJ Paper under review

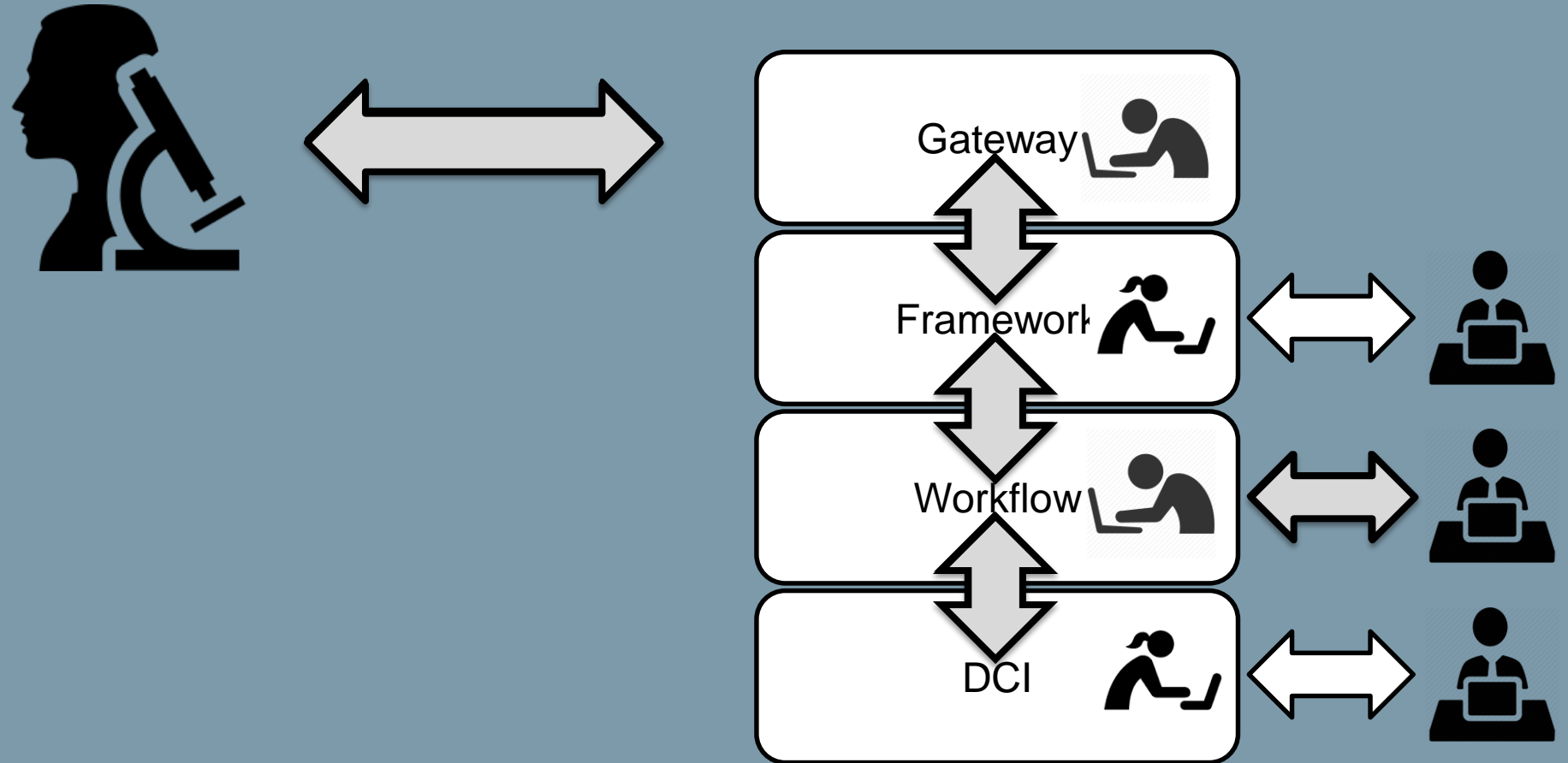
A Science Gateway (Ideally)



A Science Gateway (In reality)



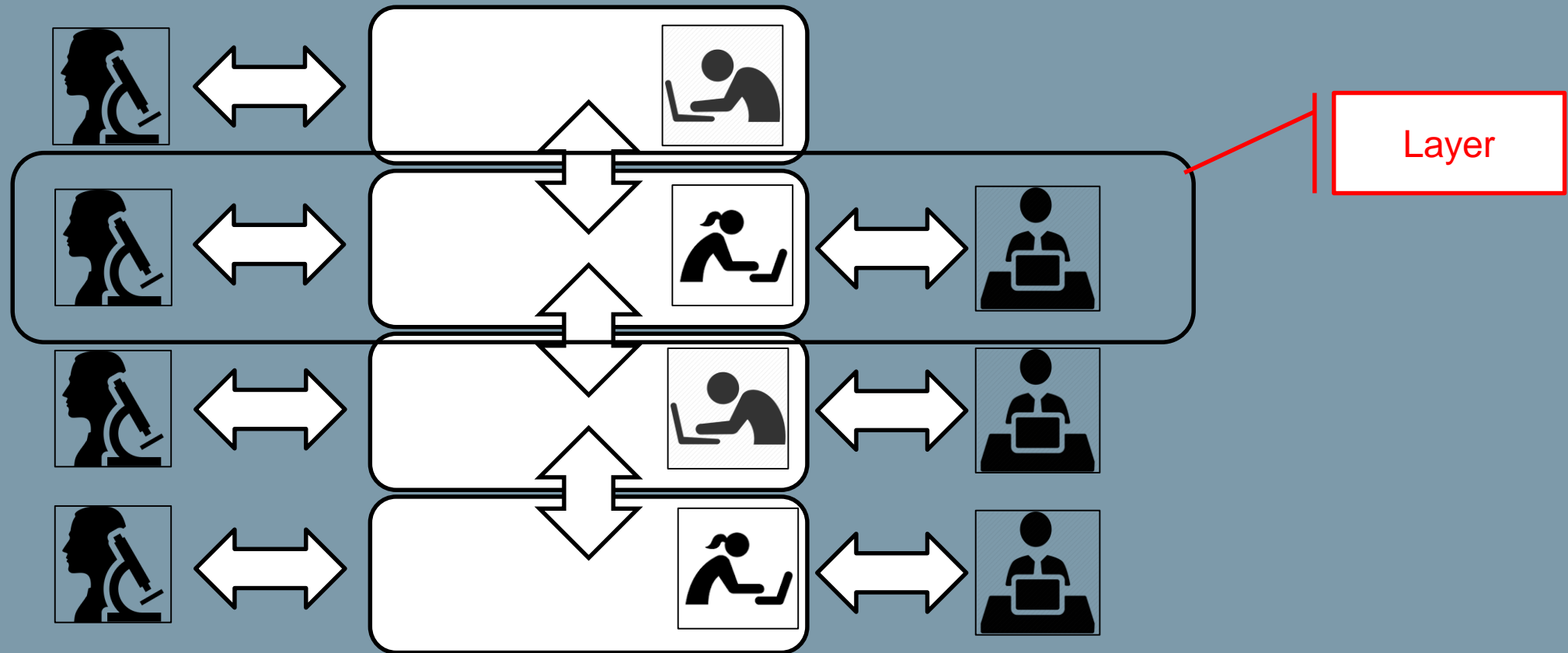
A Science Gateway (In detail)



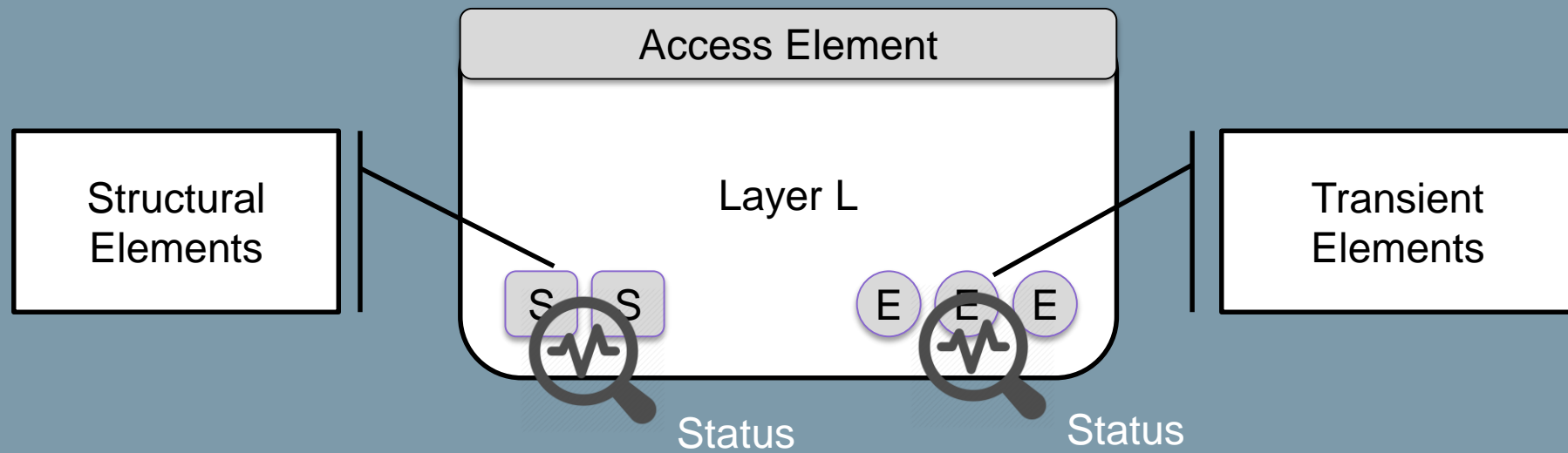
Information Flows

- Different solutions have been devised to overcome the complexity of these information flows
- If we had a comprehensive model, we could “understand” these solutions better
- If we had a comprehensive model, we could “suggest” solutions on the medium and long term

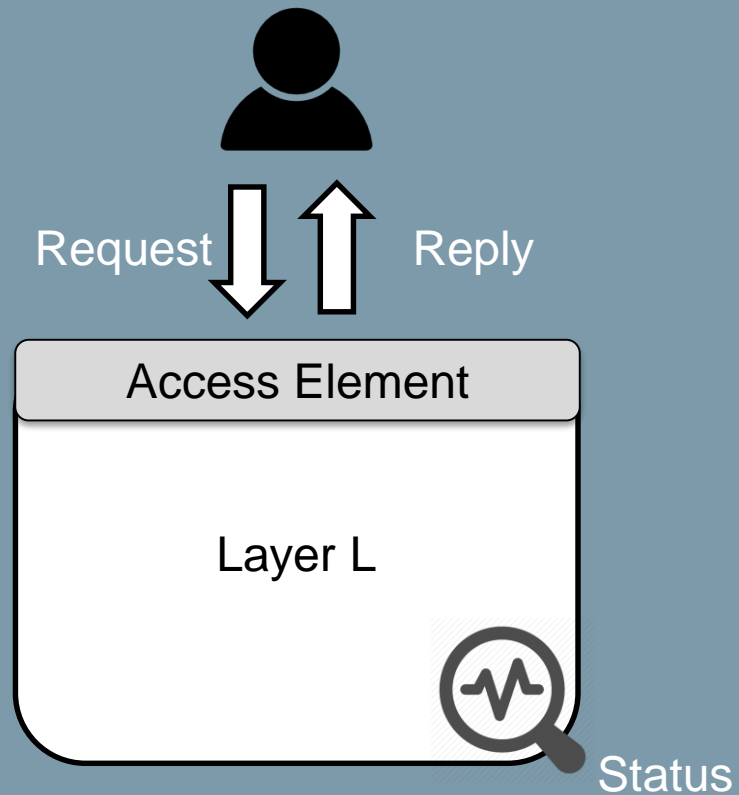
How to model Information Flows

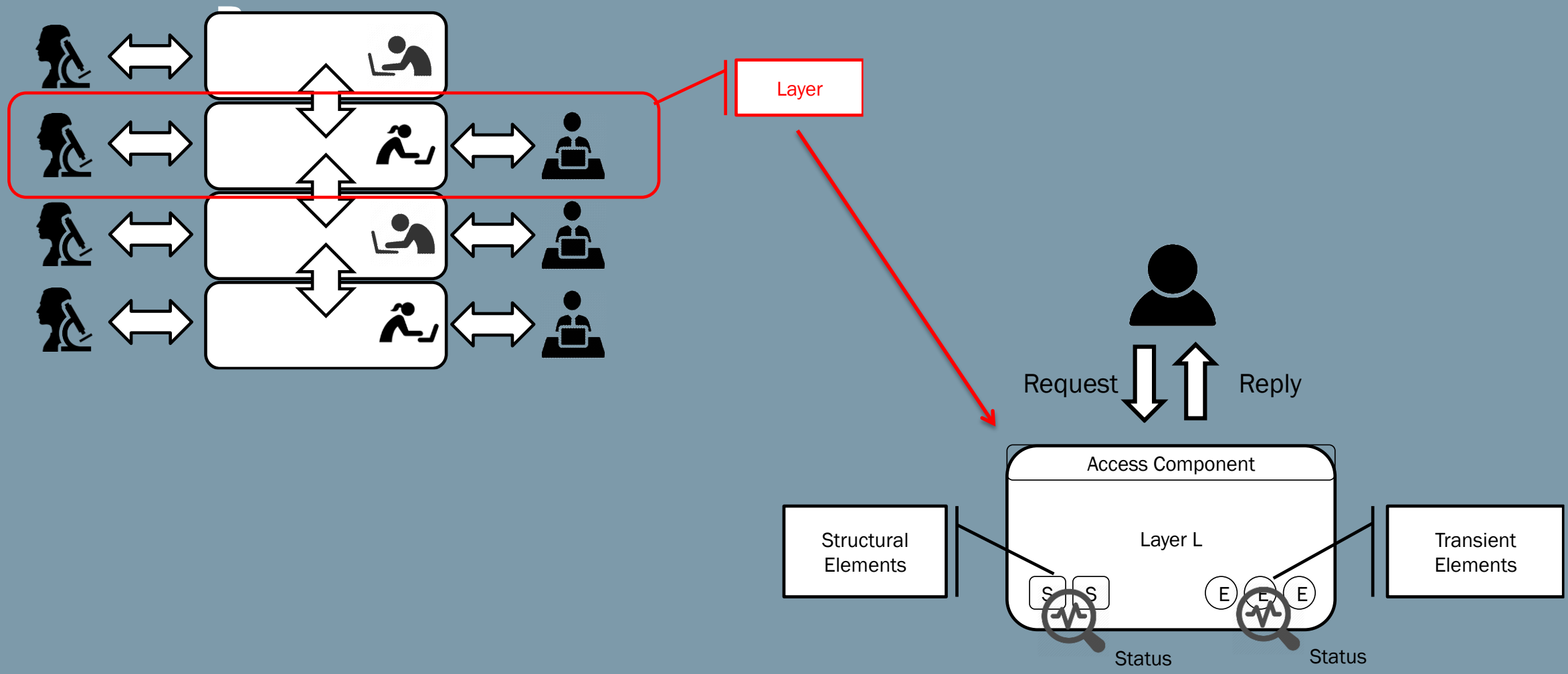


The layer – Inner Elements

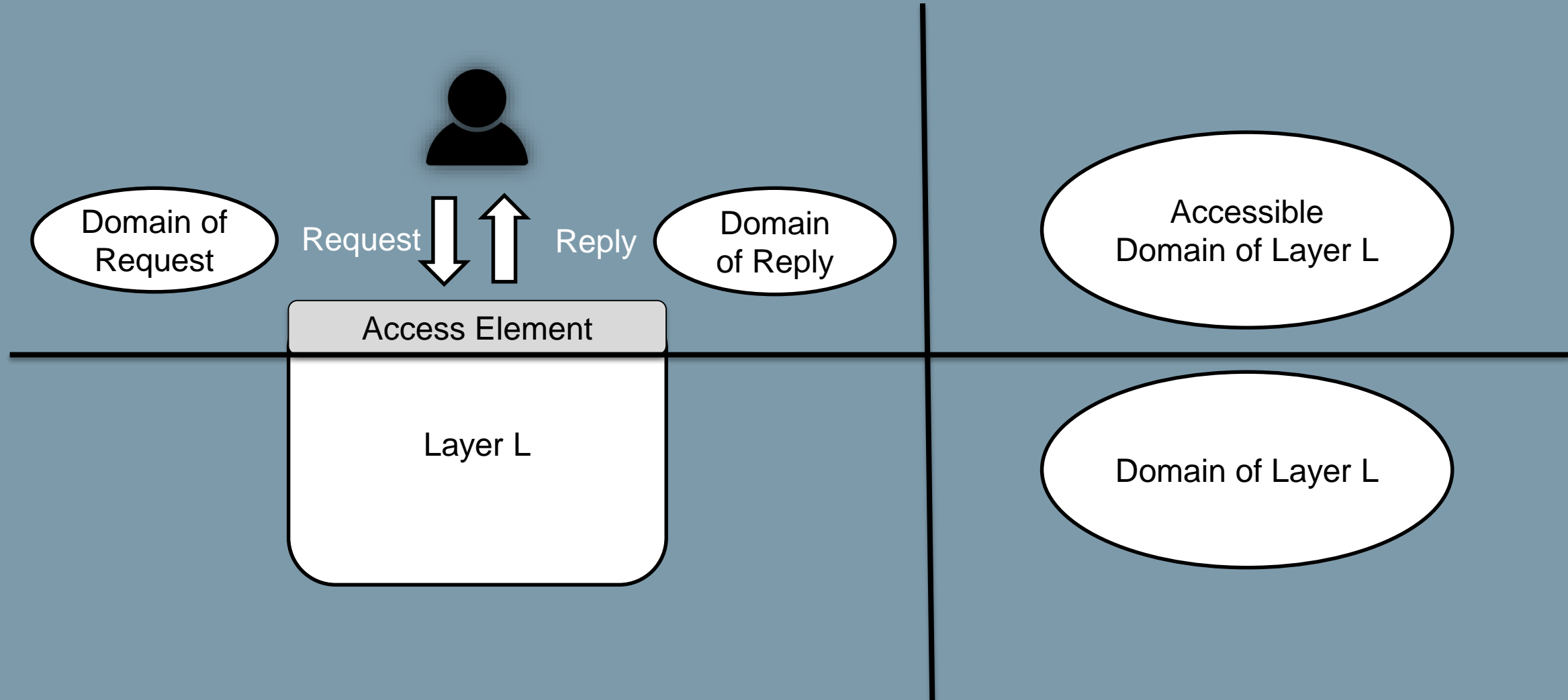


The layer – Interactions with external actors



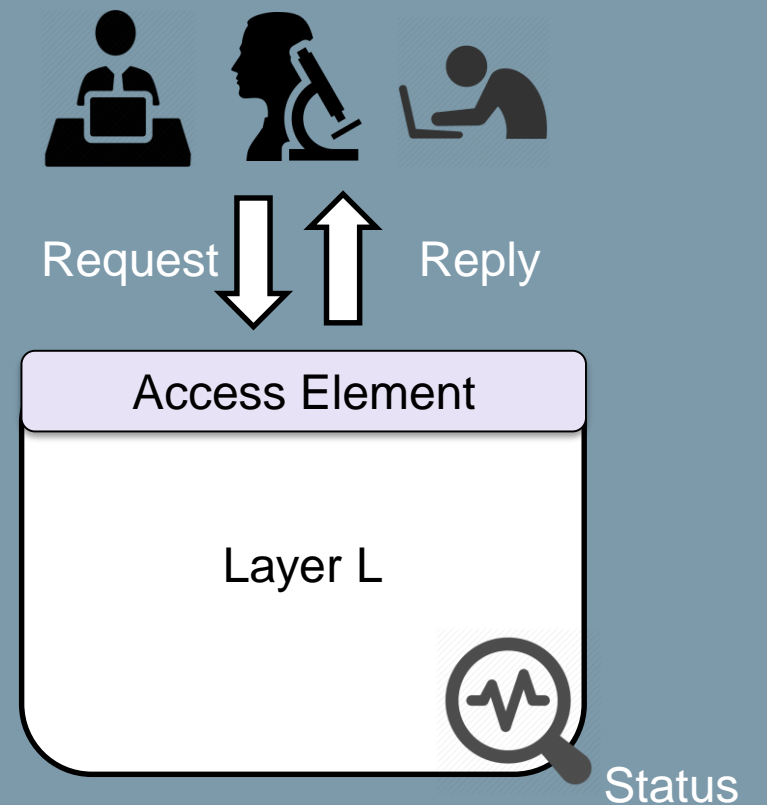


The layer - Domains



The layer – Actor Profiles

- Result-Oriented
- Layer-Oriented
- Development-Oriented



Value to the user

- Actions issued by Requests and their results returned with Replies have different values depending on the user profile.
- Also actions and results are more or less difficult to issue and to read depending on the user profiles
- We model this with Utility, Cost and Value: $\text{Value} = \text{Utility} - \text{Cost}$
- This part of the model is still unclear

Four main challenges (1/2)

– **Heterogeneous Information in one or more Replies**

To address it we have to reduce the heterogeneity within the Domain of the Reply(ies) thus making it easier to access for the various user profiles. (e.g. separate results from error codes)

– **Incomplete Information accessible through one or more Replies**

To address it, we have to extend the information domain directly available to the user. (e.g. automatically retrieve additional information if a job fails)

Four main challenges (2/2)

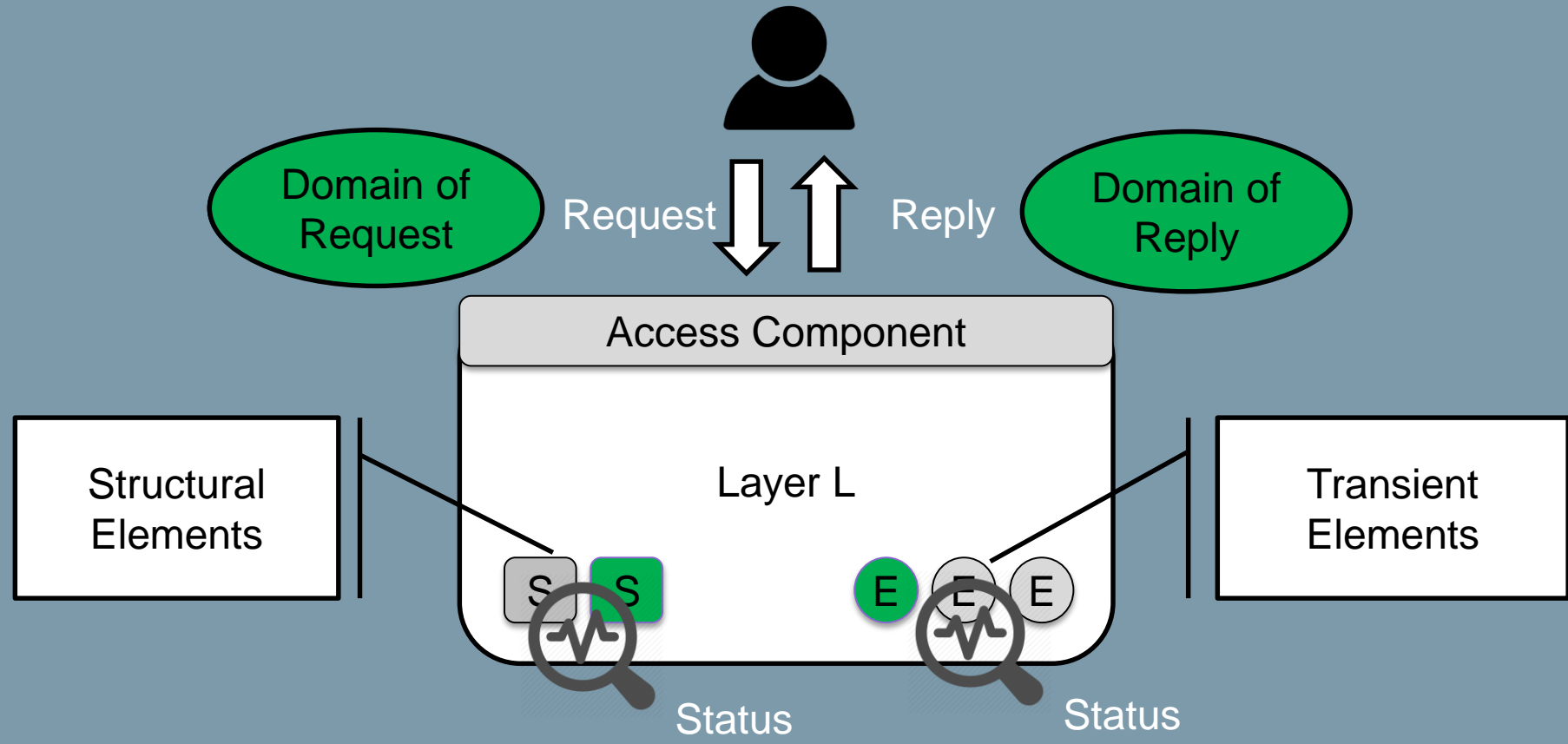
- **Compound Requests**

To address this, we need to automatically orchestrate different Requests.
(e.g. test jobs in case of failure)

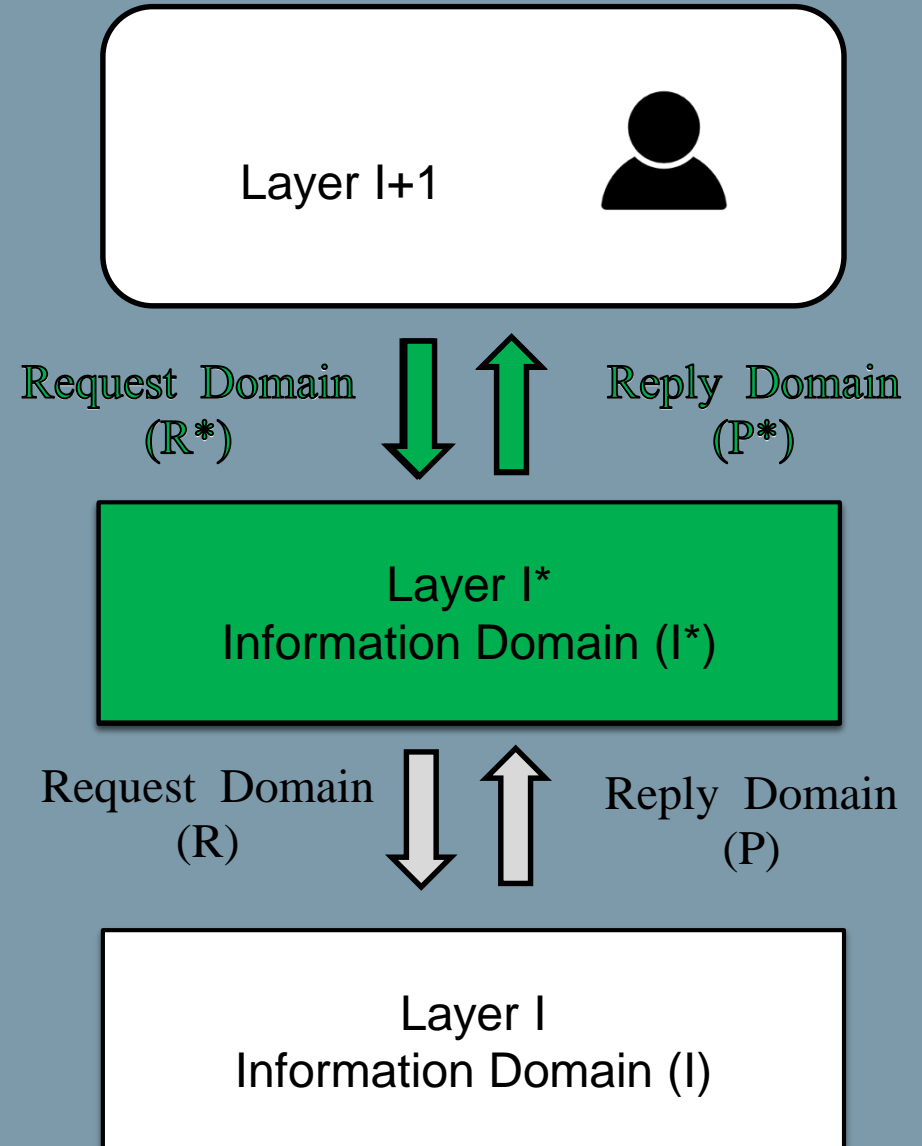
- **Interoperability**

To address this, we need to offer a unified interface to multiple
implementation of the same layer. (e.g SHIWA, DCI Bridge, CloudBroker)

Solutions (1/2)

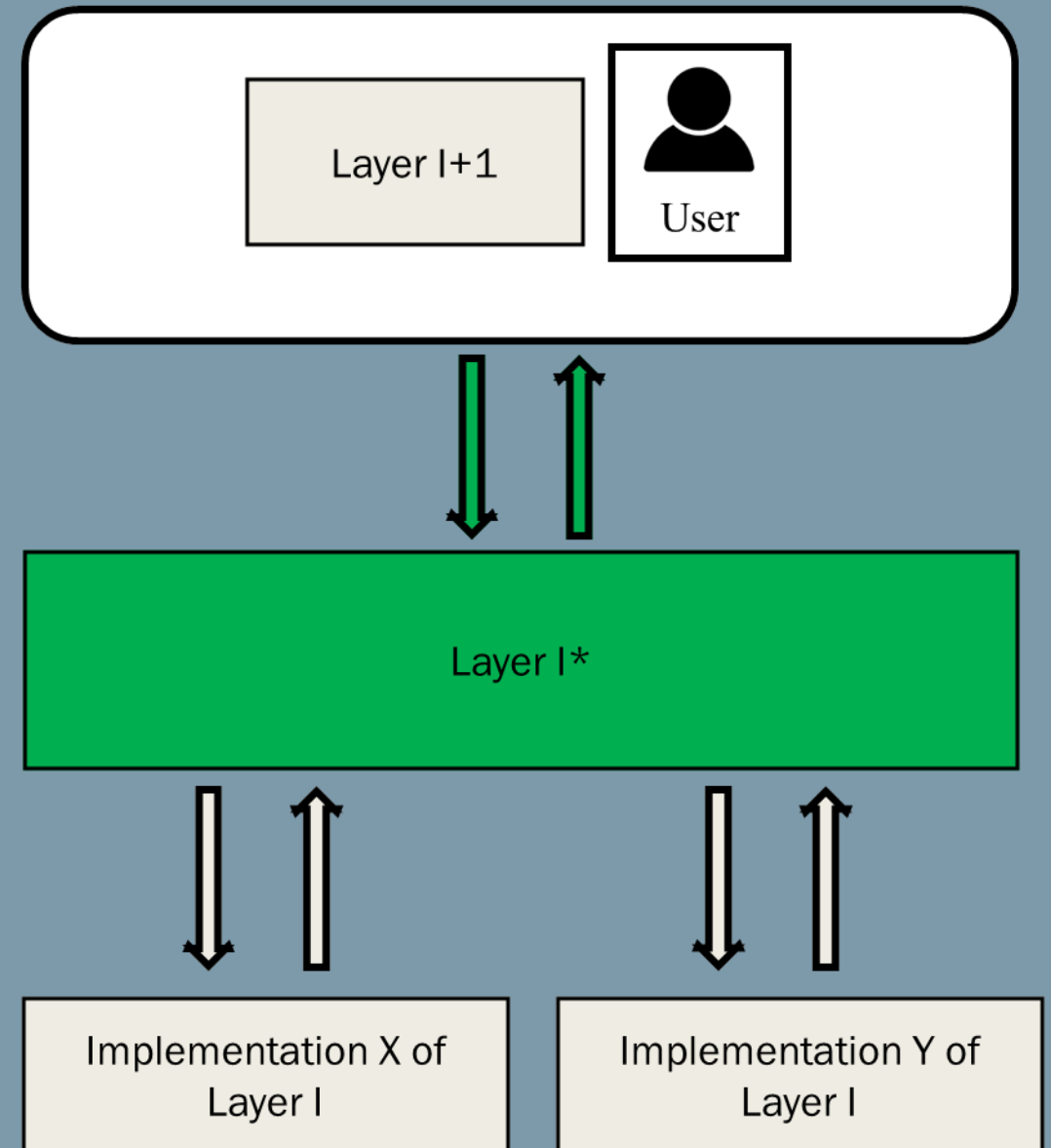


Solutions (2/2)



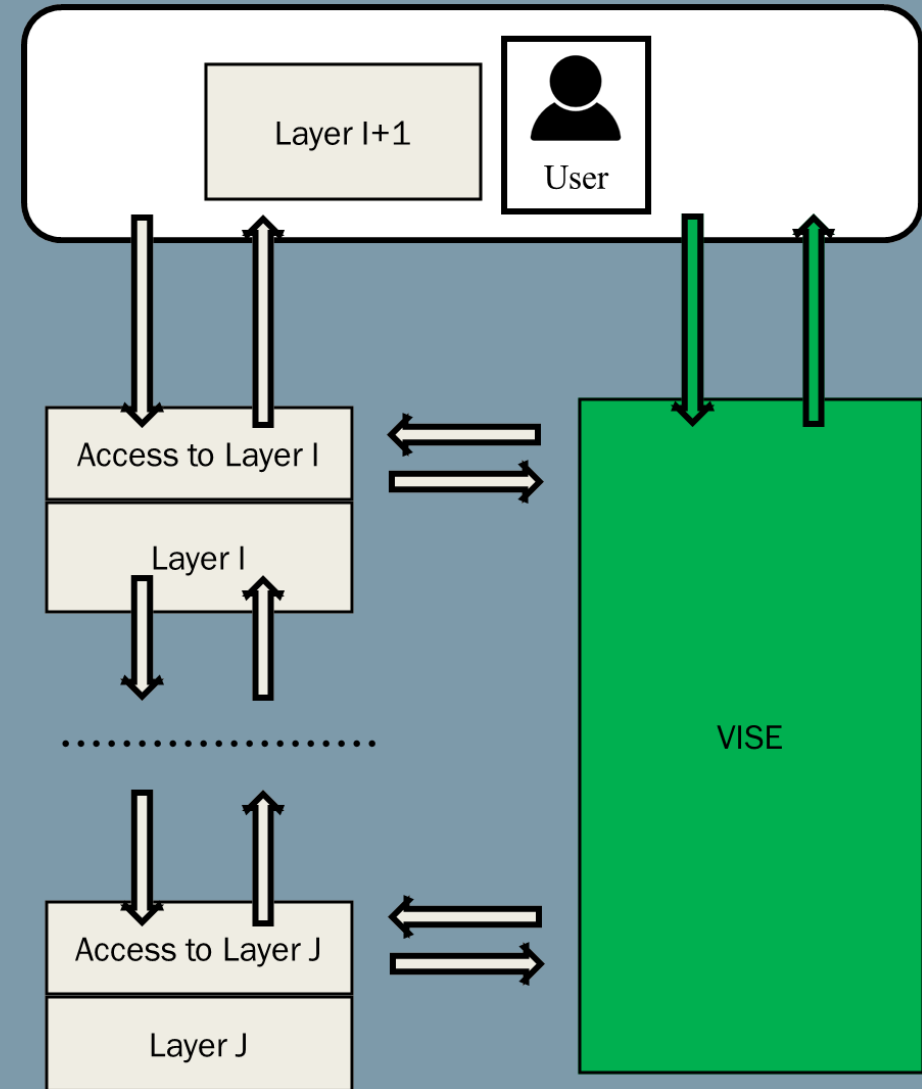
Examples (1/3)

- Interoperability:
 - SHIWA for Workflows,
 - DCI Bridge for DCIs
 - CloudBroker for Clouds



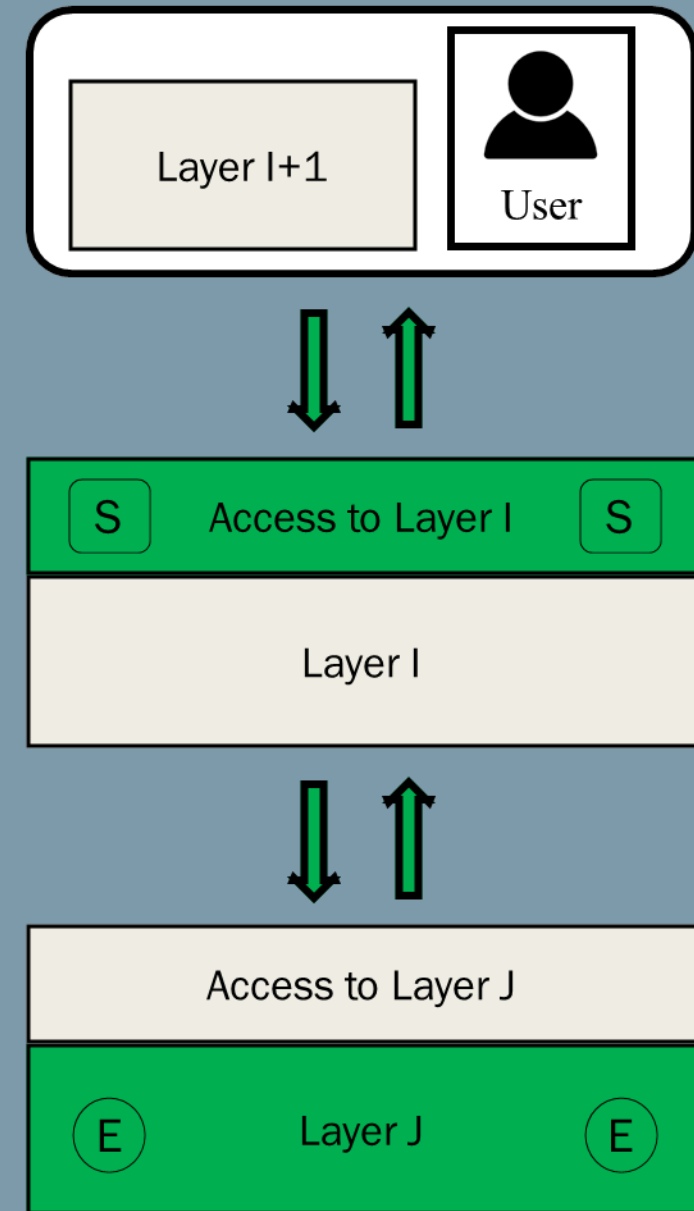
Examples (2/3)

- AMD developed a “**Processing Manager**” that spanned multiple layers.



Examples (3/3)

- Metaworkflows and rapid GUI development for HELIOPhysics



Where we are now

Future work

- Model applicable for WS-PGRADE/SHIWA/Helio
- Proof of concept for further technologies such as Galaxy, Hubzero, ...
- Further use cases
- Formalization of model
- Investigate what long term solutions (information tagging, separation of concerns) can be envisaged for more structural solutions.

Acknowledgements

- ER-FLOW, SCI-BUS and HELIO projects
- Prof. Kacsuk
- Prof. Herres-Pawlis
- Dr. Sciacca
- Dr. Becciani,
- Dr. Castelli
- Dr. Taffoni

Any questions ?

