

---

# CONTINUOUS WORKFLOWS AND CACHING

**Eliot McIntire, Ceres Barros, Ian Eddy**

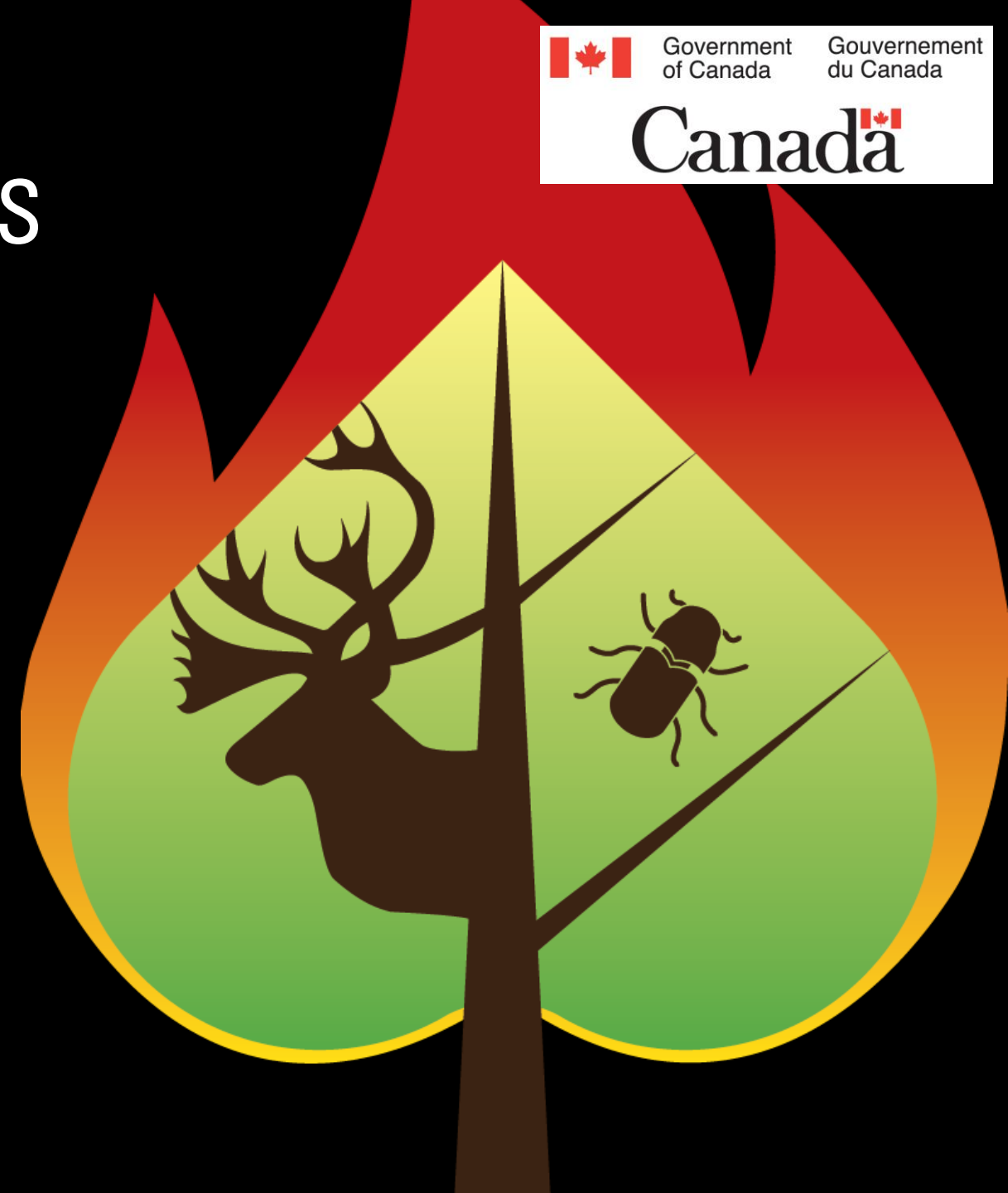
Research Scientists,

Pacific Forestry Service, Canadian Forest Service

SpaDES virtual workshop

---

June 11-13<sup>th</sup>, 2024



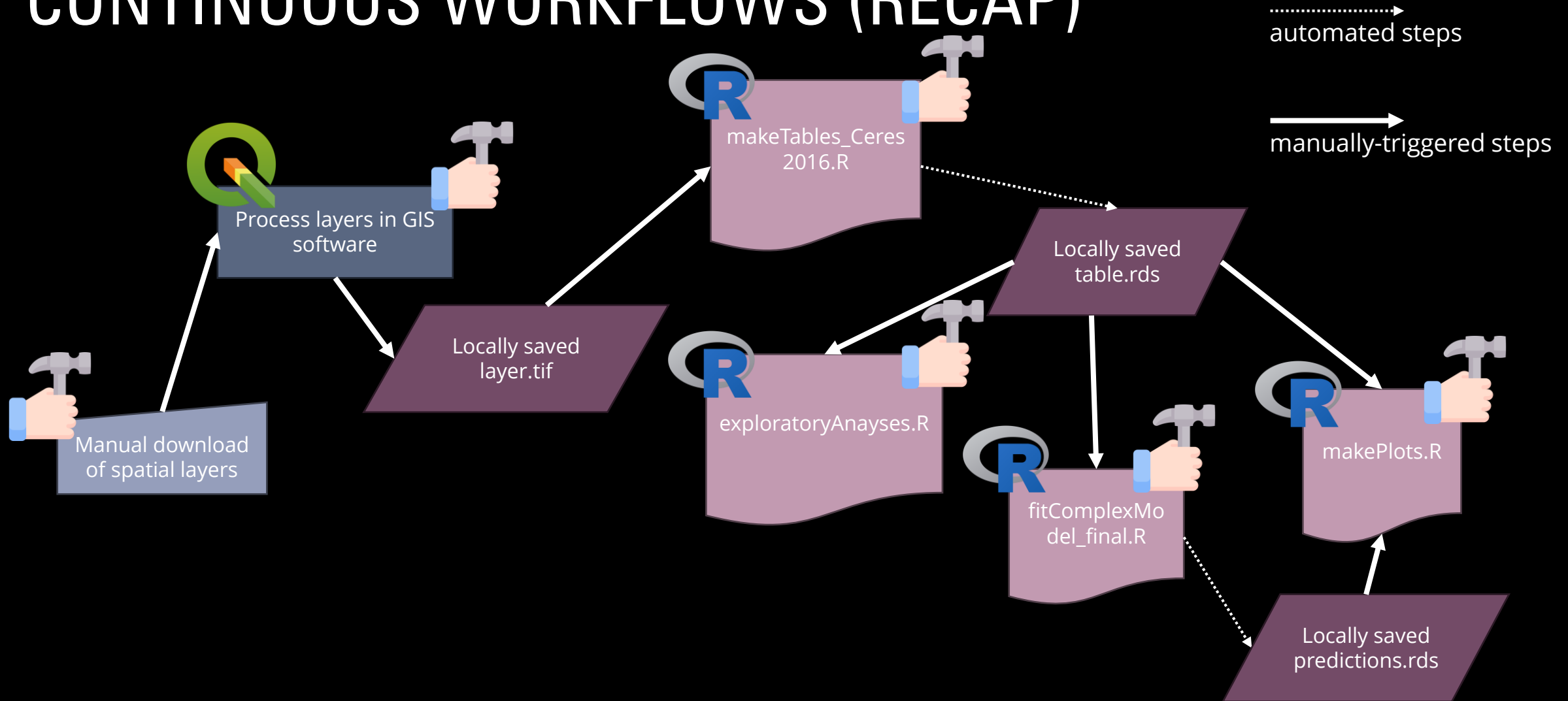
---

# OUTLINE

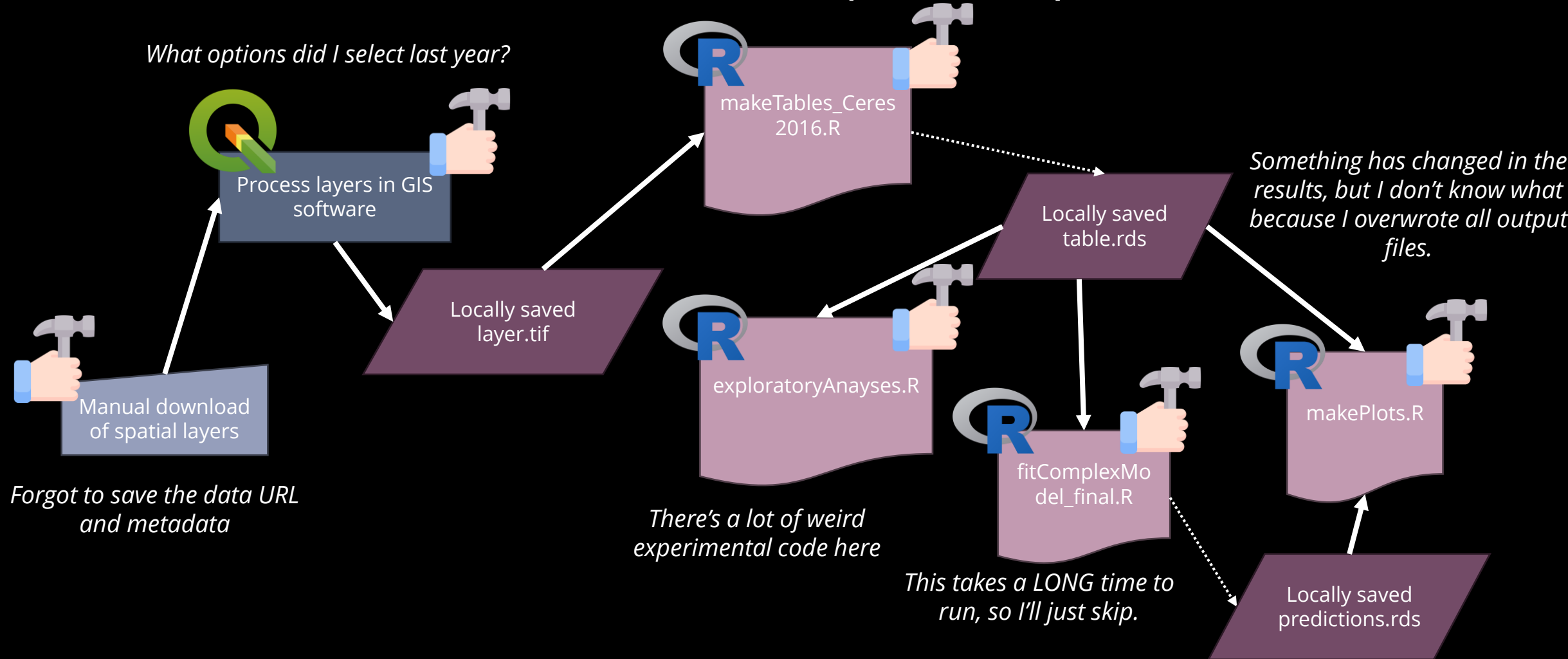
1. Continuous workflows with SpaDES
2. Caching



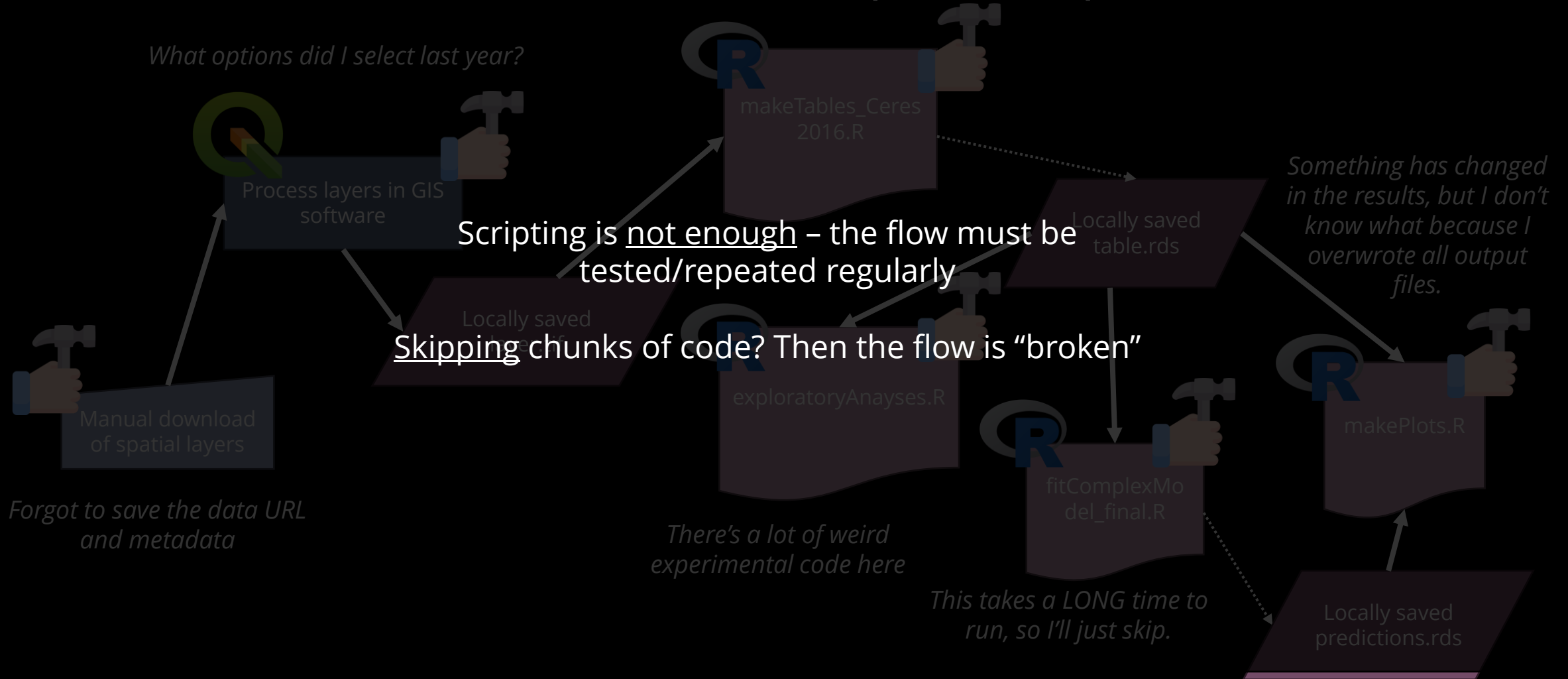
# CONTINUOUS WORKFLOWS (RECAP)



# CONTINUOUS WORKFLOWS (RECAP)



# CONTINUOUS WORKFLOWS (RECAP)

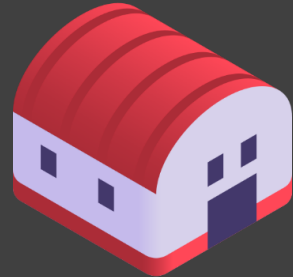


---

# CONTINUOUS WORKFLOWS (RECAP)

Self-contained

All steps, from processing *raw data* to producing *final figures* are **integrated and automated\***



Data is FAIR  
(Wilkinson *et al.* 2016)

Final outputs can be repeated  
and are integrated in the  
reporting\*\*

\*as much as possible

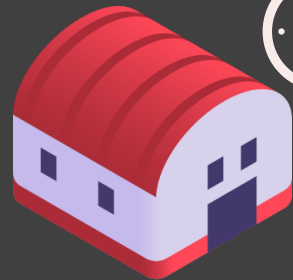
\*\*directly, or indirectly via links

---

# CONTINUOUS WORKFLOWS (RECAP)

Self-contained

All steps, from processing *raw data* to producing *final figures* are **integrated and automated**



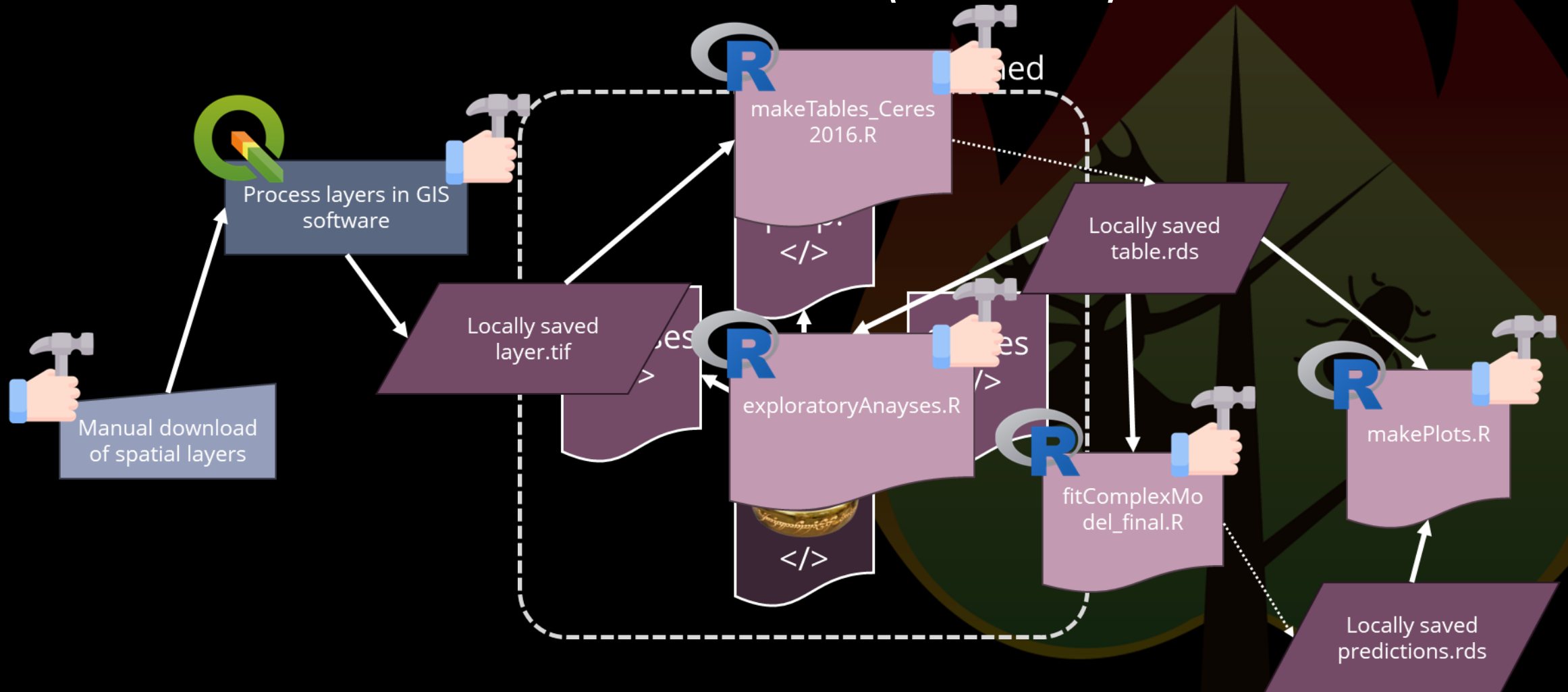
Data is FAIR  
(Wilkinson *et al.* 2016)

Final outputs can be repeated  
and are integrated in the  
reporting\*\*

Requires re-run tolerance  
(functions and workflow)

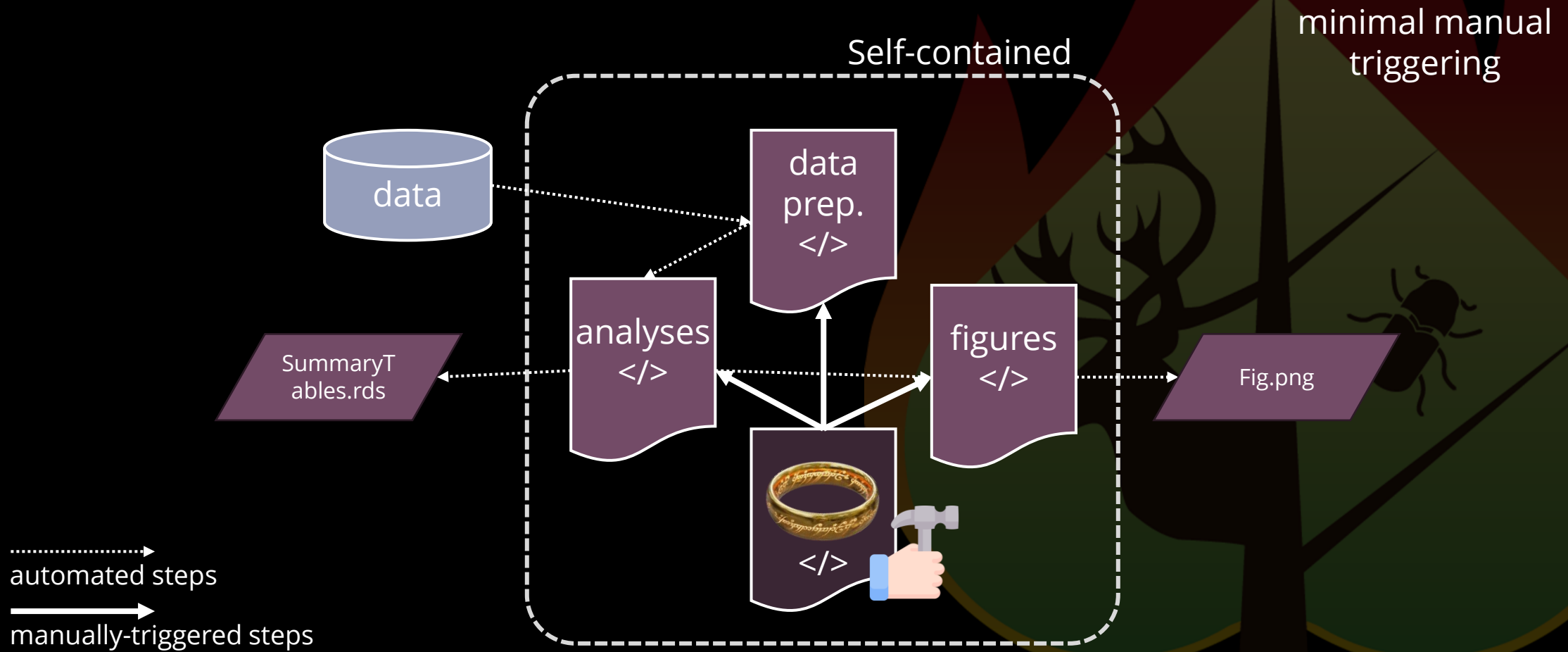
\*\*directly, or indirectly via links

# CONTINUOUS WORKFLOWS (RECAP)

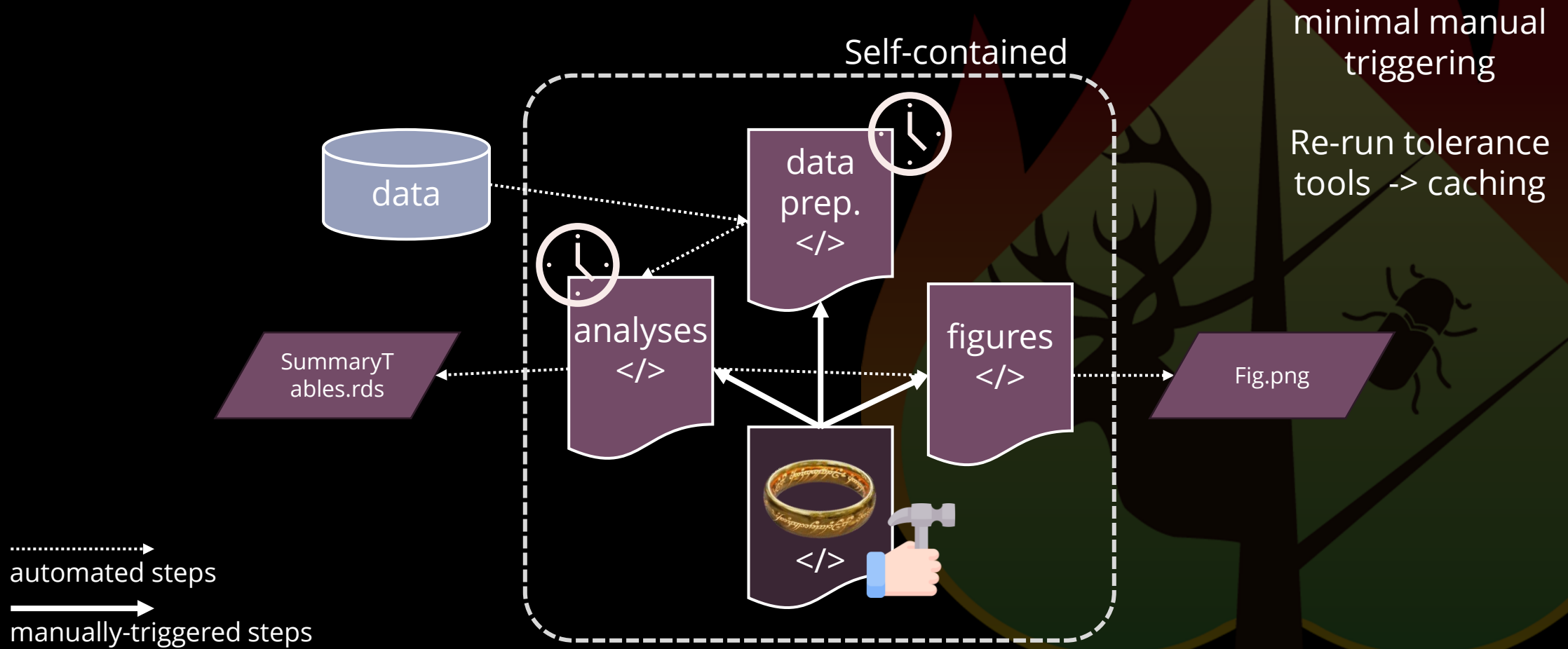




# CONTINUOUS WORKFLOWS (RECAP)



# CONTINUOUS WORKFLOWS (RECAP)



---

# CACHING

*"In computing, a cache (/kæʃ/ KASH) is a hardware or software component that stores data so that future requests for that data can be served faster; the data stored in a cache might be the result of an earlier computation or a copy of data stored elsewhere."*

[Cache \(computing\) - Wikipedia](#)

Very useful during model development

Caution -> periodic cache clears still needed  
(are you hitting a *bad* cache?)



---

# CACHING WITH REPRODUCIBLE

Caching in SpaDES workflows can happen at various levels

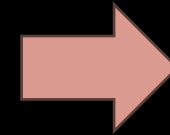
- Module level – all events are cached
- Event level – events are *selectively* cached
- Function level – any function can be cached

---

# CACHING WITH REPRODUCIBLE

Caching in SpaDES workflows can happen at various levels

- Module level – all events are cached
- Event level – events are *selectively* cached
- Function level – any function can be cached



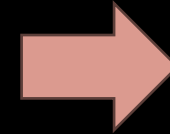
*Nested* caching

---

# CACHING WITH REPRODUCIBLE

Caching in SpaDES workflows can happen at various levels

- Module level – all events are cached
- Event level – events are *selectively* cached
- Function level – any function can be cached



*Nested* caching

```
reproducible::Cache()
```



# TIME TO LOOK AT SOME CODE



[Robust and nimble scientific workflows, using SpaDES](#)

[Workshop Agenda](#)

This is a *hands-on* workshop

Please **ask questions**, share your troubles and successes

**Last 15 min** of each WOYO is for discussion

We are always available via **Teams chat**