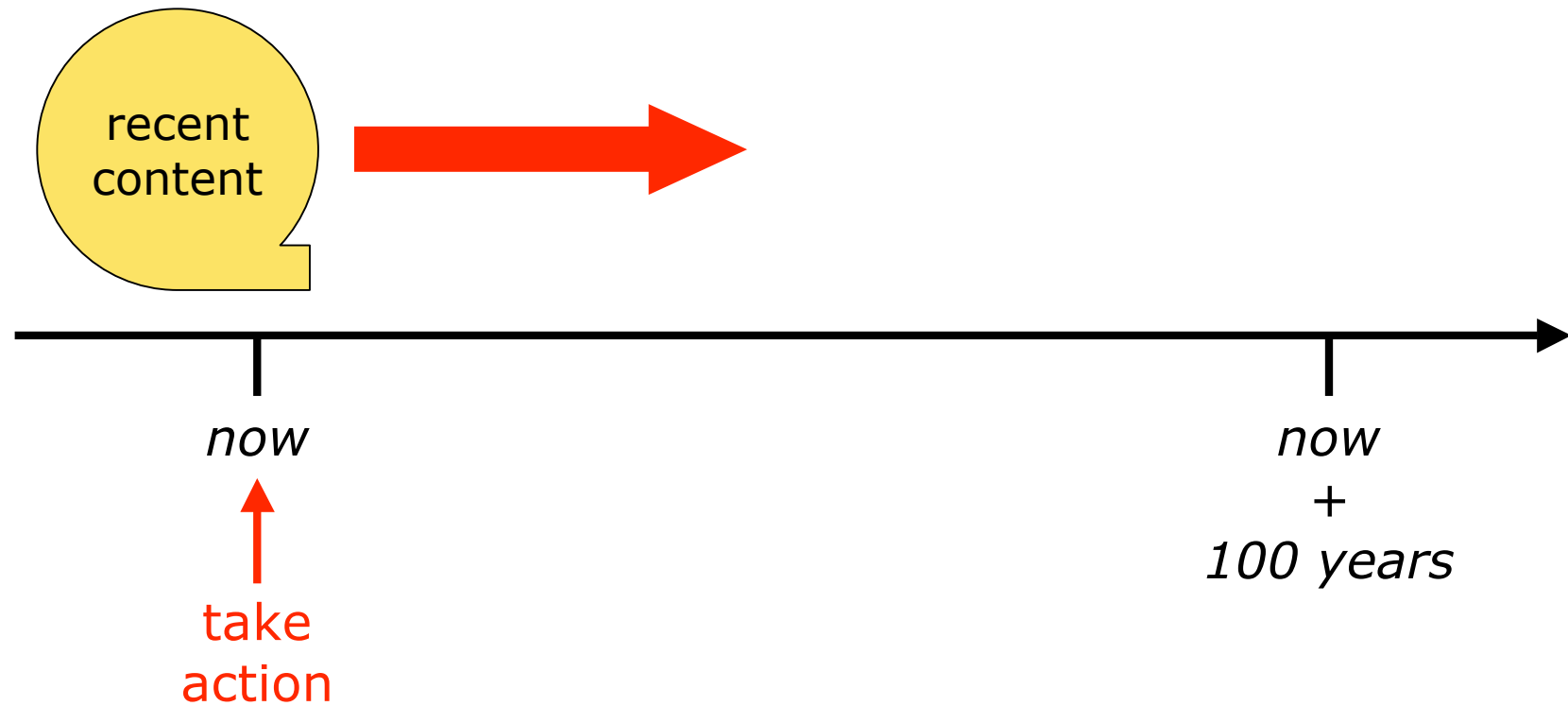


Long-term Preservation as a Relay

Greg Janée

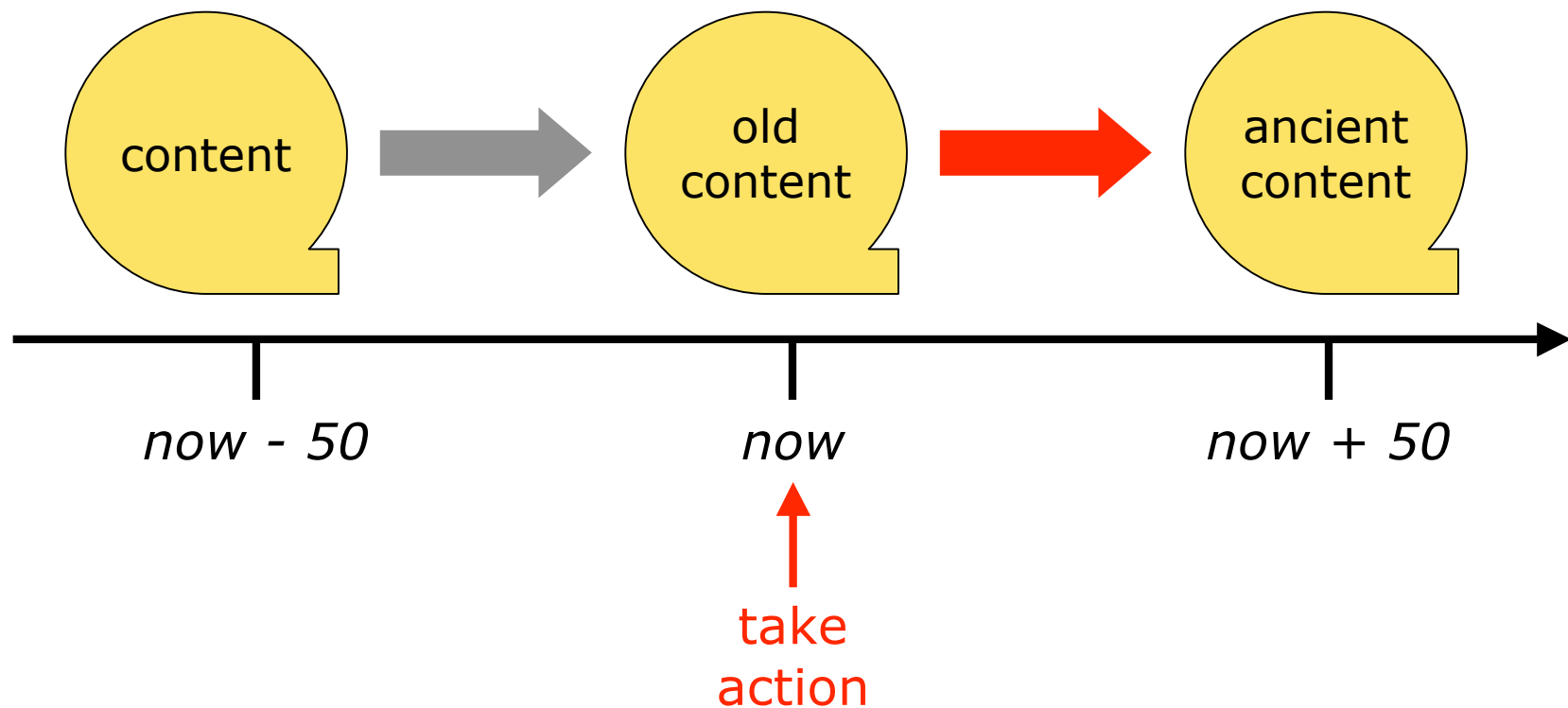
University of California at Santa Barbara

Typical preservation effort



Ein Gedankenexperiment

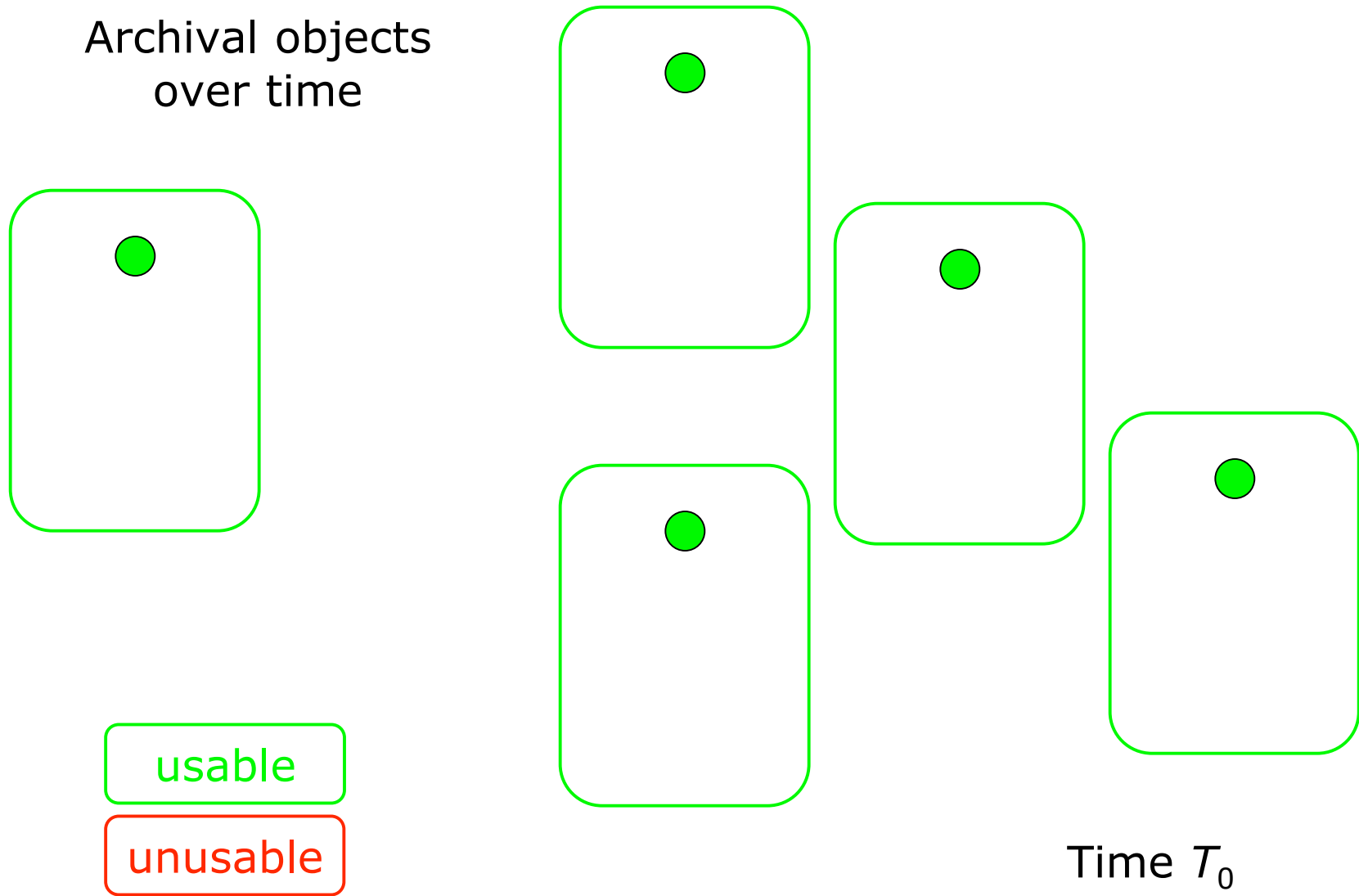
"mid-century perspective"

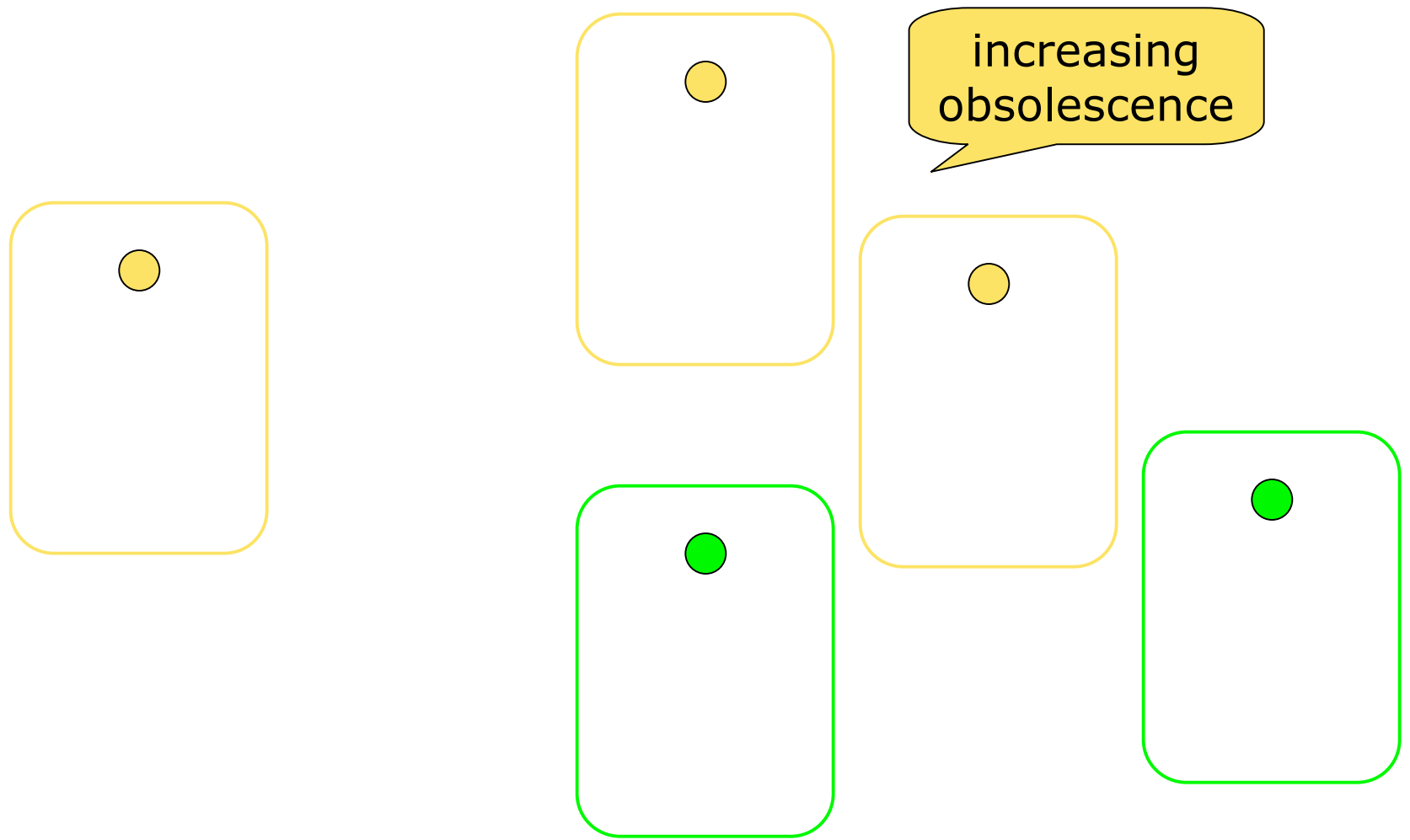


Mid-century perspective

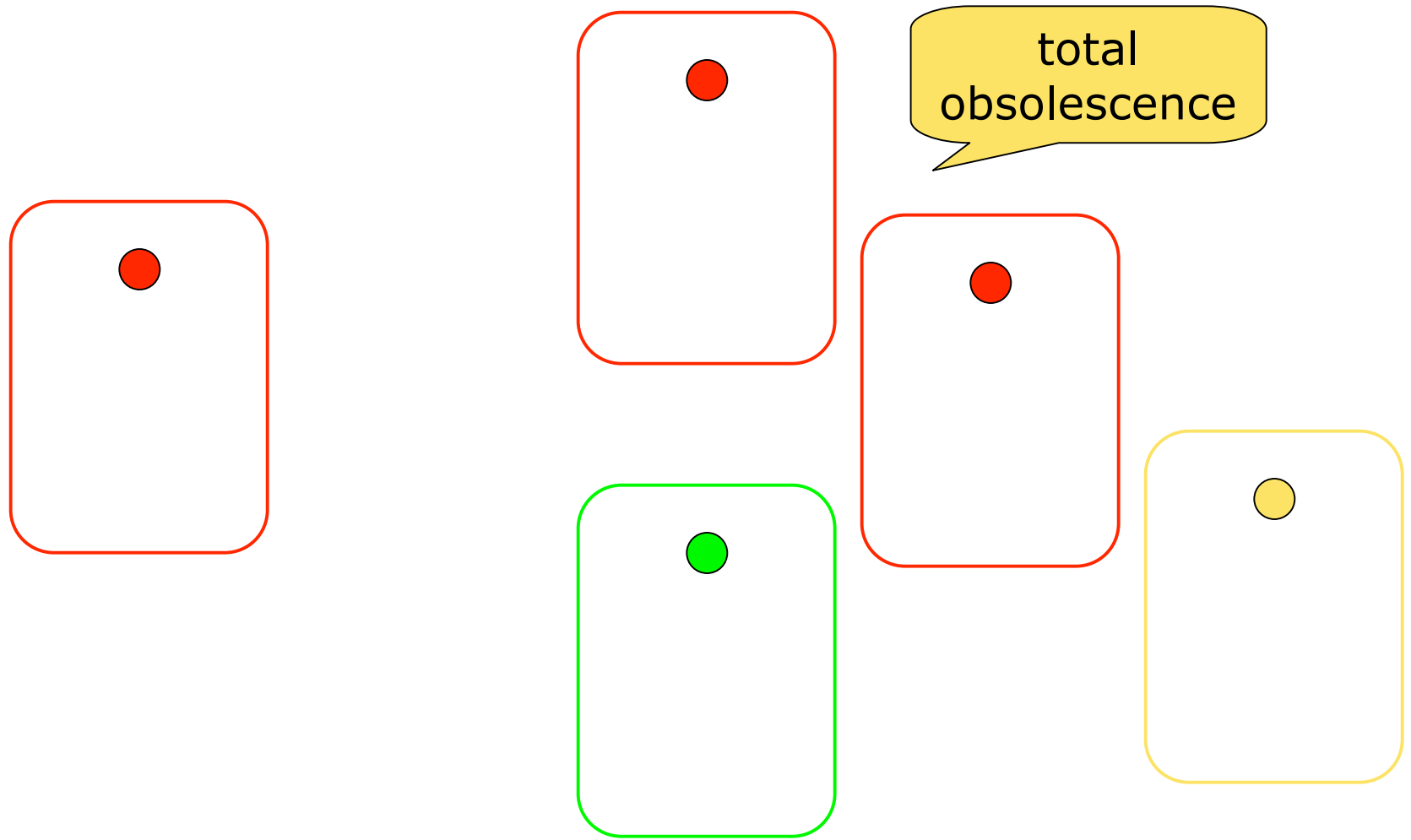
- Repeated migrations across storage media and storage systems
 - past and future
- Repeated migrations across archive systems
 - each possibly necessitating transformation and reorganization of archived content
- Repeated handoffs between institutions
 - each implementing different policies

Archival objects over time

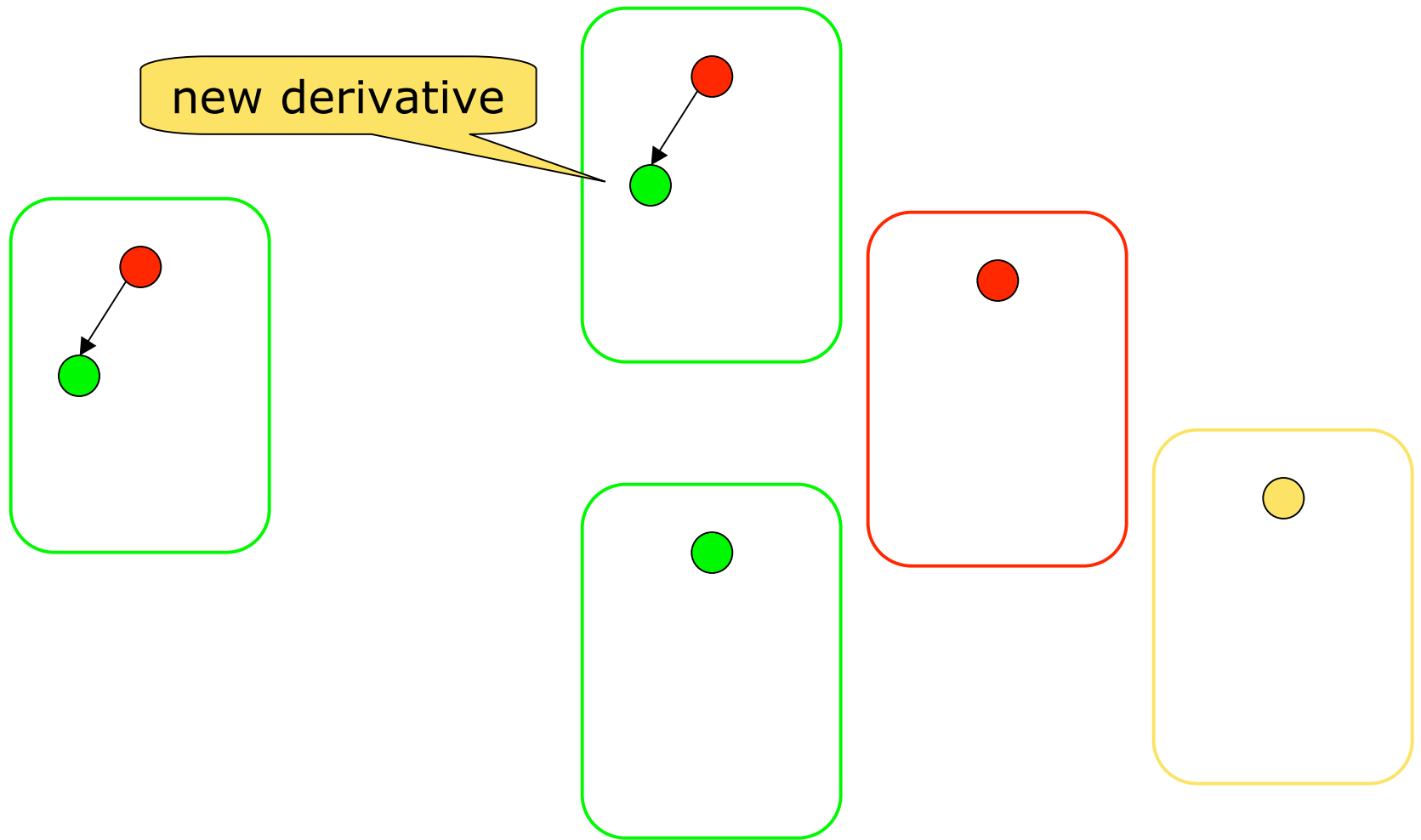




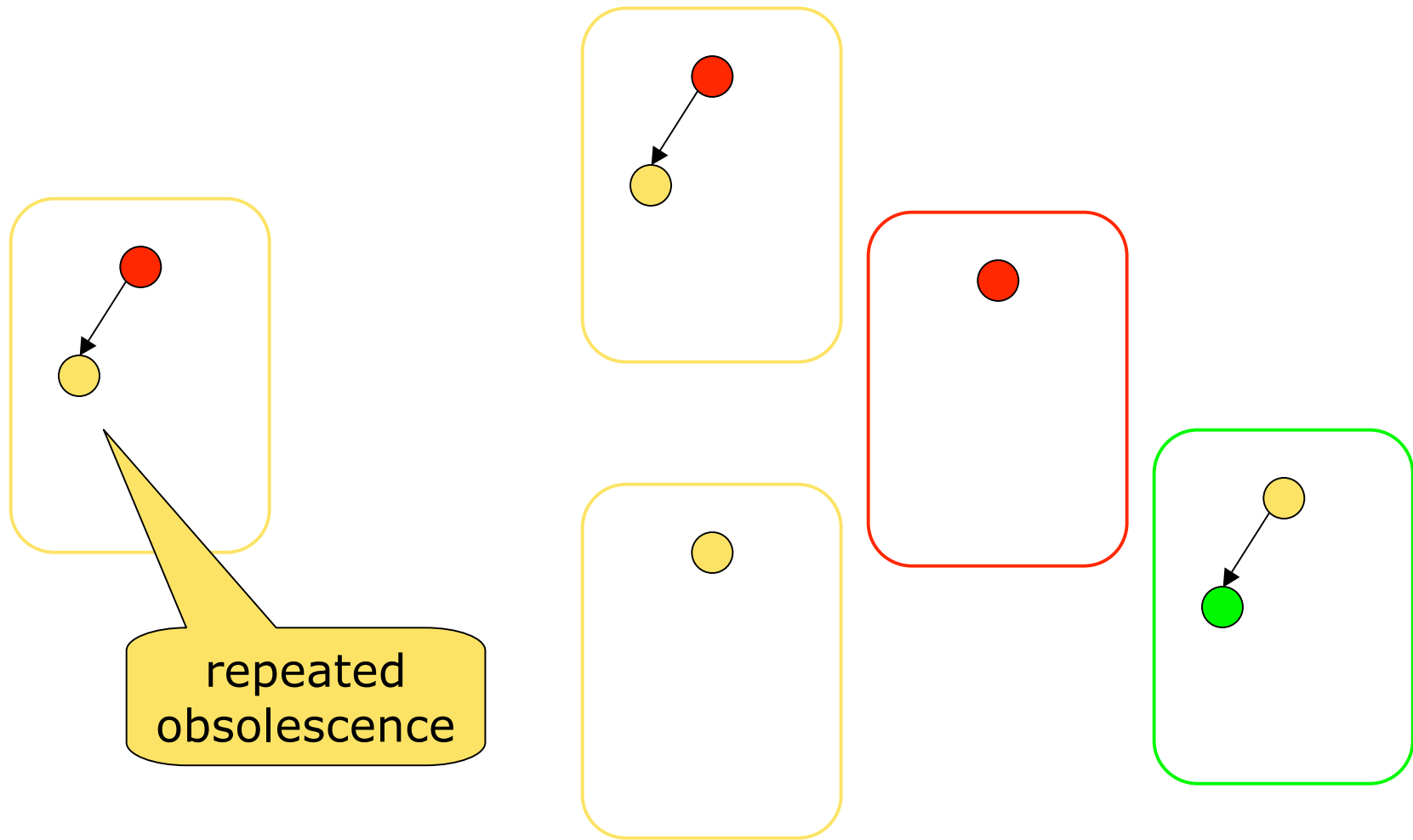
$T_0 + 10$ years



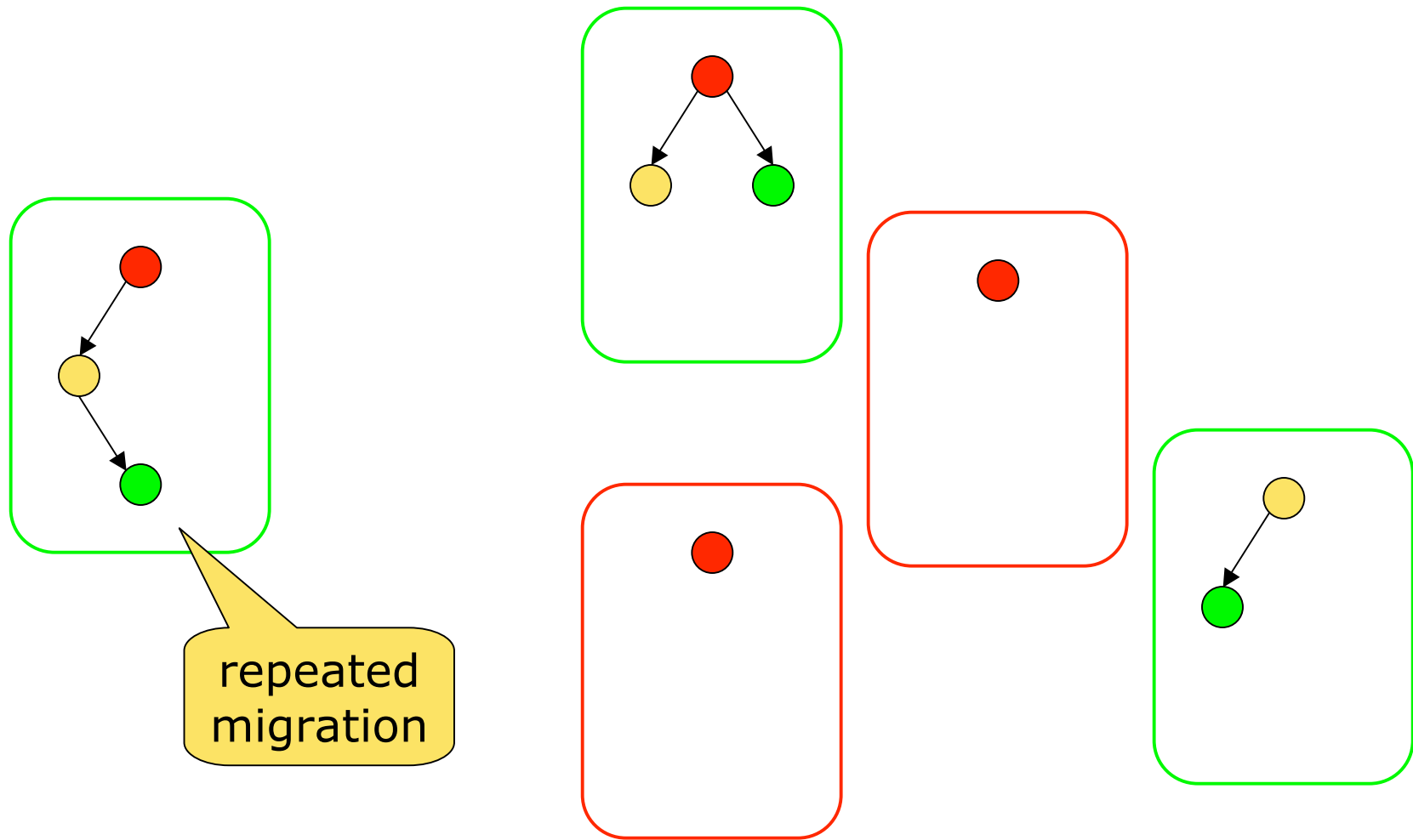
$T_0 + 20$ years



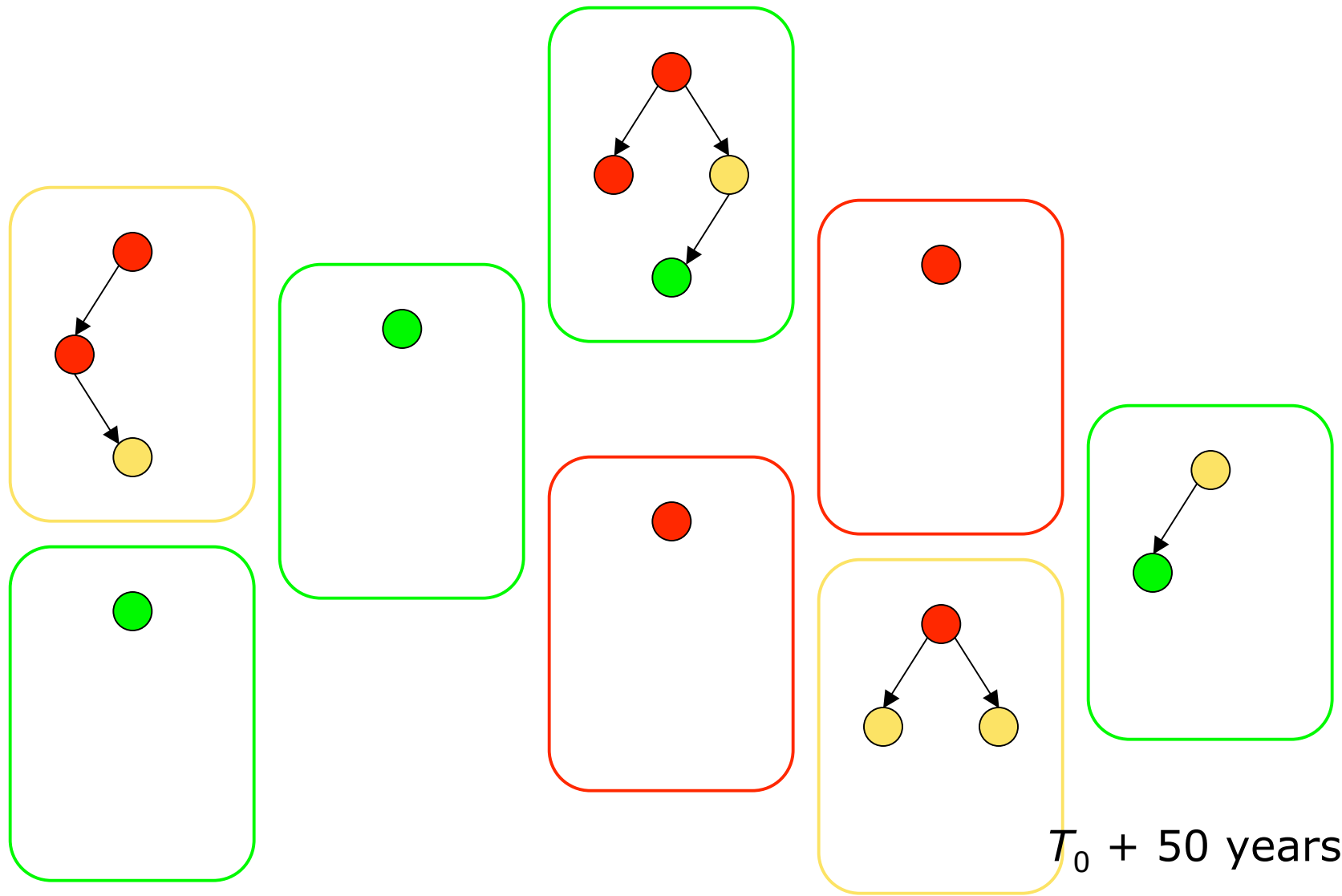
$T_0 + 20$ years



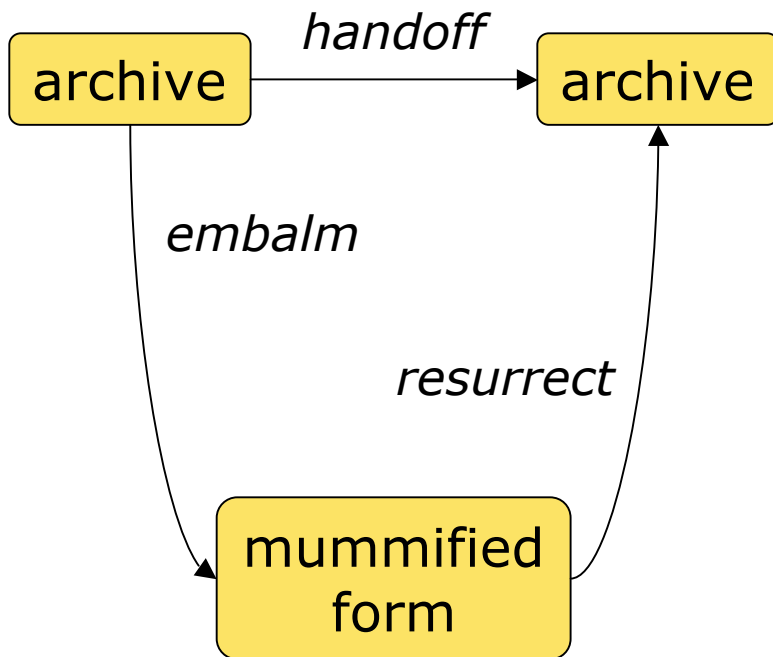
$T_0 + 30 \text{ years}$



$T_0 + 40$ years



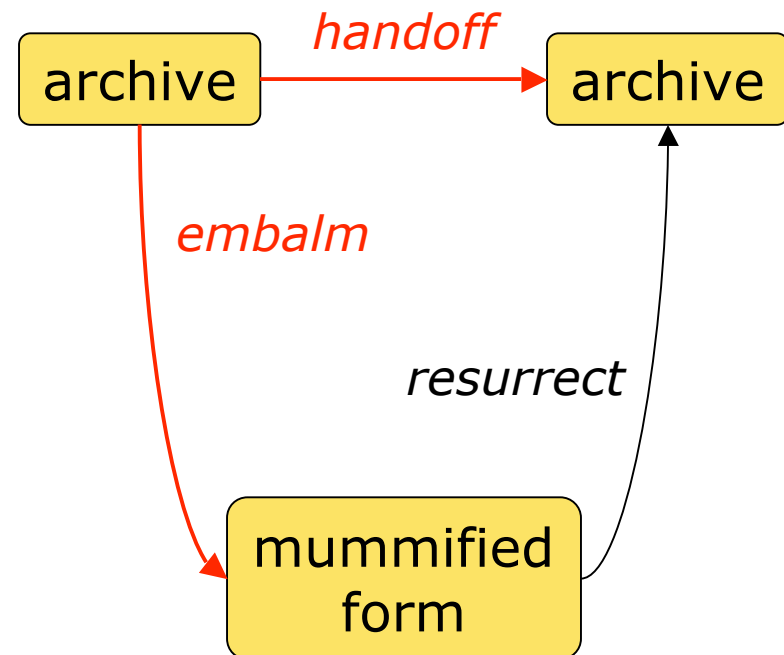
Mummification



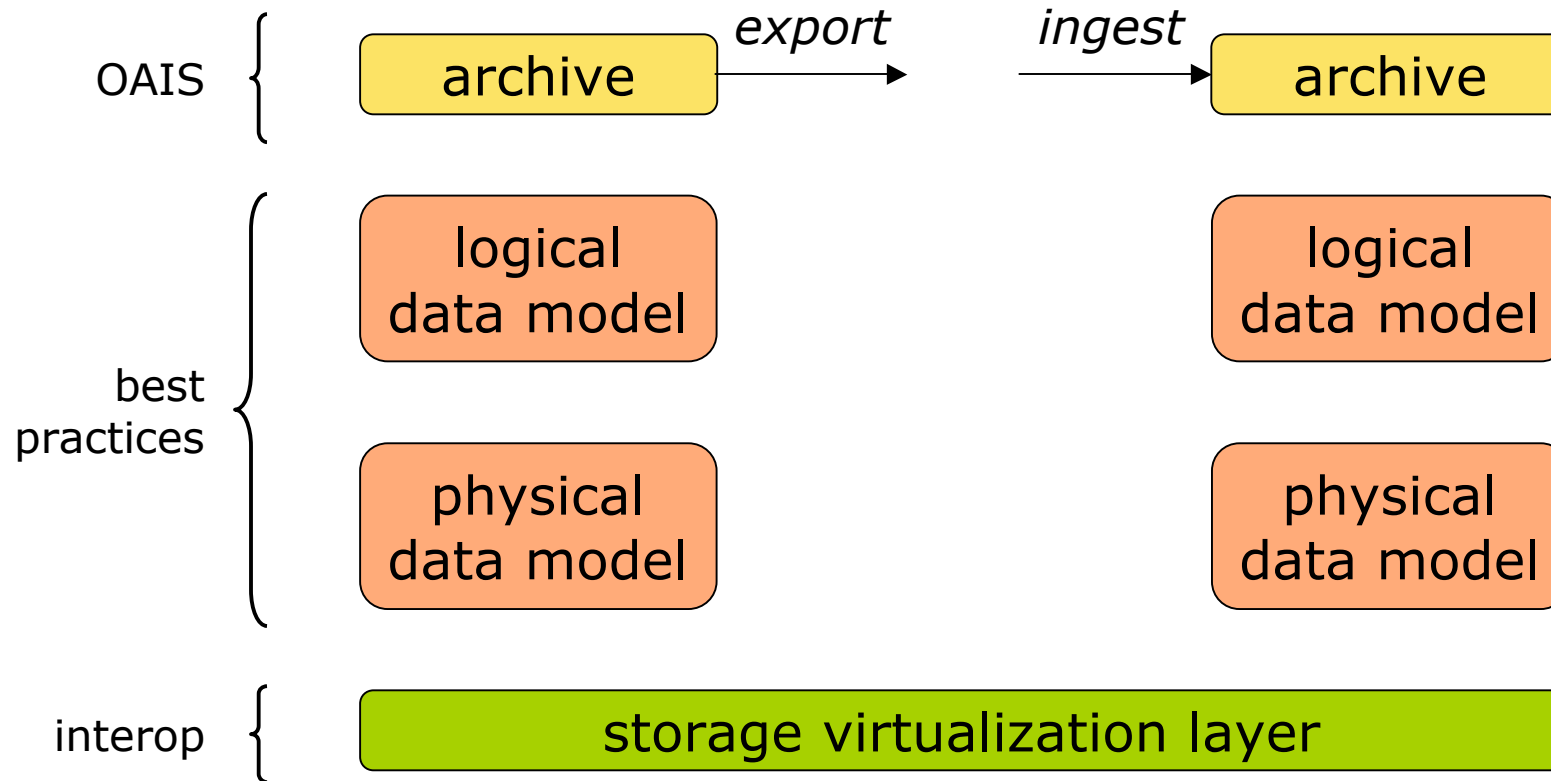
- Information value, resource availability vary over time
- Need Shirky's "option to do nothing"
- Need a representation of archival objects that is ultra-low-cost, resurrectable
- Equivalent of cardboard box in the basement

Questions

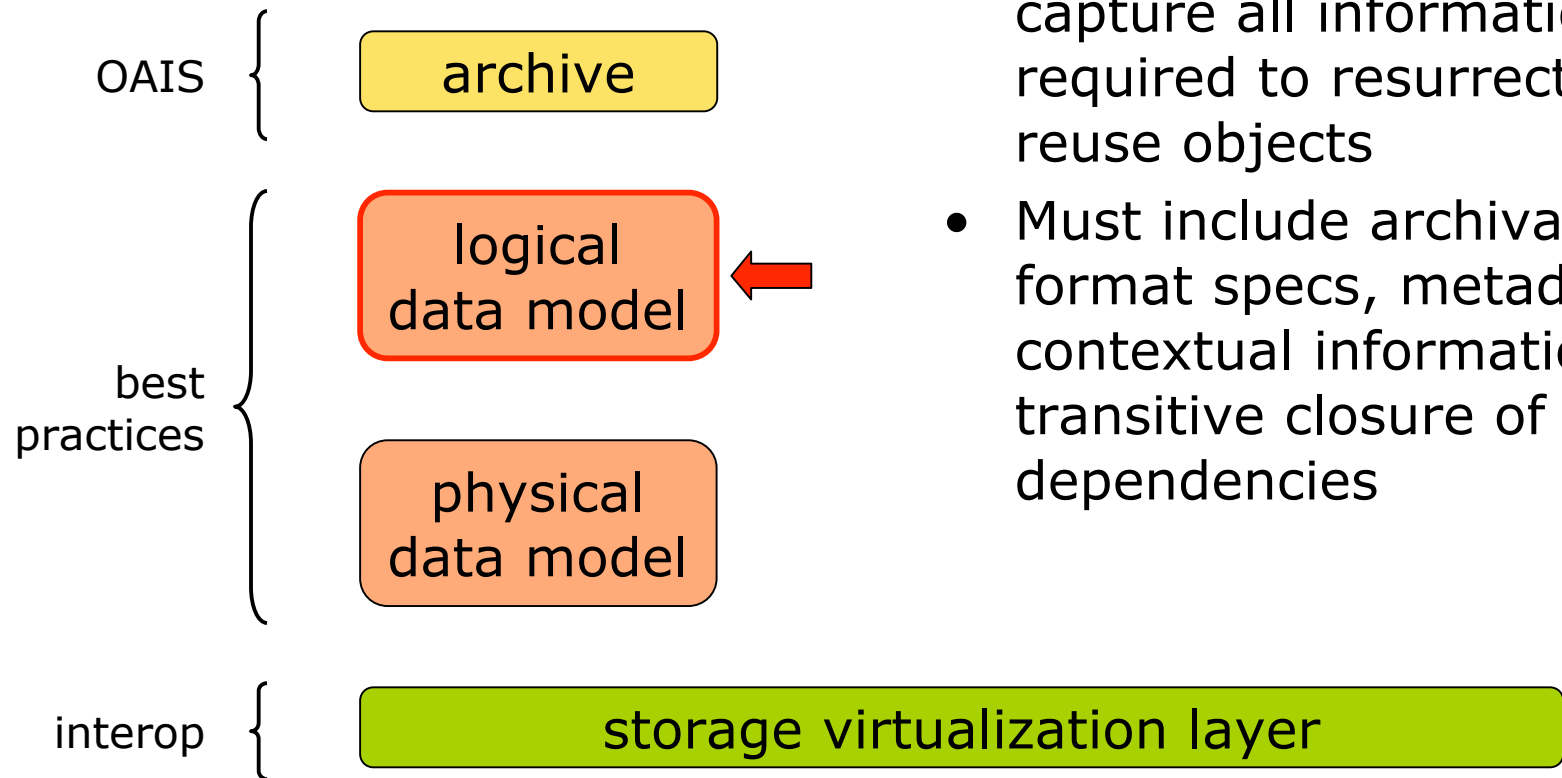
- How can we facilitate handoffs?
- How can we facilitate mummification?
 - particularly in cases of archive failure?
- How can we make resurrection possible?
 - will always be hard



Handoff architecture

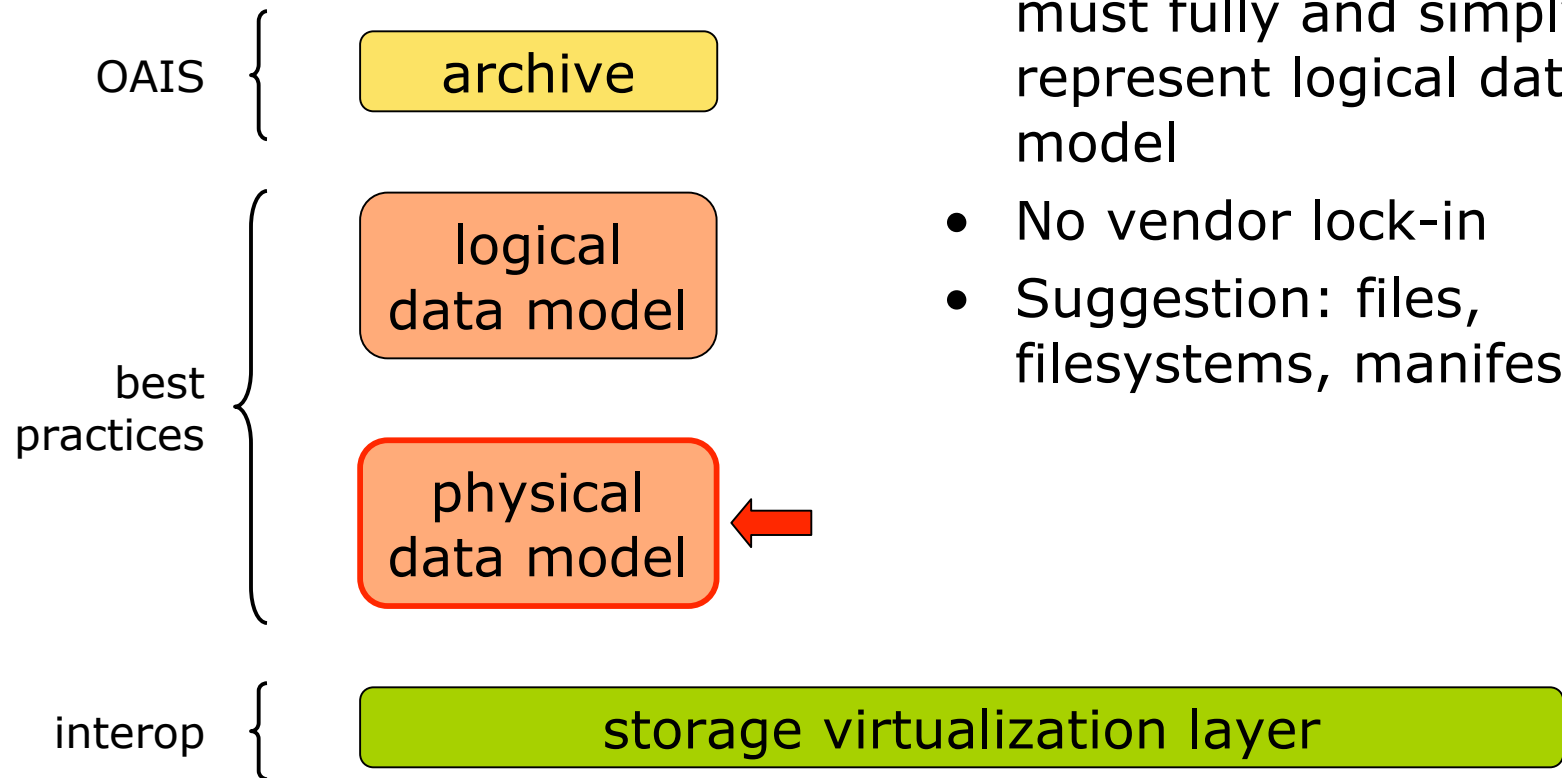


Handoff architecture



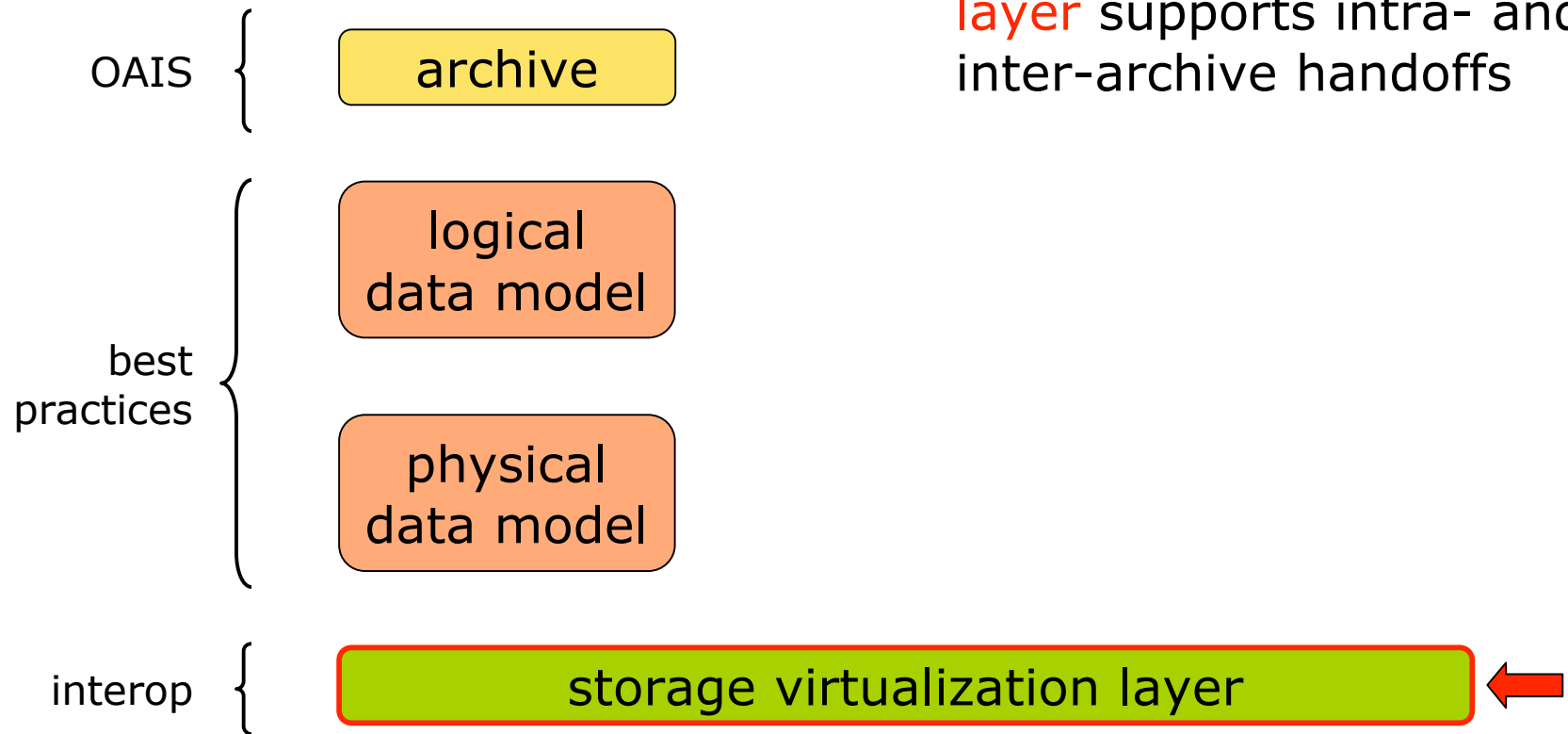
- **Logical data model** must capture all information required to resurrect, reuse objects
- Must include archival of format specs, metadata, contextual information, transitive closure of dependencies

Handoff architecture

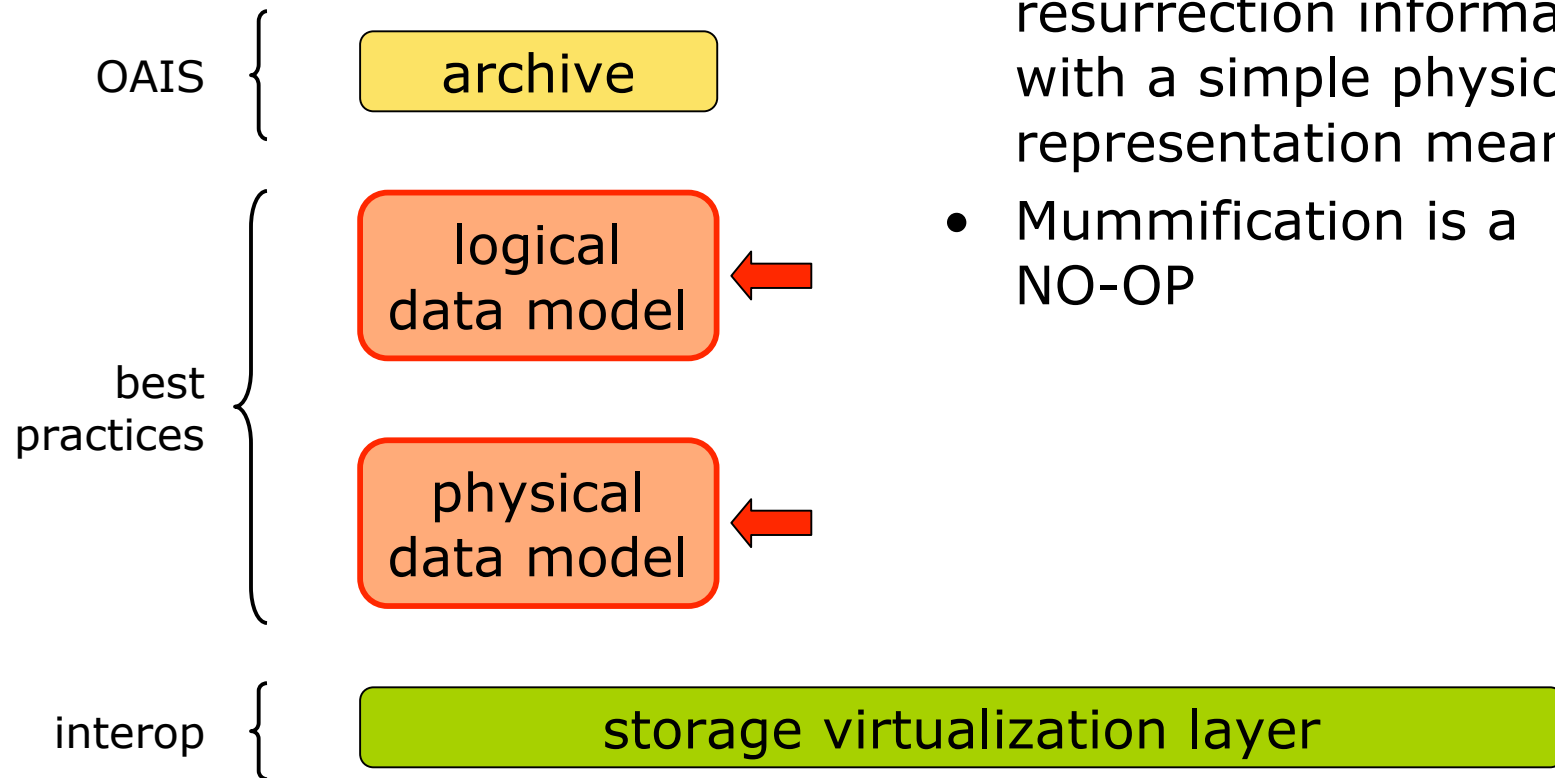


- **Physical data model** must fully and simply represent logical data model
- No vendor lock-in
- Suggestion: files, filesystems, manifests

Handoff architecture

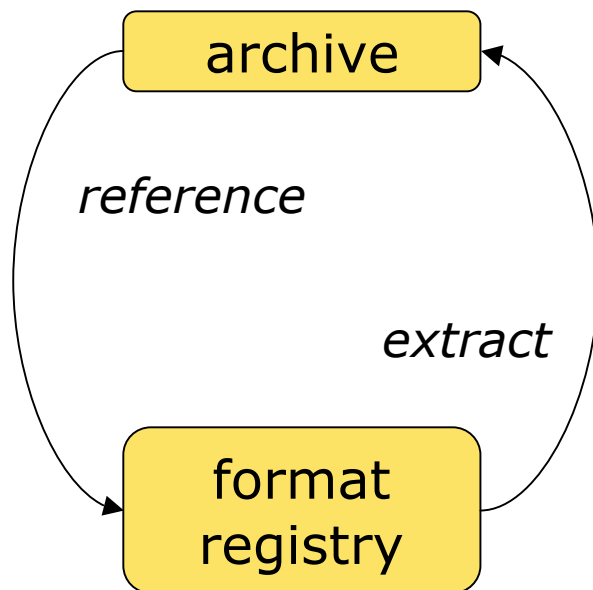


Handoff architecture



- Combination of complete resurrection information with a simple physical representation means...
- Mummification is a NO-OP

Format registries



- Archives reference registries
- Registries must capture, archive format specifications
- Must represent formats as archival objects
- Offer transitive extraction service

Discussion topics

- Characterization of ingest vs. handoff
- Filesystem-like representation for mummification?
- Standardization of logical, physical data models possible?