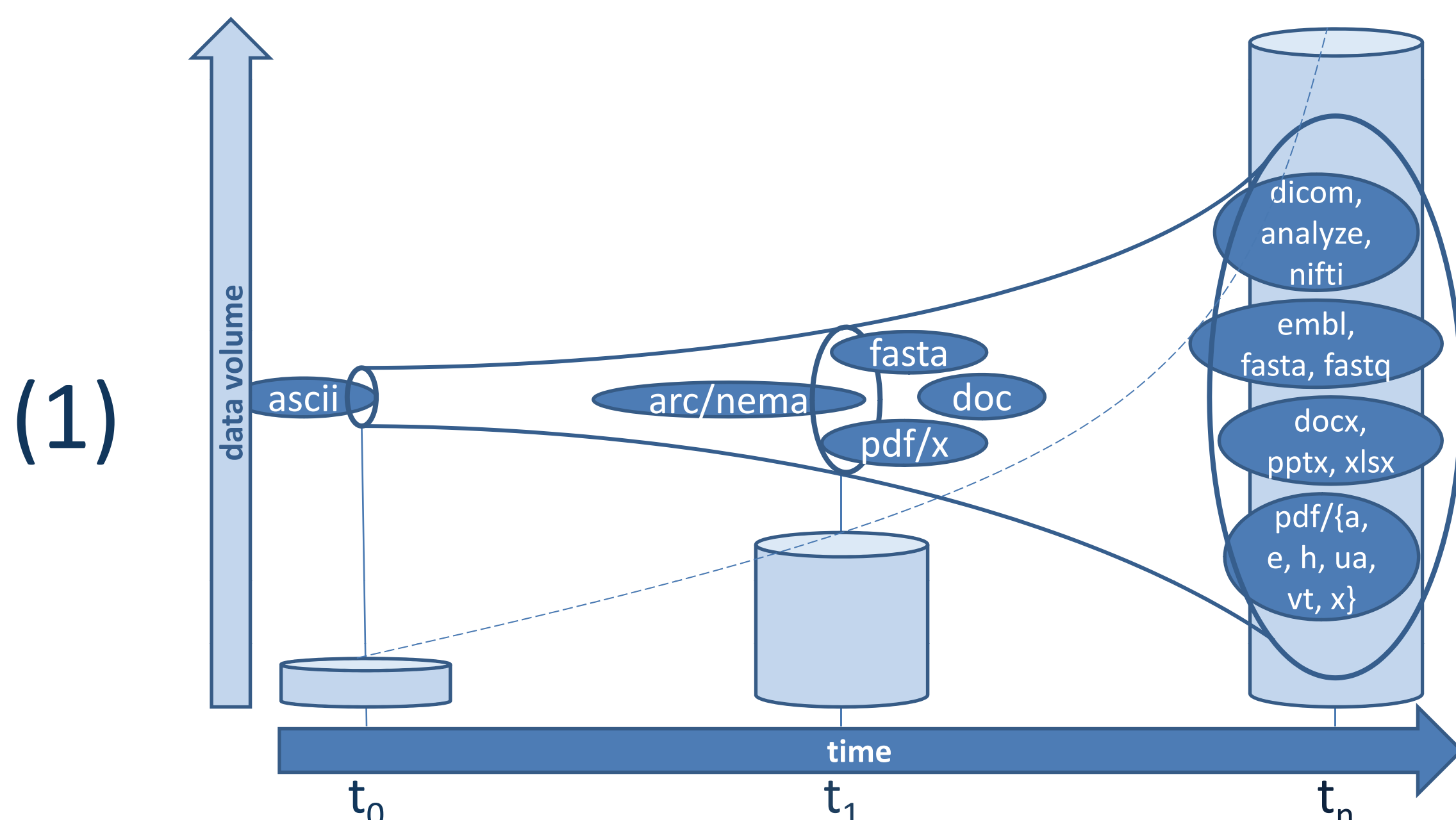


Background: Historical evolution of data

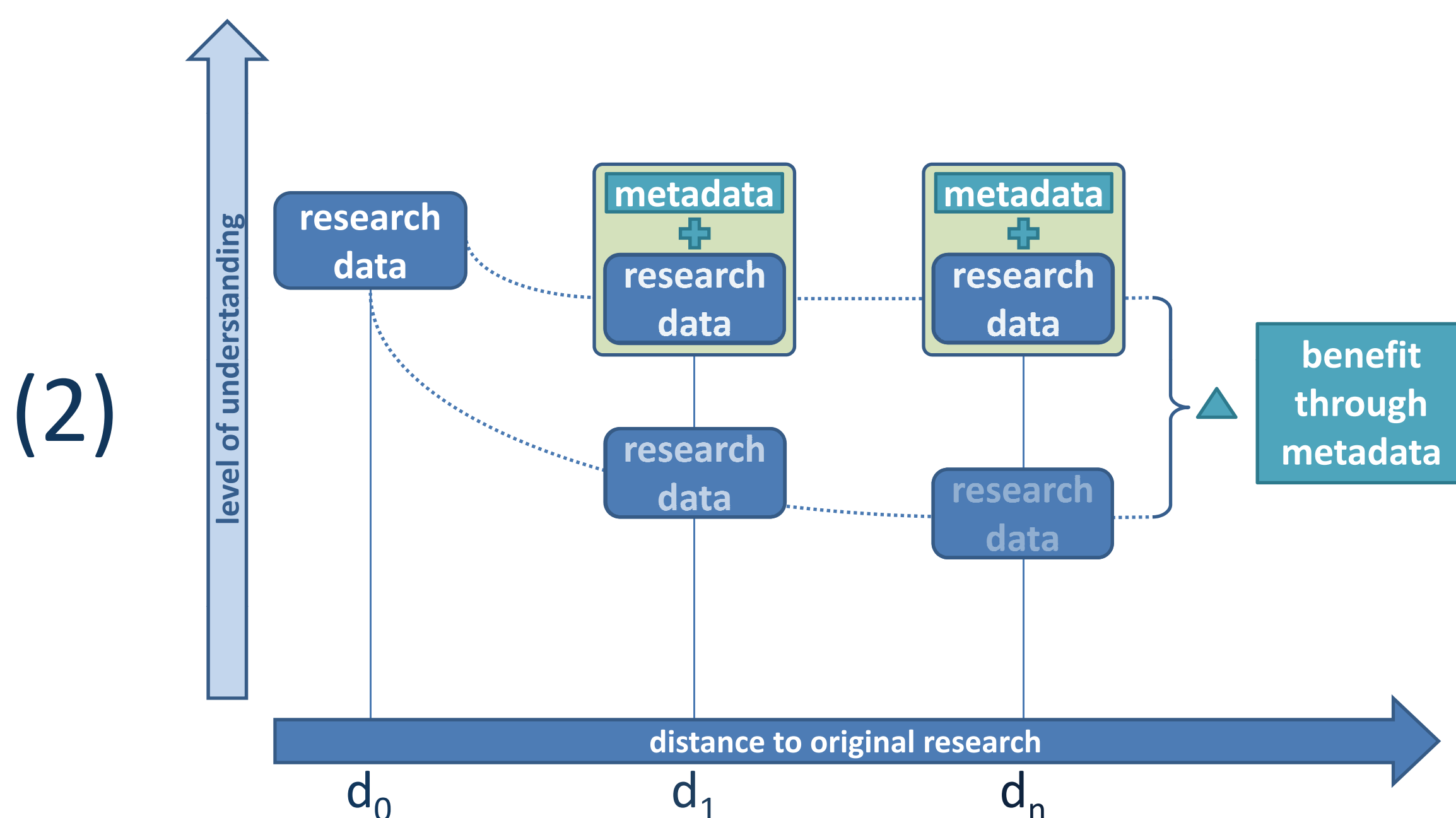


Since the invention of computers, data formats have evolved from punch cards to single small byte files – e.g. ASCII encoded – and then to complex data structures which combine multiple data types and formats. At the same time, the amount of research data produces has increased due to new possibilities and instruments.

In this context, biomedical research faces two major problems:

1. Continuing increase of data format diversity.
2. (Exponentially) increasing amounts of data.

Metadata: Approach for understanding

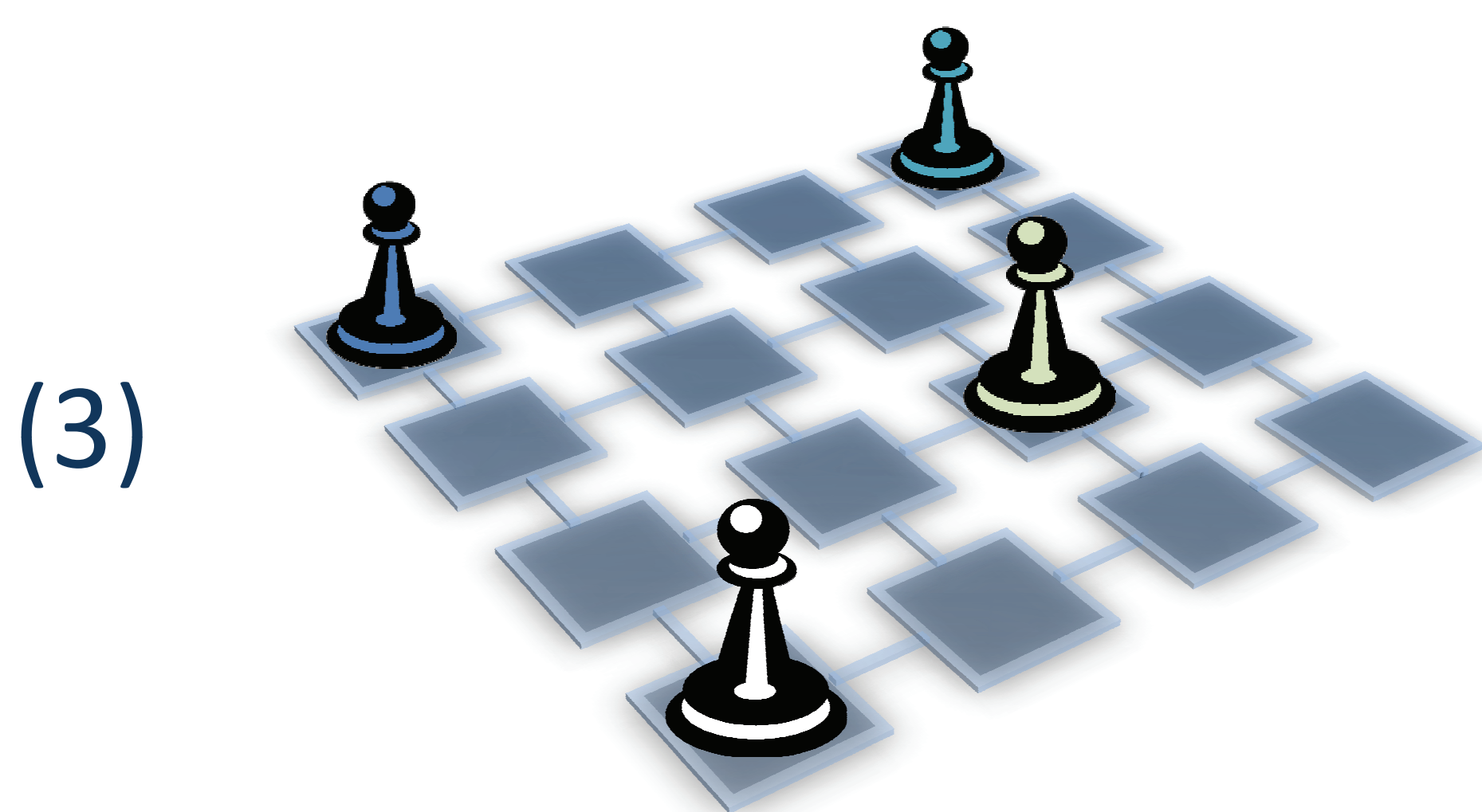


The evolution of data formats and the increase in data demand an efficient information retrieval not only within limited project time frames but also for long periods of time, e.g. 10 years and more.

The nature of metadata for the description of data contents provides an understanding of contents

- for researchers after a longer period of time in general, or
- for researchers which are locally or otherwise distant.

Grid computing: Architectural fundament



Grid computing provides a distributed computing infrastructure which interconnects vast numbers of heterogeneous computing, storage and knowledge resources. Connections, resource coordination, and infrastructure management are based on open software standards and an authentication and authorization infrastructure (AAI) based on common X.509 certificates.

Further advantages of grid infrastructures for digital preservation:

- Reliable data transfers
- Active development communities