

2017년 KSLA 학술세미나

# Research Data, Open Data

KISTI 정보기반실

최선희(sunny.choi@kisti.re.kr)

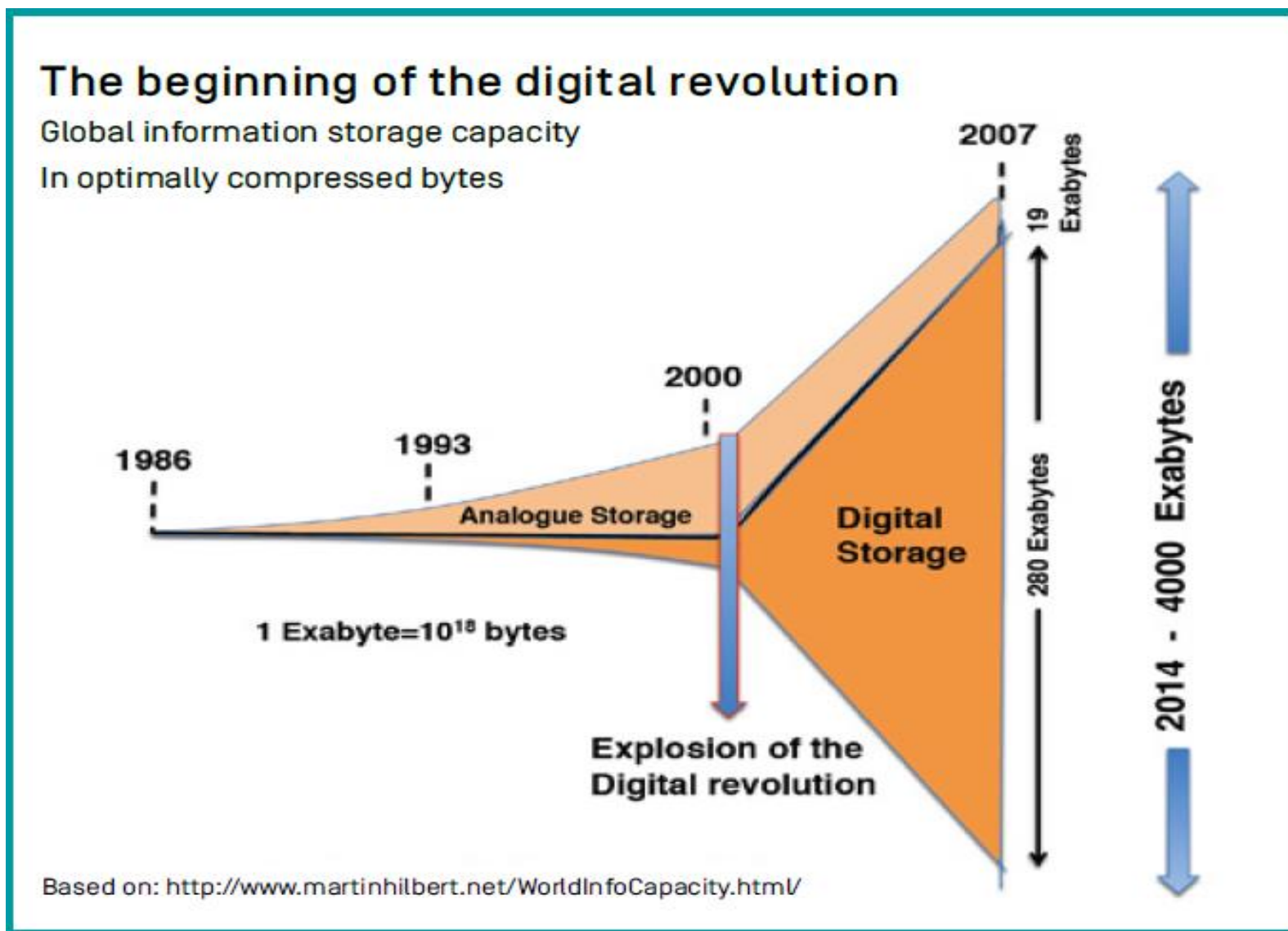
2017. 6. 29

An aerial photograph of a river delta, showing a complex network of branching waterways. A white grid is overlaid on the image, with lines intersecting at various points. The text is centered over the image.

open data as a fundamental pre-  
requisite in **maintaining the rigour** of  
scientific inquiry and **maximising**  
**public benefit** from the data revolution.

Science International (2015): Open Data in a Big Data World.

# The beginning of the digital revolution



# 과학의 핵심 가치

Openness and transparency

# 연구데이터의 핵심 가치

*Data used as evidence for a scientific claim*

1. 과학적 검증은 연구자가 자신의 업적을 증빙하기 위해 필수적
2. 연구비지원기관도 연구과제의 실효성을 검증하기 위해 필수적  
--- 연구자는 동전의 양면성과 같은 태도를 가진다  
자신의 데이터는 공개하지 않으면서 타연구자의 데이터는  
오픈되길 원한다
3. 도서관은 연구자와 연구기관, 연구비지원기관사이에서 각 이해당사자를 지원서비스하고, 대중을 위한 서비스를 제공한다



# 연구데이터 주요 범주

- The definition of open data
- Non-Replicability
- Open data and “self correction”
- Openness: the default for publicly funded research
- Boundaries of openness
- Ethical issues
- Open Data Platforms
  - *A national data platform*
  - *A disciplinary platform: ELIXIR*  
*: an integrated data support system for the life sciences*
- Open Science and Open Data, Public Knowledge or Private Knowledge?
- **Principles of Open Data**
  - Responsibilities, Boundaries of openness, Enabling practices
- **The Practice of Open Data**

# 오픈데이터 원칙 : 책임

1. 공적자금을 받은 과학자는 연관된 데이터를 고유한 부분으로 포함하는 새로운 지식의 창출과 소통을 통해 공공에게 기여할 책임이 있다.
  - 데이터는 재사용과 re-purposed한 방식으로 생산되자마자 오픈되어야 함
2. 출판된 과학적 검증(claims)을 위해 제공되는 데이터는 매우 지능적이고 공개된 양식으로 현재적이고 공적으로 유용해야 한다.
  - 데이터와 검증간에 링크 가능, 복제실험 관찰이 가능해야
3. 연구기관과 대학은 오픈데이터를 지원하기 위한 환경을 구성할 책임을 가진다.
  - 데이터관리, 보존, 분석과 적절한 기술지원, 도서관과 데이터 관리서비스를 포함한다.
  - 오픈데이터 프로세스에 포함된 경력개발을 위한 인센티브와 평가기준을 개발하여야 한다. 이러한 기준은 기관내를 넘어 국가적 글로벌적으로 동의를 얻어야 하며, 연구자의 바람직한 동기부여를 촉진하여야 한다.

# 오픈데이터 원칙 : 책임

4. 출판사는 리뷰과정동안 리뷰어에게 데이터를 이용가능하게 하고 데이터를 사용하는 출판물과 데이터를 공동으로 오픈엑세스 할 수 있도록, 데이터에 대한 인용과 완전한 참고문헌을 제공할 할 책임을 가진다.
  - 메타데이터의 제공과 텍스트/데이터 마이닝을 위한 접근을 통해 지속적인 분석이 가능하게할 책임이 있다.
5. 연구비 지원기관은 연구프로젝트에서 오픈데이터 프로세스에 소요비용을 연구수행의 본질적인 비용으로 간주하여야 한다. 데이터인프라스트럭처와 리파지토리의 장기적인 지속성을 위한 적절한 자원과 정책을 제공해야 한다.
  - 연구영향력의 평가 특히 인용지수는 데이터 생성자의 공헌
6. 전문협의회, 학술학회, 아카데미는 오픈데이터를 위한 정책과 가이드라인을 개발해야 하고, 그 회원들의 급속한 확산과 실재를 반영하는 방법으로 기회를 촉진해야 한다.
7. 도서관, 아카이브, 리파지토리는 데이터를 활용하기 원하는 이용자에게 데이터를 활용가능하게 하고 장기관 접근가능하게 하는 서비스와 기술적 표준을 개발하고 제공할 책임이 있다.



# 오픈데이터 원칙 : 공개범위

1. 공적자금이 투입된 고학을 위한 것이 디폴트 영역이 되어야 한다.
2. 예외적인 것은 사생활, 안전, 보안의 이슈와 공적인 이익이 있는 상업적인 활용에만 제한되어야 한다.
3. 예외는 case-by-case로 적용되어야 하며, blanket basis로 정당화되어서는 안된다.

# 오픈데이터 원칙 : 실체화

1. 인용과 출처 : 학술출판에서 연구자가 다른 연구자의 데이터를 활용할때, 반드시 그 원작자에 대한 인용표시와 영구디지털식별자에 대한 출처표시가 이루어져야 한다.
2. 상호운용성 : 연구데이터와 메타데이터는 접근가능하고 재사용 가능하고 최대한상호운용가능해야 한다.
3. 비제한적인 재사용 : 기공개된 데이터라면 적절한 웨이버나 비제한적인 라이선스를 통해 재사용 가능해야 한다.
4. 연결성 : 데이터는 콘텐츠를 구성하는 다른 데이터 및 그 의미적 가치를 극대화하는 맥락과도 연결되어야 한다.