

Preserving Transactional Data Briefing Day

Sara Day Thomson | @sdaythomson

17 March 2016 London



Digital**Preservation**Coalition



UK Data Service



Our digital memory accessible tomorrow...

Outline

Introduction

Definition

Approaches

Challenges

Legal Review

Case Studies

Conclusions

Tweet

Share your thoughts
on Preserving
Transactional Data
#ptd_dpc

Me @sdaythomson

DPC @dpc_chat

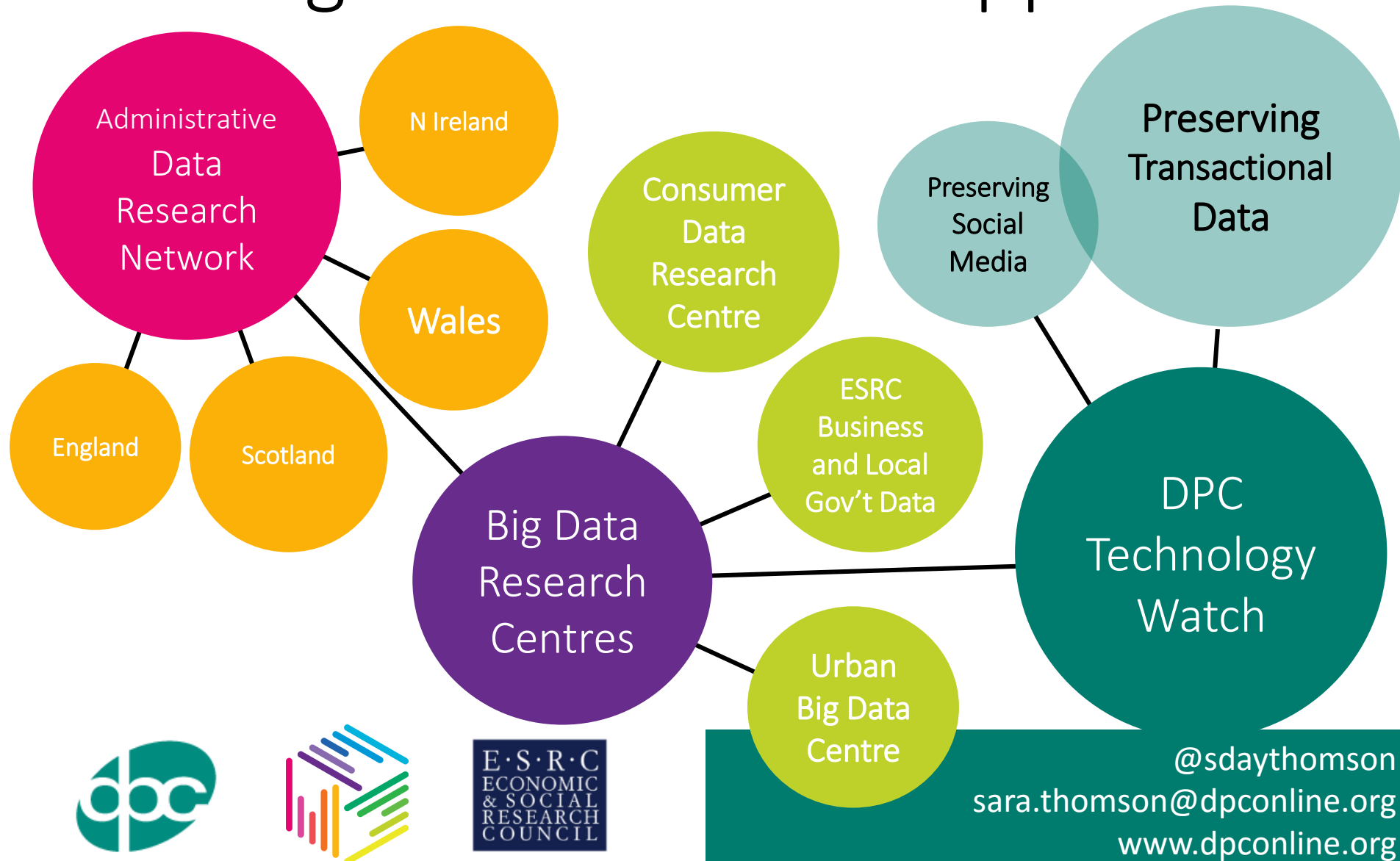
UKDS @UKDSBigData



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

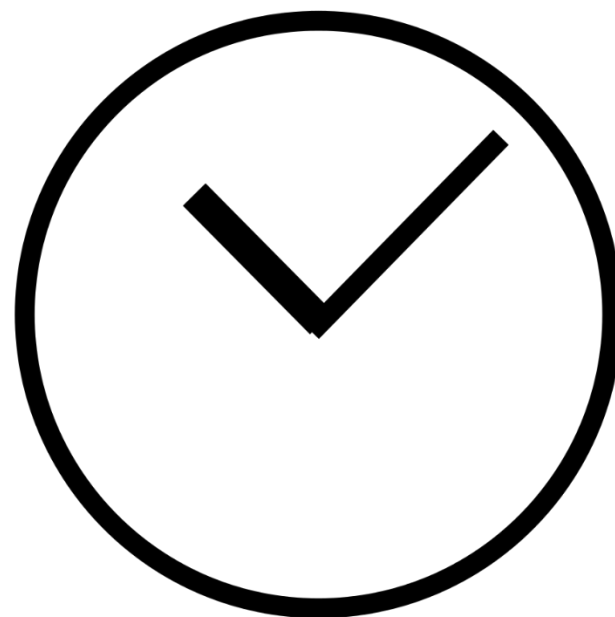
UK Data Service

Big Data Network Support

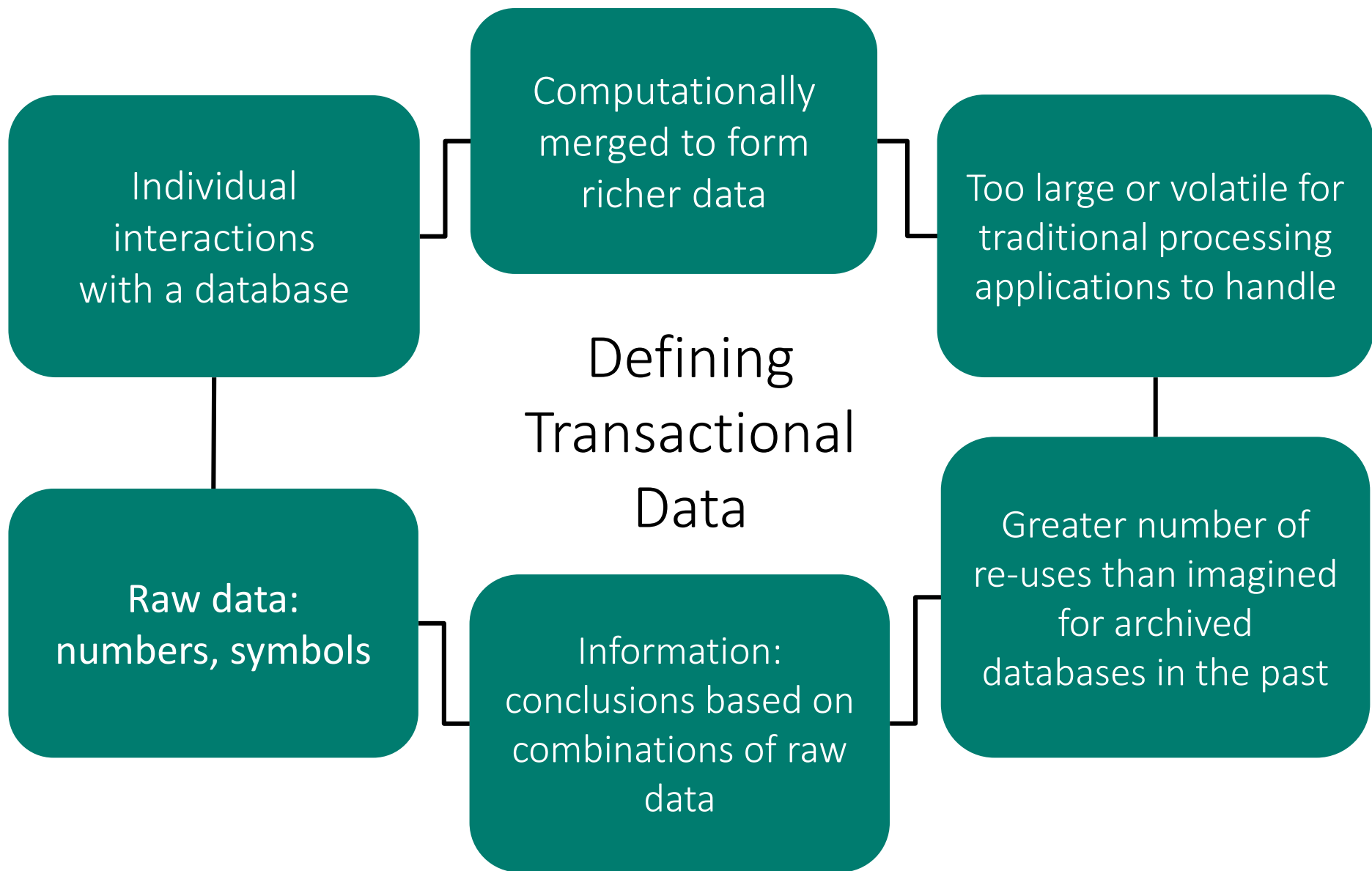


Long-term Preservation concerns?

- Reproducibility in research
- Compatible, comparable data
- Availability of historical data
- Compliance and records management



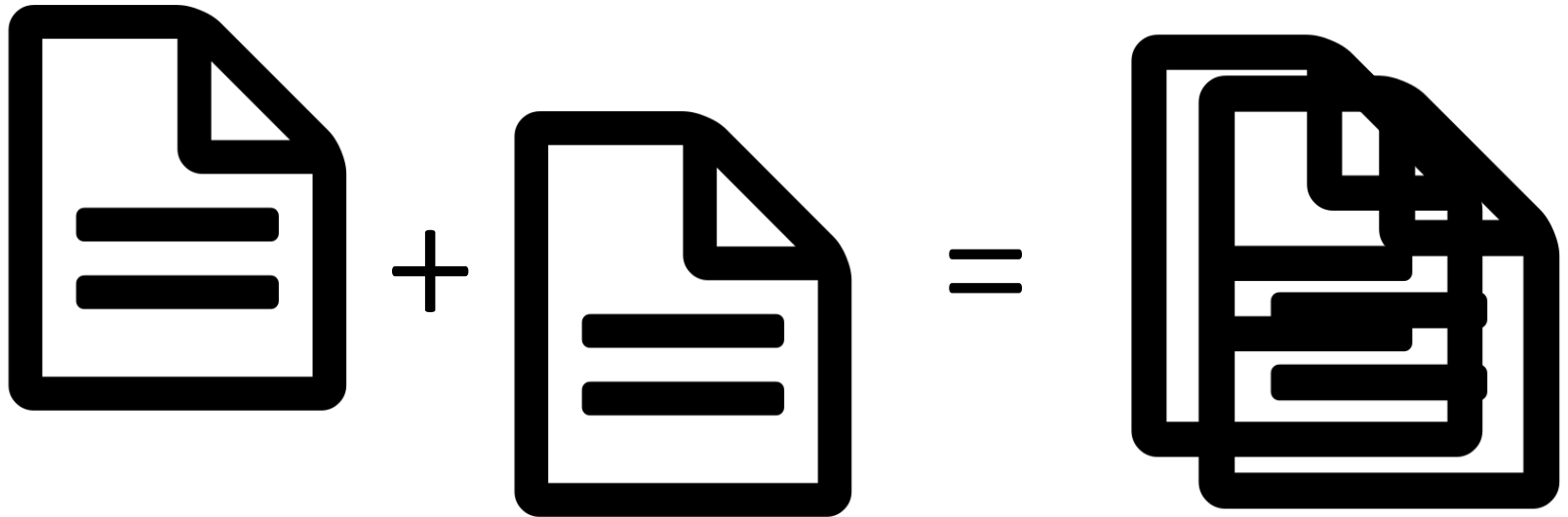
@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Computationally
merged to form
richer data

The value of 'Big' Data
is the ability
to merge, or link



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Greater number of re-uses than imagined for archived databases in the past

Regulatory Compliance

Education

Governance

Consumer Analysis

Public Services

Transport

Environmental Science

Economics

Geo-spatial Analysis

Computer Science

Social Science

Spread of Disease

Cyber Crime

Network Analysis

Urban Planning

Business Analysis

Energy Usage

Human Behaviour

Healthcare

Heritage Collections

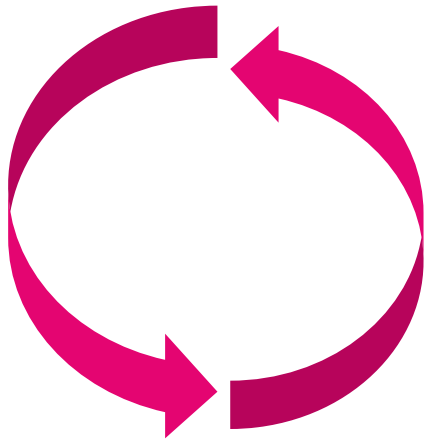


@sdaythomson

sara.thomson@dpconline.org

www.dpconline.org

Approaches



Re-use of Data



Database Preservation



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Re-use of Data

- Acquire or license from data source
- Partnership with data source (organisation who collects data)
- Open Data



Database Preservation

Encapsulation

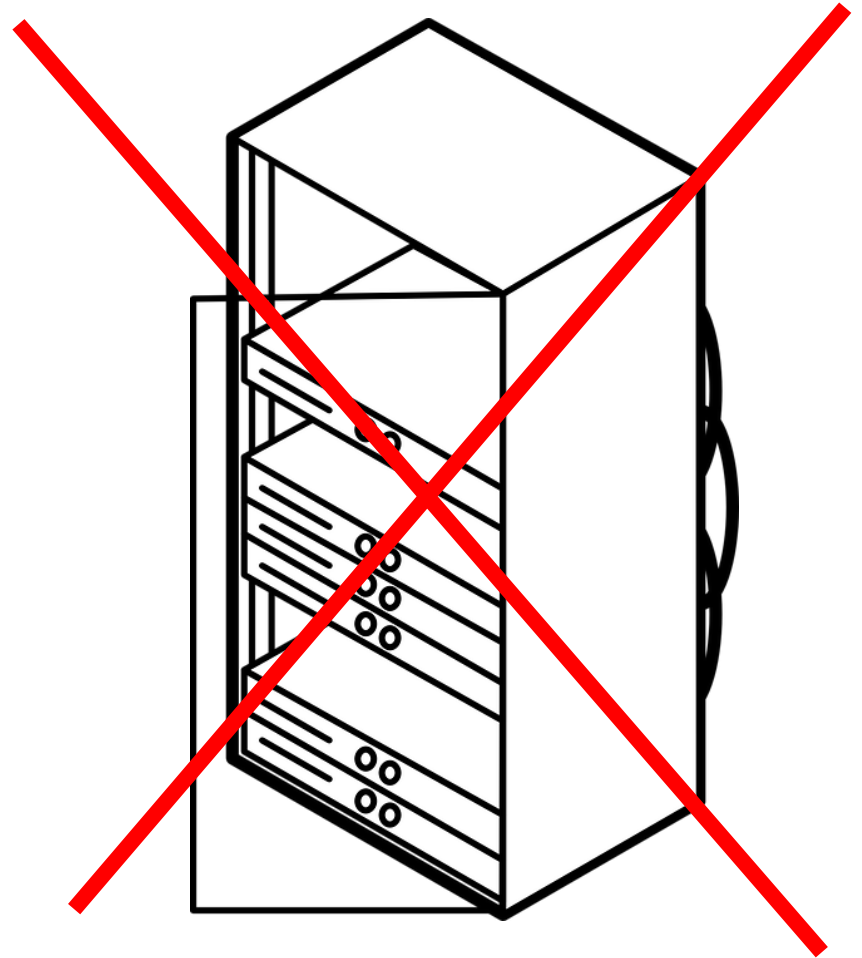
Migration /
Normalisation

Emulation



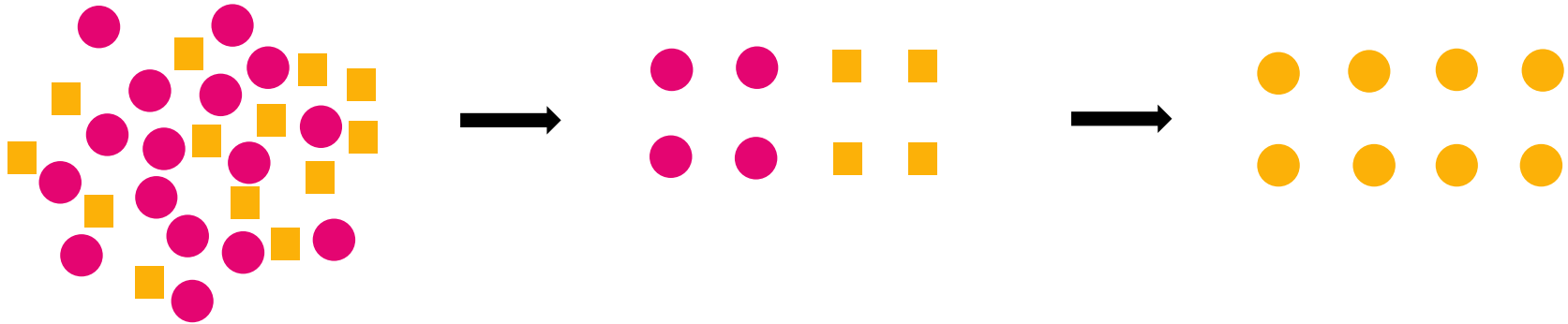
@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

What are the challenges?



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

It's all about re-use...



Selection

Compatibility

Metadata & Documentation

How can these data
be used now, and in
the future?

Legal Review

EU Directive 96/9/EC (Database Directive) (1996)

UK Human Rights Act (1998)

UK Data Protection Act (1998)

European General Data Protection Regulation (GDPR) (2018)



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Legal Review

EU Directive 96/9/EC 'Database Directive'

- Exclusive rights holder
- Copyright protection
- Sui generis

UK Data Protection Act (1998)

Protects personal data

Data collected for one purpose cannot necessarily be re-used for another

There is no
anonymization
in big data



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Legal Review

'...in the vast majority of cases...the complexity of the law, amplified by a plethora of guidance, leaves those who may wish to share data in a fog of confusion'.

Thomas and Walport,
'Data Sharing Review Report',
11 July 2008



Laurie and Stevens,
'The Administrative Data Research Centre
Scotland: A scoping report on the legal &
ethical issues arising from access & linkage
of administrative data',
Research Paper Series No 2014/35



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Enabling Future Re-use: Validation and Reproducibility

- Persistent Identifiers (DOIs)
- Standardized metadata & documentation
- Conversion to preservation-friendly format
- Preserve entire environment
- Secure preservation environment



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Case Studies

UK Data Service: Energy Demand
Research Project

GESIS: Conceptualising a Spatial Data
Infrastructure

Consumer Data Research Centre: Data Services

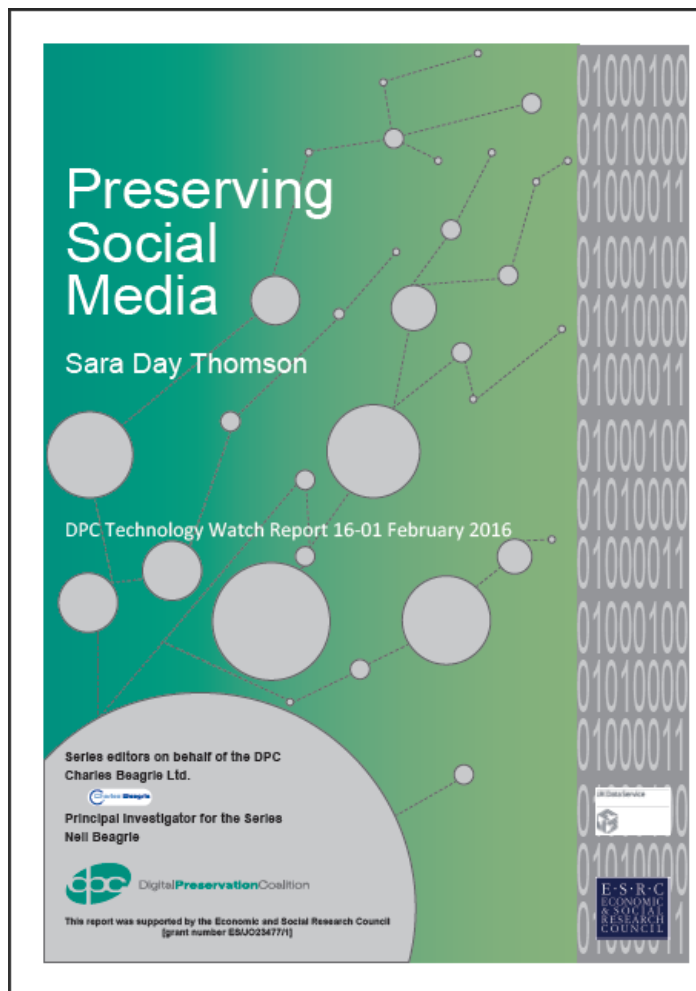


@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Conclusions

- New approaches to preservation
- Selection, compatibility, metadata, and documentation
- Preserving more than data
- Planning for broader uses





Preserving Social Media

DPC Technology Watch Report

dx.doi.org/10.7207/twr16-01



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org

Thanks!

Please fill-in your feedback form.



@sdaythomson
sara.thomson@dpconline.org
www.dpconline.org