The concept of 'research route' as a methodological basis for social analysis of web archives, with special attention to the transformations of the websites of Portuguese newspapers as a case study

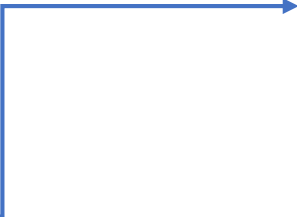Diogo Silva Cunha, University of Lisbon
cunhadiogo15@gmail.com

- "Investiga XXI"

- "Transformations of Portuguese Newspaper
  Websites: *Correio da Manhã*, *Diário de Notícias*,
  Expresso and *Público* (1996-2016)"

- Web archives:

  ➢ Reborn-digital materials (reproduction of
    offline webpages)

  ➢ Supplier of empirical and historical
    materials

- Arquivo.pt

- Transformations of the websites entangled research problems:

  ➢ How to access the materials for study?

  ➢ How to choose the materials for study?

- Intuitive answer: you type, you have results.

- First approach: 3,166 counted versions for the 4 newspaper websites

- **"Research route"**

- Cybernetics (Norbert Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine ,* 1948):
  - ➢ Greek for steersman
  - ➢ Control and communication theory in animal and machine
- Cyberspace as "navigable space" (Lev Manovich, *The Language of New Media*, 2001)
  - ➢ The navigator/ flâneur
  - ➢ The explorer
- Beyond automatic navigation

- Research Route

  ➢ Intermediate position of cartography to find the best way among possible ways of defining the analysis corpus

  ➢ A course on the internet through related websites

  ➢ Connections between the websites following criteria (what's best?)

  ➢ Analogy with a navigable route (instruments, observation, calculation, intuition)

- Research Route

  ➢ In short: a methodological concept for

    isolating research material in a web archive

    following certain criteria.

- The transformations of the websites of Portuguese newspapers as a case study → Only homepages

    ➢ Finding newspaper websites addresses in Arquivo.pt (conditioning: Arquivo.pt allows search by URL and text)

    ➢ Three search modes:

        ➢ a non-exhaustive internal list with collected data

        ➢ Search by URL (current URL)

        ➢ Search by text ("Diário de Notícias"/"DN"…)

- The transformations of the websites of Portuguese newspapers as a case study

  ➢ Exhaustive research of available material

  ➢ Creating a database in the form of tables to relate data on the addresses for documenting and ordering the available material

  ➢ Designing a more oriented route functioning as a map of preserved data, actually available, related and with representative quality

- The transformations of the websites of Portuguese

  newspapers as a case study

  - ➤ Categories:

    - ➤ Address

    - ➤ Number of versions

    - ➤ Date of the first version

    - ➤ Date of the last version

    - ➤ Comparatively high incidence (years

      with more than X versions)

    - ➤ Observations (notes after analysis of

      each webpage)

Most common note:

redirections led to the

discovery of previously

unknown addresses

- The transformations of the websites of Portuguese newspapers as a case study

  ➢ Table example:

Address	Number of versions	Date of the first version	Date of the last version	Comparatively high incidence	Observations
publico.clix.pt	1.145	February 22, 2005	May 25, 2016	2005, 2006	As of June 6, 2007, this address redirects to versions of the address ww2.publico.clix.pt. As of 2010, this address redirects to versions of the address publico.pt.

- The transformations of the websites of Portuguese newspapers as a case study
  - ➤ Remember: the first approach counted versions for the 4 newspaper websites were 3,166
  - ➤ The final count of versions after the establishment of the research route is 21.148!

  The benefits of the concept are evident

- Research route

  ➢ It was used as a methodological basis for research

  ➢ It may also be used for organizing materials with specific purposes, e.g. collections of thematic news or certain institutions (the explorer's research route would turn into a flanêur's leisure route)

THANK YOU!

Diogo Silva Cunha, University of Lisbon
cunhadiogo15@gmail.com