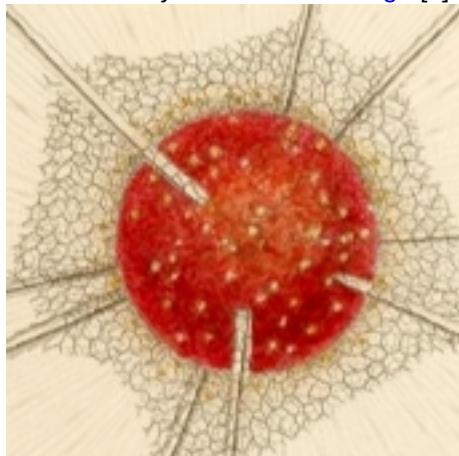


Big Data i humanitats digitals: de la computació social als reptes de la cultura connectada

Published by [Maria Farràs Drago](#) [1] on 23/10/2013 - 12:44 | Last modification: 05/02/2026 - 14:01



Un post de [Sandra Álvaro](#) [2]

El Big Data és el nou mitjà d'aquest segon decenni de segle. Un nou conjunt de tecnologies computacionals que, com les anteriors, ha canviat la manera en què accedim a la realitat. En un moment en què la web social és el nou laboratori de la producció cultural, les humanitats digitals s'aboquen cap a l'anàlisi de la producció i distribució massiva de productes culturals, per comprometre's en el disseny i qüestionament dels mitjans que l'han fet possible. D'aquesta manera, adopten un enfocament més centrat a veure com es produeix i distribueix la cultura, que les col·loca davant els reptes d'una nova cultura connectada.

5.264.802 documents de text, 1.735.435 arxius d'àudio, 1.403.785 vídeos i més de dos bilions de pàgines web accessibles a través de la WayBack Machine, constitueixen l'inventari actual de l'[Internet Archive](#) [3]. Hi podem sumar les obres de més de 7.500 artistes d'avantguarda, arxivades en forma de vídeos, pdf, arxius sonors, programes de televisió i ràdio en [UBUWEB](#) [4], o les més de 4.346.267 entrades en 241 idiomes, facilitades pels 127.156 usuaris actius que elaboren la [Wikipedia](#) [5] i les contínues contribucions dels més de cinc-cents milions d'usuaris del [Twitter](#) [6]. Aquests són sols alguns exemples dels nous espais virtuals on s'emmagatzemen i comparteixen coneixements. Arxius digitals, d'accés lliure i contribuïts de manera col·lectiva, wikis i xarxes socials, on conviuen hibridacions i trobades entre diferents mitjans i múltiples continguts. Donen lloc a un entorn complex, on la nostra cultura es mostra com un treball sempre en procés.

En els anys noranta l'ordinador era vist com una «maquina de remediació», és a dir, un mitjà capaç de reproduir els ja existents. Llavors les humanitats digitals es van centrar en la traducció dels documents del nostre patrimoni cultural, tributaris de la cultura de la impremta, al mitjà digital. Un procés de mediació en què els documents van ser reduïts a dades discretes, llegibles i operables per màquines. La computació en humanitats, tal com la descriu Roberto A. Busa en la introducció a [A Companion to Digital Humanities](#) [7], és l'automatització de tota anàlisi possible de l'expressió humana. Aquesta automatització va aportar noves capacitats als documents, que van anar mutant cap a un nou missatge performatiu, que va desplaçar el document escrit com a vehicle predominant del coneixement humà. Alhora, els mateixos processos dedicats a l'anàlisi i reproducció dels textos es van aplicar al desenvolupament d'eines que permetessin accedir a aquests continguts i compartir-los, amb la qual cosa es va produir un canvi de l'arxiu a la plataforma. Ambdues operacions canviaran la manera en què es duu a terme la recerca en les humanitats i definiran el contingut d'unes humanitats digitals 2.0.

La conversió de documents a codi binari mitjançant l'escanejat o la transcripció, el seu emmagatzematge en bases de dades, aprofitant les capacitats d'aquestes per a la recerca i la recuperació de la informació i l'aplicació de descriptors, etiquetes i metadades, van donar lloc a l'enllaçat i interoperabilitat dels mitjans. Es van crear arxius digitals com el [Salem Witch Trial](#) [8], dirigit per Benjamin Ray a la Universitat de Virgínia el 2002, un repositori de documentació referida a la caça de bruixes de Salem, o [The Valley of Shadow](#) [9] construït pel Centre per a la

Història Digital a la Universitat de Virgínia, i que documenta les vides de la gent del Comtat d'Augusta a Virgínia i del de Franklin a Pennsilvània, durant el període de la Guerra Civil americana. Es converteixen en espais virtuals navegables, en què s'accedeix a la informació de manera activa, a través d'una estructura que permet enllaçar continguts de diferents bases de dades, emmagatzemades en diferents formats, siguin imatges, textos o arxius sonors. Al mateix temps, la realització d'aquests repositoris, accessibles de manera gairebé ubíqua, suposa un treball col·laboratiu entre professionals de diverses disciplines, des d'historiadors, lingüistes i geògrafs fins a dissenyadors i enginyers informàtics. Aquesta trobada donarà lloc a pràctiques convergents i a la postdisciplinarietat que definirà les humanitats digitals. En aquest treball col·laboratiu en què s'hibriden procediments i formes de representació, apareixen nous formats on contextualitzar la informació, des de l'esquema cronològic fins als mapes geogràfics i les timelines. En aquest sentit, cal esmentar el projecte Digital Roman Forum, dut a terme entre el 1997 i el 2003 pel [Cultural Virtual Reality Laboratory](#) [10] (CVRLab) de la Universitat de Califòrnia, Los Angeles (UCLA). Aquest mostra una nova manera d'espacialitzar la informació, creant un model tridimensional del Fòrum romà que funciona com a interfície i que conté diferents càmeres, enfocades als diversos monuments reproduïts. Aquest dispositiu permet comparar la reproducció històrica amb fotografies actuals, al mateix temps que ofereix una relació dels diferents documents que fan referència a aquests espais i que han estat utilitzats per a la realització de la reproducció.

Aquesta capacitat d'accés i enllaç és portada més enllà de l'arxiu per projectes com Perseus o Pelagios, espais en què s'accedeix als continguts i s'hi contribueix de manera col·lectiva i lliure. Aquests fan ús dels estàndards sorgits a les comunitats de pràctica, per relacionar els seus continguts a través de diferents recursos en línia, amb la qual cosa esdevenen, més que no pas repositoris, autèntiques plataformes on produir i compartir coneixement. Així, la llibreria digital [Perseus](#) [11], iniciada el 1985, es basa en l'elaboració d'un programari open-source, que permet un sistema d'operació de dades extensible en un espai de treball en xarxa, a través d'un sistema que funciona a dos nivells, un d'accessible als éssers humans, que permet afegir-hi continguts i descriptors, i un altre que incorpora la capacitat de generar coneixement de les màquines. Aquesta plataforma dona accés als documents originals, relacionant-los amb múltiples informacions, com ara les seves traduccions i reedicions posteriors, edicions de comentaris, mapes dels espais citats... Tota aquesta informació és exportable en arxius de dades en format xml. El projecte [Pelagios](#) [12], dedicat a la reconstrucció del món antic, elabora un mapa en el qual dades geoespaciales històriques es relacionen amb continguts en altres fonts de recursos en línia. L'accés a un punt del mapa-interfície ens condueix a un conjunt heterogeni format per imatges, traduccions, citacions, bibliografies o altres mapes, tot exportable en diversos formats d'arxius de dades com són xml, json, atom o klm.

[>> SEGUEIX LLEGINT \[13\]](#)

- [14]

Source URL: <https://interaccio.diba.cat/en/blogs/2013/10/23/big-data-humanitats-digitals-de-computacio-social-als-reptes-de-cultura-connectada>

Links:

- [1] <https://interaccio.diba.cat/en/members/farrasdm>
- [2] <http://blogs.cccb.org/lab/author/sandra-alvaro/>
- [3] <http://archive.org/index.php>
- [4] <http://www.ubu.com/>
- [5] <http://ca.wikipedia.org/>
- [6] <https://twitter.com/>
- [7] <http://www.digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&hunk.id=ss1-1-2&toc.depth=1&toc.id=ss1-1-2&brand=default>
- [8] <http://saalem.lib.virginia.edu/home.html>
- [9] <http://valley.lib.virginia.edu/VoS/usingvalley/valleyguide.html>
- [10] <https://mail.cccb.org/owa/redir.aspx?C=b755dbe334d048a0bcd74a6ae49d14c5&URL=http%3A%2F%2Fwww.cvrlab.org>
- [11] <http://www.perseus.tufts.edu/hopper/>
- [12] <http://pelagios-project.blogspot.com.es/>
- [13] http://blogs.cccb.org/lab/article_big-data-i-humanitats-digitals-de-la-computacio-social-als-reptes-de-la-cultura-connectada/
- [14] <https://interaccio.diba.cat/node/4382>

