

# Cóncavo y convexo: Documentación y Humanidades Digitales, punto de inflexión

BOSCH, Mela / Directora del Centro Argentino de Información Científica y Tecnológica (CAICYT).  
CONICET – [mbosch@caicyt.gov.ar](mailto:mbosch@caicyt.gov.ar)

---

» Palabras clave: humano, social, software, documentación, infraestructuras de información.

## › **Cóncavo y convexo**

En este trabajo, que es una suerte de diálogo de compañeros de camino, van a aparecer viejos convidados, parte de mi historia personal y por supuesto la perspectiva de la institución de cuya gestión me ocupo, el Centro Argentino de Información Científica y Tecnológica (CAICYT), que forma parte de la colmena del CONICET.

Justamente los conceptos de *cóncavo* y *convexo* en este contexto los he tomado de la Dra. Dora Barrancos, Directora del Área de Ciencias Sociales y Humanas de CONICET, quien los mencionó en el curso de una clarificadora intervención durante una reunión de especialistas. En ese entonces, Barrancos se refirió a la reflexión como un espacio convexo en relación con la concavidad de los repositorios documentales. En lo personal me resultó inspirador para visualizar este ámbito que nos ocupa.

Esto en cuanto a referencia en el espacio. En cuanto al tiempo, me debo retrotraer más: en noviembre de 1971 se reunieron para debatir en un programa de televisión holandés dos de los más grandes intelectuales del siglo pasado, hablando en inglés uno y el otro en francés, Chomsky y Foucault (en riguroso orden alfabético), abordaron el tema de la *naturaleza humana*:



Imagen 1. Captura de pantalla del programa *On Human Nature*. Fuente: YouTube.

[https://www.youtube.com/watch?v=n\\_AScyqFUCc](https://www.youtube.com/watch?v=n_AScyqFUCc).

Para Chomsky, este es el conjunto de esquematismos innatos, de principios organizadores apoyado basado en un determinismo biológico, altamente discutido. Foucault, en cambio, se preocupa no por el *qué es* la naturaleza humana sino en *qué es ese concepto*, cómo hemos llegado a él, lo considera un indicador epistemológico más que un concepto científico.

Durante el intercambio Chomsky presenta como inherentes a la naturaleza humana las propiedades de las estructuras cognitivas que permiten a un niño adquirir estos complicados sistemas y crear nuevos pensamientos a partir de ellos. El concepto de creatividad queda entonces en el centro de la discusión entre ambos.

Foucault nuevamente se mueve del *qué* al *cómo*, hace notar que la historia del conocimiento ha debido enfrentar dos exigencias: la primera es la *atribución*, no solo la necesidad de datar y situar, sino también la de imputar a alguien lo realizado. Lo *no atribuido* se difumina en los conceptos de tradición o mentalidad: la creatividad queda anclada con la originalidad. La otra exigencia es la de *verdad*, con la paradoja de que lo que se considera verdad para serlo no puede depender de la voluntad de un individuo que la crea de manera original, aparece como un efecto complejo, múltiple, una grilla de acumulación de saberes y con esta acumulación llegamos al concepto de *archivo*.

Podría seguir hablando por mucho tiempo sobre el debate de estos colosos en aquellos años 70. Yo lo escuché grabado en una cinta mucho después, en 1978. Eran las épocas en que los exiliados internos en Buenos Aires respirábamos gracias a las cintas grabadas venidas de diferentes partes del mundo, que traían aire a los nichos de resistencia y resiliencia de Argentina. Mientras, Buenos Aires vivía la embriaguez del mundial de fútbol, al salir de la Alianza Francesa, donde había sido reproducido el debate, caminé hasta el hall de este Centro Cultural San Martín. Me refugié allí,

sintiendo que, a pesar de la dolorosa euforia que vivíamos en nuestro país, estábamos como humanidad asomándonos a una época impredecible y asombrosa; me sentí feliz por primera vez en mucho tiempo. A la vez descubrí que el debate entre ambos estaba instalado en mí, en la esquizoide cotidianidad de esa época de exiliada pero dentro de mi país.

¿Por qué traigo esta anécdota aquí? En Chomsky y Foucault, como en Piaget y Vygotsky quizás, parece estar esta tensión entre factores internos *humanos innatos* y *externos sociales*, donde la acumulación aparece como forma de hacer evidente lo social. Existe una tensión entre el *hardware* de la razón humana y el *software* de nuestras metodologías que, tal como insiste Foucault, necesitan de la acumulación de conocimiento. Un punto de inflexión, más allá del *cliché*, entre una concavidad y una convexidad.

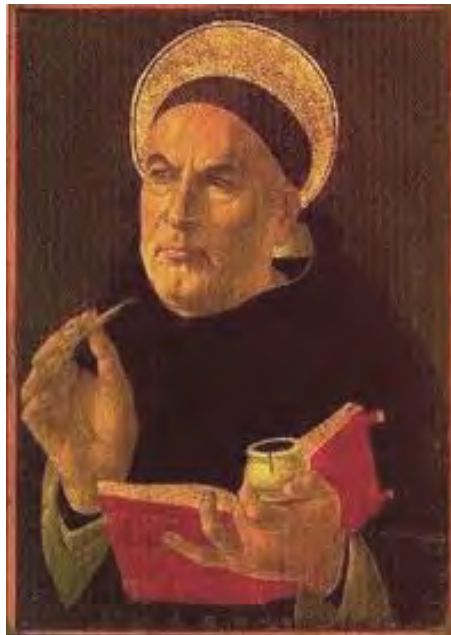
Me permito seguir con mi itinerario personal: diez años después de aquel 1978 tuve el privilegio de participar como lingüista en la primera experiencia argentina de desarrollo de una base textual documental en disco láser. En 1988 aún no había difusión masiva de internet, la editorial *El Derecho* decidió digitalizar su impresionante acervo de fallos, así como textos fundamentales del cuerpo legal argentino, compilando entonces por primera vez legislación, jurisprudencia y doctrina. Todo ello en una impresionante masa de 350 megabytes. Cantidad que deja libre 3 cuartas partes de mi *pen drive* de ahora, sin embargo, era por entonces una empresa no realizada en el país, y solo conocida en pocos lugares del mundo. El proyecto en el aspecto informático fue llevado adelante gracias a un pionero a quien recuerdo hoy con agradecimiento: Joza Vrljicak:



Imagen 2. *El Derecho en disco láser*. Fuente: fotografía propia.

Durante este desarrollo pude experimentar en el hacer diario la dinámica entre concavidad

y convexidad: primero fue imprescindible la captura de la masa documental analógica a digital y armar nuestro reservorio: debimos digitalizar con personas dactilografiando día y noche los 21 volúmenes impresos, más los códigos civil, comercial y procesal, y otros documentos. No había posibilidad de corrección manual ni ojos humanos que lo pudieran hacer en breve término considerando la cantidad de texto. Empezamos a reflexionar sobre los errores típicos de humanos tipeando, armando listados de estadísticas de ocurrencias y, con esto, generamos diccionarios de corrección. Con este archivo generamos otros donde esas palabras formaban parte del contexto semántico para uno de los primeros lenguajes de interrogación de bases documentales en lenguaje natural en español, tratamiento con pocos precedentes en lenguas latinas, excepto la experiencia del *Index Thomisticus* del Padre Busa que era nuestro mentor:



*Imagen 3. Padre Roberto Busa. Fuente: fotografía propia.*

Durante esa experiencia, que ahora tiene 25 años, y al vivir en el día a día la dinámica entre la acumulación de texto digital que por primera vez yo veía en una cantidad no manejable por manos humanas y que podía ser interrogado en el lenguaje natural de los textos, llegué al convencimiento de que era posible la emergencia de esa creatividad, de lo impensado que inquietaba a Foucault, a la vez que esperaba en una nueva capacidad de extraer esquematismos formales de esa masa de datos que nos permitieran generar nuevo conocimiento como esperaba Chomsky.

Para retomar el hilo de esta historia, y volviendo al histórico debate, quien era el moderador del mismo, Fons Elders, dijo algo importantísimo en aquel momento:

Perhaps the best way to compare both philosophers would be to see them as tunnellers through a mountain working at opposite sides of the same mountain with different tools, without even knowing if they are working in each other's directions (Rajchman, 2006: 1).

Aquí tenemos a estos *mineros* del conocimiento: por una vía, Chomsky nos dio las bases para realizar la mayor abstracción sobre la lengua humana. Por otra vía, Foucault avanzó recopilando documentación con su Grupo de Información en las Prisiones y de allí nos regaló el broche terrible del siglo XX, *Vigilar y Castigar*:



Imagen 3. *Mineros del conocimiento. Fuente: fotografía propia.*

A esta empresa me permito sumar otros *mineros*: desde la concavidad de la acumulación de conocimientos avanza Otlet, *el hombre que quería clasificar el mundo* (Levie, 2006), en cuya hermosa propuesta universal de diseminación de conocimiento, su túnel, siguió mucho después que su *Mundaneum* fue condenado al olvido por los nazis, pero sin pausa nos llevó hasta los lenguajes controlados que conducen a las ontologías de la web semántica.

Avanza también Suzanne Briet que con su impresionante labor de más de treinta años dio centralidad al concepto de *documento* unido al de *evidencia*. Esta feminista militante, participante de los movimientos de educación popular y convencida de la importancia de la democratización de la información, mantuvo una pasión por unir el documento a los hechos. Se hizo famosa su *metáfora del antílope* (Briet, 1951), que se presentó luego como un exceso de pragmatismo, comprensible quizás porque le tocó a ella realizar para la naciente UNESCO los inventarios sobre los recursos bibliográficos destruidos por la guerra.

Ahora nos planteamos adónde nos llevan estos túneles: ¿está primero el espacio convexo de la reflexión humana acerca de nuestra capacidad creativa? ¿nos permite este dar un salto cualitativo la concavidad de nuestros reservorios de memoria, ahora digitales? ¿cuál es la relación de estas dos topologías?

El espacio cóncavo de la documentación ha obrado de manera autorreferencial, tal como

hace notar la crítica de Ronald Day (2014) al enfoque positivista y pragmático de las escuelas bibliotecológicas herederas de Otlet. En su obra *Indexing it all*, Day critica el instrumentalismo de Otlet y el de Briet y, utilizando el concepto de Foucault, propone una *episteme de la documentación*. Examina el paso de las estructuras de índices que, de ser estructuras explícitas de organización, mediadoras entre los usuarios y los documentos, se transformaron en infraestructuras implícitas en los dispositivos que todos usamos en nuestros actos cotidianos de uso de información e incluso para la comunicación social. Estos lenguajes controlados y sus mecanismos de etiquetado se generalizaron en esta década del siglo XXI como instrumentos lógicos no solo para la indización y para la recuperación de información sino para la creación de *sujetos sujetos* en las redes sociales (Day, 2005).

Otra aproximación con una visión menos disruptiva, en un intento de integración, es la que ofrece Chris Alen Sula (2012) del Pratt Institute de New York. Sula intenta un modelo que dé cuenta de la vinculación entre Humanidades Digitales y Bibliotecas. Retoma en parte la distinción de Suzanne Druet dejando para los documentalistas el mundo de los documentos secundarios y para los humanistas digitales los primarios:

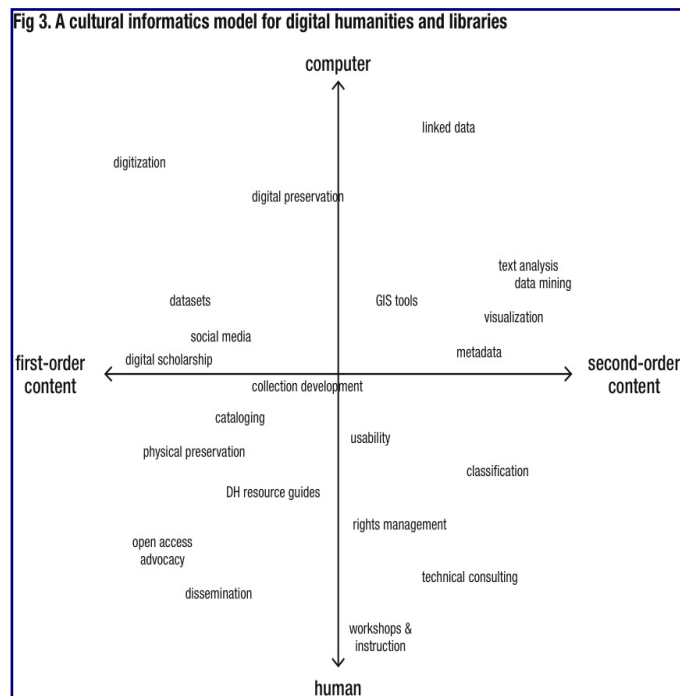


Imagen 4. Fuente: Sula (2012).

### › **Sobre las diversas competencias de las Humanidades Digitales**

Las Humanidades Digitales se han preguntado mucho sobre su *naturaleza* y así también lo

ha hecho la Información Documental, que ni siquiera logra un nombre (Ciencias de la Información, Ciencias Documentales, etc.) y tampoco logra etiquetar su protagonista: el especialista en información, o el hermoso, llano y denostado bibliotecario.

Creo que en esta inflexión estamos en un momento donde podemos ver qué tenemos para dar de parte de cada una de nuestras comunidades. En ese sentido, Sula (2012) hace un estudio sobre las competencias relacionadas con las Humanidades Digitales, toma como fuente las *core competencies* desarrolladas por ALA (2009), que, a partir de lo expuesto por Sula, me he permitido traducir y sintetizar:

1. Ciclo de vida de las colecciones de información en los diferentes estadios de selección, compra, uso y disposición.
2. Mantenimiento, preservación y conservación de colecciones.
3. Desarrollo de herramientas para la organización del conocimiento registrado en recursos de información.
4. Desarrollo de sistemas de catalogación, metadata, indización.
5. Técnicas de identificación y análisis de tecnologías emergentes e innovaciones para reconocer e implementar mejoras tecnológicas necesarias en Humanidades Digitales.
6. Competencias en técnicas y métodos estadísticos aplicados a la información documental.
7. Competencias para alcanzar audiencias específicas para presentar conceptos y servicios.
8. Competencias para evaluar y responder a diferentes necesidades de comunidades y preferencias de usuarios.

## › ***Las competencias de las Humanidades Digitales y las líneas de investigación del CAICYT***

Ahora veamos cómo es posible vincular estas competencias con la actual línea de trabajo que tenemos en CAICYT.

En los últimos años de uso intensivo de las tecnologías, nos encontramos en un escenario que considera a la información como *commodity*, como mercancía de uso comercial, como bien producido masivamente cuyo valor está en la relación oferta-demanda, sin diferenciación en función de quién o dónde se produce.

Salir de esta presión es un desafío para los que formamos parte de la comunidad de actores de la producción de información científica, que nos requiere convertirnos en desarrolladores de capacidades diversificadas que no sean transformables en *commodities*. Esta estrategia se basa en incidir en la cadena de valor con mayor calidad.

Como institución, esta meta nos lleva a mirar hacia el pasado y hacia adelante: el CAICYT tuvo su origen en 1958 como Biblioteca del CONICET, pasó a ser un centro de servicios documentales en 1976 y, recientemente, en 2012, fue promovido a unidad de investigación, además de centro de servicios.

Esto nos presenta un promisorio horizonte y, a la vez, nos impone enfatizar la consideración de la información como un producto social que se implementa en servicios a la comunidad altamente diferenciados.

Por ello estamos trabajando en cuatro ejes:

1. Apoyo a *competencias para la gestión de la Comunicación Científica y Tecnológica*: Implica la puesta en disponibilidad de servicios operacionales como ayuda, por ejemplo, al manejo de citas, así como tutoriales y materiales de autoformación.
2. *Consolidación de las infraestructuras de información*: Esto implica la disponibilidad uniforme y persistente de estructuras de descripción estables (metadatos) e instrumentos de denominación apropiados (listas de valores, vocabularios controlados, tesauros, ontologías) lo cual permitirá expresar las líneas de investigación y acción científica del CONICET en forma consistente, y favorecer el diálogo entre agendas científicas a nivel global.
3. Promoción de *entornos de exposición de la comunicación científica y tecnológica*: Se trata de proveer de medios técnicos acordes a los actuales modelos de mostración y circulación de publicaciones en la comunidad científica global. Se suministran los medios técnicos de asignación de metadatos para digitalizar recursos, así como para su persistencia en repositorios y en el desarrollo de flujos documentales no redundantes, significativos y reusables.
4. Apoyo a la *creación de instrumentos de autoconocimiento del sistema científico*: Propende el desarrollo de métricas útiles para la caracterización de la producción formalizada en ciencia y técnica. Esta línea de trabajo está orientada a proveer de instrumentos para la toma de decisiones como así también disponer de modelos que permitan condiciones de comparabilidad entre actores científicos en el curso del tiempo.

También estamos llevando a cabo, junto con otros investigadores CONICET, como la Dra. Gimena del Rio Riande (SECRET-IIBICRIT), y junto con las Áreas Comunicación Institucional y de Tecnología Documental de CAICYT e investigadores de distintas universidades argentinas, el proyecto institucional *Metodologías en Herramientas Digitales para la Investigación* (MHeDi). El proyecto contempla una línea temática dirigida a trabajar en general con metodologías en herramientas para producción de datos y desarrollo de esquemas de metadatos para proyectos de investigación. En esta línea CAICYT está trabajando en el desarrollo de un marco de verificación de calidad de metadata e infraestructura para datos primarios científicos. La otra línea temática está orientada al estudio de metodologías para el uso de herramientas digitales en el ámbito de Ciencias Sociales y Humanidades. Comenzando con la localización de productos centrales para el desarrollo de las Humanidades Digitales en los países de habla hispana, se ha procedido a la traducción del



directorio de herramientas digitales del directorio DIRT (<http://dirtdirectory.org/es>) y a la traducción y exposición de la Taxonomía en Humanidades Digitales de DARIAH (TADIRAH), desde el Servidor Semántico del CAICYT (<http://vocabularios.caicyt.gov.ar/portal/?v=TAD>).

Para todo esto debemos individualizar a los actores de la información científica argentina: autores, investigadores, áreas de desarrollo científico y directorio de CONICET, editores científicos, bibliotecas especializadas en ciencia y tecnología, etc. Así, nuestro interés es el de dar continuidad, mejora e incremento en la relación con cada una de esas comunidades. El énfasis está puesto justamente en cada uno de los ejes que indicamos arriba, que requieren una explicación mayor.

## › ***Las Humanidades Digitales y el desarrollo de infraestructuras de información***

Nos vamos a detener ahora, por su fuerte vinculación con las Humanidades Digitales en nuestra política de desarrollo de infraestructuras de información.

Respecto de qué son las estructuras de información, cabe decir que es un concepto que hemos tomado de Bowker *et al.* (2010) quienes parten del concepto general de infraestructura como el conjunto de equipamientos necesarios para la actividad humana.

Pero estos equipamientos, las infraestructuras muestran rasgos específicos: tienen que estar embebidas dentro de otras estructuras sociales, tecnológicas, etc., por lo que resultan transparentes. Tienen asimismo un ámbito de aplicación espacial o temporal más allá de una práctica aislada.

Otros aspectos importantes de las infraestructuras es que deben ser aprendidas de manera inclusiva en una determinada comunidad de práctica, no funcionan de forma aislada. Además se articulan a través de estándares y normas. Esto hace que su cambio sea gradual y requiera negociación y procesos de ajuste y para funcionar se deben articular sobre bases de infraestructuras ya instaladas. Finalmente, algo que todos sabemos: se hacen visibles cuando fallan.

Es evidente que en esta caracterización entran: infraestructuras físicas, como edificios, caminos o puertos, a redes de transporte y comunicación como cables, etc. Así también como las grandes infraestructuras en la ciencia, como aceleradores de partículas, telescopios orbitales entre otros.

Apoyándose en otros, estos autores extienden el concepto llevándolo a entidades más abstractas como protocolos, estándares y memoria (Bowker *et al.*, 2010: 97-117).

Por esta vía llegamos nuevamente a nuestras *concauidades*, proponiendo estas infraestructuras de información, también denominadas *ciberestructuras*, que tienen una presencia invisible como los cableados, pero que modelan y permiten construir, transportar y acumular conocimiento. Así tenemos las redes de computadoras, pero también el software para desarrollo de aplicaciones con alta normalización y protocolos que permiten el reuso entre plataformas.

Son estas ciberestructuras no transparentes desde los lenguajes de marcado hasta las geolocalizaciones y los datos enlazados las que nos permiten hoy estar en este escenario de las Humanidades digitales.

Y aquí estamos los horribles bibliotecarios con nuestra experiencia y nuestras instituciones que se transforman en sedes de infraestructura.

Se trata de una empresa de construcción colaborativa de conocimiento. Salimos de la pasividad de ofrecer nuestras competencias tan prolijamente presentadas por ALA a pensar formas de viabilizar y de apoyar estrategias de aprovechamiento colectivo y sustentable.

CAICYT ha iniciado el desarrollo de algunas de estas infraestructuras: el Portal de Publicaciones Periódicas, <http://ppct.caicyt.gov.ar/>, para facilitar la puesta en marcha de publicaciones científicas, que tiene embebidas las prácticas de referato en ella. Ofrecemos también en el Servidor Semántico, <http://vocabularios.caicyt.gov.ar/portal/home.php>, que nos permite brindar una plataforma que suministra de manera simple e intuitiva listas de descriptores de tesauros, taxonomías, glosarios y ontologías terminológicas. En el ya mencionado Servidor Semántico cada recurso tiene un referente científico para su localización, uso y reuso. Es una infraestructura porque permite servicios de consulta, desarrollo, descarga y/o explotación de servicios web semánticos para que cada actor pueda desplegar los modelos y dinámicas de utilización que considere necesarias para facilitar la producción, transmisión, tratamiento y gestión del conocimiento especializado.

Otras infraestructuras que ponemos a disposición son una plataforma para análisis de datos documentales, que es utilizada para el proyecto de autoconocimiento del sistema de bibliotecas CONICET (ReFeRI), así como estudios de usuarios y relevamientos documentales. También plataformas para formación y para trabajo colaborativo. Además, ofrecemos una plataforma para datos documentales que puede funcionar como repositorio de un amplio espectro de tipos documentales y soportes.

Creo que aquí podemos finalmente decir que CAICYT espera ser un referente estable y consistente, con claro reconocimiento, poniendo a disposición servicios con valor agregado, sustentados de una infraestructura de recursos de información y humanos altamente profesionales abiertos a la comunidad con responsabilidad social.

## › **Bibliografía**

American Library Association (ALA) (2009). *ALA's core competencies of librarianship*. Recuperado de <http://www.ala.org/educationcareers/careers/corecomp/corecompetences> el 31/08/2015

Bowker, G., Baker, K., Mitterand, F. & Ribes, D. (2010). Toward Information Infrastructure Studies. Ways of Knowing in a Networked Environment. En Hunsinger, J., Klastrup L. & Allen, M. M. (Eds.), *International Handbook of Internet Research* (97-117). Heidelberg-London-New York: Springer.

Briet, S. (1951). *Qu'est-ce que la documentation?* Paris: Éditions documentaires, industrielles et techniques.

Day, R. E. (2005). Poststructuralism and Information Studies. En Cronin, B. (Ed.), *Annual Review of Information Science and Technology*, 39, 575-609.

Day, R. E. (2014). *Indexing It All: The Subject in the Age of Documentation, Information, and Data (History and Foundations of Information Science)*. Cambridge-Massachusetts: Verlag-MIT Press.

Levie, F. (2006). *L'homme qui voulait classer le monde: Paul Otlet et la Mundaneum*. Bruxelles: Impressions Nouvelles.

Rajchman, J. (2006). *The Chomsky-Foucault Debate. On Human Nature*. New York: New Press.

Sula, C. A. (2012). *Digital Humanities and Libraries: A Conceptual Model*. New York: School of Information & Library Science, Pratt Institute. Recuperado de <http://chrisalensula.org/digital-humanities-and-libraries-a-conceptual-model/> el 31/08/2015