Letture Magistrali in Biblioteconomia

XVI

La *Lectio Magistralis* in Biblioteconomia è promossa da Università degli studi di Firenze, Dipartimento di Storia, Archeologia, Geografia, Arti e Spettacolo (SAGAS), Associazione italiana biblioteche, Biblioteca nazionale centrale di Firenze, Biblioteche oggi, Casalini Libri, ICCU, JLIS.it, SISBB

La serie *Letture Magistrali in Biblioteconomia* è curata da Mauro Guerrini

Homo Deus
Techno-Humanism vs Dataism and the place of Information in Society

Lectio magistralis in Library Science

by

Philip E. Schreur

Florence, Italy
University of Florence
7th March 2023

Casalini Libri
2023

Schreur, Philip Evan, 1957-
Homo Deus : Techno-Humanism vs Dataism and the place of Information in Society : Lectio magistralis in Library Science = Homo Deus : Tecno-Umanesimo vs Dataismo e il ruolo dell'informazione nella società : Lectio magistralis in biblioteconomia / Philip E. Schreur.
Fiesole (Firenze) : Casalini Libri, 2023.
52 p. ; 24 cm. — (Letture magistrali in biblioteconomia; 16)

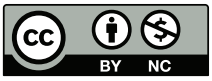
1. Information science
2. Information technology
3. Metadata
4. Information society
5. Harari, Yuval N. Hiṣṭoryah shel ha-maḥar

1. Scienze dell'informazione
2. Tecnologia dell'informazione e della comunicazione
3. Metadati
4. Harari, Yuval N. Hiṣṭoryah shel ha-maḥar

ISBN 978-88-7656-102-3 (print)
ISBN 978-88-7656-103-0 (ebook)

Z665.S37 2023
020

Pubblicazione disponibile in accesso aperto al permalink http://digital.casalini.it/9788876561030 sul portale TorrossaOpen (oa.torrossa.com) in base ai termini della licenza Creative Commons Attribuzione – Non commerciale.



Disponibile anche in versione a stampa, inviando richieste a orders@casalini.it.

ISBN: 978-88-7656-102-3 (print).

Copyright © 2023, Casalini Libri, Fiesole (FI), Italia.
Prima edizione: marzo 2023.

Tutti i numeri della serie Letture Magistrali in Biblioteconomia sono disponibili sui portali access.torrossa.com e www.torrossa.com.

L'editore aderisce al Servizio nazionale coordinato di conservazione e accesso a lungo termine per le risorse digitali (depositolegale.it).

Lecture Series: *Magisterial Lessons in Library Science*

The Department of Storia, Archeologia, Geografia, Arti e Spettacolo (SAGAS), the Masters Program in Archival Studies, Library Science and Codicology, and the Library Network of the University of Florence are pleased to present a series of lectures entitled *Letture Magistrali in Biblioteconomia* (Magisterial Lessons in Library Science), in collaboration with the Biblioteca Nazionale Centrale di Firenze, *Biblioteche oggi*, Casalini Libri, Italian Library Association (AIB), *JLIS.it*, and SISBB.

The lectures are geared toward scholars engaged in the critical and methodological study of the theories and principles of library science and toward professionals involved in the experimentation, verification and application of scholarly hypotheses within the laboratory/library. The necessary and positive relationship between scholars and professionals creates a virtuous circle that favours the advancement of the discipline.

The *Magisterial Lessons* are also open to students in Library Science and for all those interested in the proposed topics.

Mauro Guerrini
Professor in Library Science,
Cataloguing and Metadata Creation
University of Florence, Italy

Homo Deus: Techno-Humanism vs Dataism and the place of Information in Society

Philip E. Schreur

Stanford University

Abstract

Sixty years after the release of one of the most impactful science fiction movies ever made, *Forbidden Planet*, Yuval Harari published the English translation of his book, *Homo Deus*. The future of homo sapiens predicted in *Homo Deus* has striking parallels with the fate of the inhabitants of the forbidden planet called the Krell. Driven by an unquenchable desire for immortality and power, both civilizations have embraced a movement Harari defines as Techno-Humanism based on the enhancement of mind and body using mind-altering substances and mechanical prosthetics. But Harari proposes an alternative to this potentially disastrous fate, a new religion called Dataism. Dataism shifts the focus from the individual, as has long been supported by Humanism, to the flow of information across society as a whole. Libraries, and Information Science, will have a key role in this new world through the championing of access to data, linking of data, and discovery of data. This uninhibited flow of data, and the intersubjective webs of meaning built upon them, will help us reach the stars.

1. Introduction - *Forbidden Planet*

On March 3rd, 1956, arguably one of the most impactful science fiction movies ever produced, *Forbidden Planet*, held its premiere in Charlotte, North Carolina. The screening took place as part of the second annual Southeastern Science Fiction Conference at a small Metro-Goldwyn-Mayer (MGM) screening room that held approximately forty people.

Forbidden Planet was a film of many firsts: it was the first to include faster-than-light travel, it was the first to be set entirely on a planet other than earth, it was the first to depict a robot with a personality, namely Robbie the Robot, and it was the first movie of any kind to include a score that was generated entirely electronically.

The plot of *Forbidden Planet* is deceptively simple. According to the Internet Movie Database (IMDb), a starship crew in the 23rd century is sent to investigate the sudden cessation of communication from an outpost established by Earth on a distant planet circling a star named Altair IV. Once they arrive, the captain discovers only two human survivors, a powerful robot, and the remains of a lost civilization. Shortly after the starship lands, it is revealed that all but two of the original colonists have been horribly murdered and the starship's crew soon begins to die in the same manner.

One of the two surviving colonists, Dr. Morbius, has discovered the remains of an extinct civilization called the Krell. Apparently, at the apex of their scientific development, the perfect blending of the organic with the machine, their entire civilization crumbled in one bloody night. And whatever was responsible for the extinction of the Krell has broken loose once more on the planet, relentlessly tearing all human life to pieces. The rescue team, with all of their powerful tools and technology, is powerless to stop it.

2. Homo Deus

Sixty years after the premiere of *Forbidden Planet*, Yuval Harari, a professor at the Hebrew University in Jerusalem, published the English language version of his book, *Homo Deus*. This book, whose

subtitle is "a brief history of tomorrow", could have been written as the rise and fall of the Krell. The parallel between the Krell and Homo sapiens is striking. The Krell focused all of their attention on scientific development. And as their civilization spread over the entire planet, they both sought a source of unlimited power and the ability to merge with the machine, eventually leaving organic matter behind. It was their ultimate success that led to their downfall. Harari sees human civilization rapidly approaching this same juncture. Companies such as Calico, a subsidiary of Google, aggressively combat aging and eventually may engineer immortality. Prosthetic limbs are becoming powerful replacements for missing arms and legs. The Meta corporation, the parent company of Facebook, has created an Artificial Intelligence (AI) called CICERO that can engage in a game of strategy, called Diplomacy, and negotiate equally as well as human players. An object called a stentrode, the combination of a stent and an electrode, when situated in the brain, could directly connect human thought and the machine. Homo sapiens seems insistent on pursuing the pathway of the Krell. In his book, Harari clearly illuminates the road we are on and offers two possible futures. But what has led us to this point in time, and what is our role as information professionals in negotiating our future?

2.1 Early Human History

Throughout human history, according to Harari, we have been driven by three implacable forces: famine, plague, and war. And throughout our evolution from hunter-gatherers to an agrarian society, and then to the global culture of today, these three have relentlessly driven our history.

Famine has been an ever-present scourge. Our transition from hunter-gatherers to an agrarian society has given us more predictability in our food sources, but the explosion of the human population and unpredictability of the climate have always strained our resources. Similarly, the explosion of the human population and the creation of a truly global culture have made pandemics such as the Spanish Flu, or HIV, or COVID much more likely. And human aggression to others of its kind, highlighted by the current war in

Ukraine, appears ceaseless. But for all this, Harari demonstrates that we have left this part of our history behind. Statistically, these events are having far less effect on the human population than they have in the past.

2.2 What is a Human?

But before Harari turns his eye from the past to the future, he ponders two key elements we must explore, the first being, what does it mean to be human? What are those elements that have allowed both Homo sapiens and the Krell to dominate their respective worlds?

There have been many theories over time as to what makes us human: some have attributed it to the presence of a soul, some to our consciousness, others to the ability to communicate through speech. All of these, however, have been proven either to be false, such as the exclusive ability to communicate orally, or only to be accepted through faith, such as the presence of a soul. Instead, Harari prefers to focus on two characteristics of humans that make us unique: first, the ability to work together in large groups for a common cause, and second, our ability to create intersubjective webs of meaning.

It is true that many other creatures live and work in larger groups: ants form colonies, bees form hives under the governance of a queen, wolves form packs and stake out territories in which to roam, but the similarities stop there. Pods of whales across the world's oceans do not cooperate to farm krill as a food source, wolves do not create passports to safely travel through each other's territory, bees do not explore alternative ways of governing their hives. Humans' unique ability to work together in large groups to fulfill common goals has allowed us to dominate our world. And according to Harari, our mechanism for accomplishing this is the second characteristic that makes us unique: our ability to create intersubjective webs of meaning.

What gives a rectangular piece of paper we call the Euro its value? Does it have any intrinsic worth? Very little. It is our common belief

that it has value that allows us to use it as a means of barter for goods across cities, states, and countries. Does the country of France actually exist? It only exists in that the people of France believe that it exists and others in the world accept that belief as well. We place so much trust in these intersubjective entities that we even give them legal rights. Intersubjective entities such as governments can own property; business entities such as Google can be sued; academic entities such as libraries can impose classification systems. These intersubjective webs of meaning give purpose to our lives, allow our civilizations to grow, and even allow us to build starships to reach forbidden planets.

And in our own world of Information Science, what gives the Paris Principles its authority? Does the Library Reference Model hold sway by divine right? Does Resource Description and Access, either in its current or former state, have any absolute meaning? It is the firm belief of individuals such as us in the benefits of adhering to common principles that gives them their value.

2.3 Humanism

A second key element in the rise of Humans and our Information Age is the focus on the individual, that is, Humanism. Humanism, simply put, gives inherent value to the human as opposed to the divine. And the philosophical movement of the 18th century known as the Enlightenment turned this humanistic focus to science, reason, and intellectualism.

Humanism itself has taken three main turns in the twentieth century. The first and most familiar is Liberal Humanism. Liberal Humanism gives value to the individual and the individual experience. Your internal feelings and judgments are as important as anyone else's, and your own personal thoughts and experiences validate your feelings and judgments. Social humanists put more emphasis on the collective human experience than the individual. Evolutionary humanists would say that neither the individual experience nor the collective experience are what is important. They believe that certain individuals, or races, thoughts and feelings carry more weight than others because they are more evolved.

It appears, though, that Humanism and the focus on the individual may have run its course. Although we are still told that our personal experience is important, economic superiority is created through the mechanization of industry through AI. War is increasingly driven by mechanical means such as drones as opposed to troops on the ground. Social media such as Twitter turn the collective against the individual.

We have reached a critical juncture in our history; our pathway is diverging. According to Harari, the path forged by Humans and Humanism has reached its apex and two new paths are emerging: Techno-humanism and Dataism. And information and knowledge management will be critical to both.

2.4 Techno-Humanism

Techno-humanism, in a way, is the ultimate expression of humanism as it still sees humanity as the pinnacle of creation. Humans are rapidly fulfilling their perceived destiny by conquering and exploiting the universe. And yet, at the same time, we have reached the limits of this path. With the current devaluation of the individual and the dizzying explosion of technology, we are approaching a breaking point in our evolution. The quickened pace in the development of AI and Machine Learning (ML) have produced something Harari describes as the separation of intelligence from consciousness. This separation leaves us with only one recourse, and that is to force nature's hand and take control of our own evolution.

As part of this evolution, Harari sees humans replacing the three fears of famine, plague, and war, with the three new goals of immortality, happiness, and divinity. And these three goals can be realized, just as the Krell from *Forbidden Planet* realized them, through the fusion of the human and the mechanical. According to Harari,

> Homo deus will retain some essential human features but will also enjoy upgraded physical and mental abilities that will enable it to hold its own against the most sophisticated non-conscious algorithms. Since intelligence is decoupling from consciousness,

> and since non-conscious intelligence is developing at breakneck speed, humans must actively upgrade their minds if they want to stay in the game.

Open science, open data, open technology will be key to realizing this revolution.

But although this evolution seems inevitable, it will have two distinct disadvantages. First, it will be expensive. The genetic and mechanical upgrading of our bodies, the fusion of our minds with supercomputers, will be cost prohibitive. This will be an evolution for the rich; the poor need not apply. The second disadvantage can be learned from *Forbidden Planet* and the fall of the super-race called the Krell. For the Krell, the loss of their corporeal form, together with the ultimate merger of their minds with the unlimited power and intelligence of the machine, allowed both their evolved consciousness and their primitive subconsciousness to have equal access to limitless power. And in a single bloody night, the Krell destroyed themselves. Harari sees the human mind as equally complex, a mixture of conflicting desires barely kept in check. To allow all parts of this fragmented consciousness equal access to unlimited power would prove disastrous.

Fortunately, Harari sees an alternative path for us, Dataism.

3. Dataism

Dataism is defined as a new religion that deifies neither gods nor man, it deifies data, data in all its forms: economic data, medical data, personal data... And Dataism has one and only one commandment: thou shalt privilege the right of information to circulate freely over the rights of humans to own data and to restrict its movement. In other words, there is no private data. All data should be open and shared. And Harari implies that this data is not necessarily data that is collected and anonymized. It includes your personal data attributable to you. It is through this unrestricted flow of information that we achieve our new goal: the Internet-of-All-Things. The individual loses their position as the center of the universe as Humanism would have us believe,

and instead takes a role as a collection of data points in the vast global network of humanity. Here, then, we begin to see our role as information scientists in this global network of data. We become essential workers in the flow of information, ensuring that it is open, understandable, and equally available to all.

3.1 Expression of Dataism in the Libraries

Information science has become a key focus in almost all Library science degree programs. What is surprising, is that this term first arose nearly fifty years ago. In the United States, the Association for Information Science and Technology (ASIS&T) was founded in 1937 as the American Documentation Institute. One of its earliest definitions of Information Science comes from 1968:

> Information science is that discipline that investigates the properties and behavior of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability. It is concerned with the body of knowledge relating to the origination, collection, organization, storage, retrieval, interpretation, transmission, and utilization of information.

A definition from 2009 has changed very little:

> Information science is the science and practice dealing with the effective collection, storage, retrieval, and use of information. It is concerned with recordable information and knowledge, and the technologies and related services that facilitate their management and use. More specifically, information science is a field of professional practice and scientific inquiry addressing the effective communication of information and information objects, particularly knowledge records, among humans in the context of social, organizational, and individual need for and use of information. The domain of information science is the transmission of the universe of human knowledge in recorded form, centering on manipulation (representation, organization, and retrieval) of information, rather than knowing information.

It is interesting that both these definitions focus on the flow of information, not the understanding of what the information contains.

3.2 Dataism at Stanford

I'd like to spend the second part of my talk speaking more practically to Dataism in our current environment. As you know from my background, I have spent much of my career at the Stanford University Libraries. It has proven to be a wonderful environment both for the more traditional areas of libraries such as cataloging and collection development, but also for dynamic experimentation into the future of libraries. Considering how we support the flow of information at the University is key to the library's mission, and our efforts can be divided into three broad strategies: access to data, linking of data, and discovery of data. In the following section, I'd like to share some concrete examples of how we will be supporting each of these strategies in the next few years.

3.2.1 Access to Data - OSTP Memo

First, access to data. In August of 2022, the White House Office of Science & Technology Policy (OSTP) issued a new policy called the Nelson Memo. The memo begins by referring to an earlier policy issued from the office in 2013. This earlier policy, called the Memorandum on Increasing Access to the Results of Federally Funded Research:

> directed all US federal departments and agencies with more than $100 million in annual research and development expenditures to develop a plan to support increased public access to the results of federally funded research, with particular focus on access to scholarly publications and digital data resulting from such research.

Although difficult to implement, this policy resulted in more than 8 million scholarly publications becoming openly available to the

public with over 3 million people reading these articles for free every day. This original policy, however, did allow for a 12-month embargo on free public access to these scholarly publications allowing publishers to develop funding models that eventually could encompass this federally mandated open access policy.

According to the new memo issued this year, in 180 days after the memo's publication for federal agencies with a budget of over $100 million dollars and within 360 days for federal agencies with a budget of $100 million or less, all federal agencies must produce a plan by which their federally funded research becomes publicly available with **no embargo period** or delay after publication. Because of leeway in putting these plans into action, there will be a rolling implementation date of between 2024 and 2028, but at the end of this time, all federally funded research of any kind by any agency must become freely available at point of publication.

The memo claims to be driven by six main elements: equitable access, rapid sharing, open science, public trust, taxpayer return, and research integrity. And in order to achieve these goals, publications must be made freely available, the data supporting them must be made freely available, metadata for the articles and the data supporting them must be made freely available, and persistent identifiers must be used for individuals, awards, organizations, and outputs.

Here is a chart showing the percentage of federally supported research at major research institutions in the United States. You can see from this chart that Stanford will be heavily impacted by this memo. The chart records total publications vs federally funded publications from 2017-2021. Stanford University is near the top of the curve with 75,672 total publications of which 32,493 received federal support. In other words, 43% of Stanford publications receive federal funding.

How the Stanford Libraries respond to this new mandate will be critical for its implementation. We see this new policy affecting Stanford Libraries in four main areas: our Open Access Policy, our Digital Repository, our Research Data Services, and our Identifier Services.

First, our Open Access policy mandates that the libraries help faculty researchers understand these new policies and navigate complex publisher requirements. We have recently hired a new Director for the Office of Scholarly Communications and the implementation of the Nelson memo will be her top priority. We suspect that there will be:

- increased urgency around defining a consistent approach to open-related licensing and Stanford support for researcher publication charges,
- increased need for publication tracking, reporting, and publication feeds to faculty webpages,
- and an increased need to investigate cost models for research data & repository services.

These three foci will drive much of her initial outreach efforts.

Second, our Digital Repository will be enhanced to support the memo requirements by providing researchers with:

- unique persistent identifiers,
- metadata,
- free access to content both to Stanford University and beyond,
- authentication during access,
- long-term sustainability of research and the data supporting it,
- reuse of content,
- security,
- a common format,
- and a data retention policy.

Third, our new library division, Research Data Services, will provide:

- data management planning,
- research data curation consulting services,
- metadata & data description services,
- and support for data access and manipulation.

And last, our identifier services will need to be enhanced. We already provide Open Researcher and Contributor Identifiers, or ORCIDs, and Digital Object Identifiers, or DOIs, but we will need to expand these services to identifiers for awards, organizations, and outputs.

Although the Nelson memo will not mandate open access and data flow for all data, it will ensure open access to the bulk of research data produced at Stanford, and in the United States, and the services the Stanford Libraries develop to provide this access could become the backbone for a future network of all data.

3.2.2 Linking of Data - Linked Data for Production (LD4P)

The second great prong in the Library's approach to Dataism is the linking of this new, global world of data to which we will have access, and to me, that has always implied the use of linked data. For years now, I have pursued the transition of data traditionally held by libraries and captured in Machine Readable Cataloging, or MARC, to linked open data. This interest culminated in a series of grant proposals to the Mellon Foundation called Linked Data for Production, or LD4P. But LD4P has deeper roots at Stanford.

It began in 2011, when the Stanford University Libraries and the Council on Libraries and Information Resources (CLIR) held a weeklong conference on the prospects for a large scale, multi-national, multi-institutional prototype of a linked data environment focused on libraries and the resources of most need to them. From the very beginning, our interest was in this open, global ecosystem. Although the knowledge a university creates advances our understanding of the world, it's only a part of a global web of knowledge that Harari discusses and derives its ultimate importance from its place in that open Web.

One of the great outcomes of this first conference was a concise summary of the benefits of linked data:

- First, the Linked Data approach offers significant advantages over current practices for creating and delivering library data while providing a natural extension to the collaborative sharing models historically employed by libraries.

- And second, linked data, and especially Linked Open Data, is **sharable**, **extensible**, and **re-usable**. It supports multilingual functionality for data and user services, such as the labeling of concepts identified by language agnostic URIs.

Soon after this conference, I turned my attention to defining how this transition might practically happen in libraries. And with the help of the library community, both within the United States and beyond, and with valued commercial partners such as the SHARE-Virtual Discovery Environment, or SHARE-VDE, we began.

LD4P has had three stages so far. The first was simply called *Linked Data for Production* and lasted from 2016-2018. This first phase focused on assessing the tools and architecture we would need to move into a true production system.

The next phase was called *Pathway to Implementation* and went from 2018-2020. This second phase focused on the completion of Minimal Viable Products (MVPs) for the tools we were developing and clearing the pathway to production. It also allowed us to expand the project to a cohort of libraries pulled from membership in the Program for Cooperative Cataloging (PCC).

And the current phase of LD4P is called *Closing the Loop* (2020-2022). Its focus is on the lifecycle of library data, from Acquisitions thru to discovery, to ensure that all aspects are supported in a cohesive environment.

All the grants have focused on the transition of basic Technical Services workflows from their current infrastructure built on MARC to linked open data and the Web. Our goal, then, was to transform a library's data and infrastructure into something that was both semantically articulate but could also link to the other freely open data appearing on the Web. LD4P has grown tremendously over the past few years and has had many partners, both nationally and internationally, but our focus has always remained on open data and open access tools.

The most critical development was the creation of an environment that allowed catalogers to catalog directly in linked data. And it was essential to us that this tool would be available via the Cloud as many

libraries don't have the technical staff to set up the environment locally. Although still being refined, that editor is ready for use and is accessible to anyone worldwide for experimentation at sinopia.io.

As of this year, Stanford has begun cataloging production making use of BIBFRAME, a linked data ontology, and Sinopia, focusing initially on literary works and lectures. Over the course of the spring, we will be establishing the entire workflow and sending these BIBFRAME descriptions to OCLC and SHARE-VDE so that they can develop means for ingesting our catalog production in BIBFRAME.

Questioning Authority (QA) is our open-source tool designed for searching and bringing entities into the Sinopia environment. The development for QA is being done by teams at Cornell and the School of Library and Information Science at the University of Iowa. Data on entities from various authoritative open sources are standardized and cached so that the cataloger can query, explore, and import data in a standardized way.

The final goal of Linked Data for Production, helping to fulfill Harari's vision of a true, global network of data, was the creation of an open community we call simply LD4. The vision of the community is quite simple:

- The world enriched with library data; libraries enriched with the world's data.

The home page for this new community can be found at LD4.io and there are instructions there as to how to join. In addition to an annual conference, LD4 currently supports a number of Linked Data Affinity Groups in areas such as Discovery, Ethics, Non-Latin Scripts, Rare Materials, Serials, and Wikidata.

By working with the open community to incorporate our tools and technology into the global concerns of linked data in libraries, we hope to help libraries take their place in the global community that Dataism implies.

3.2.3 Discovery of Data

For Harari's vision of Dataism to be realized, we need three elements: Access to Data, Linking of Data, and Discovery of Data. We have now seen how access to data is being advanced through such mandates as the Nelson Memo, how linking of data is being accomplished by the transition to linked data, but discovery has always seemed the most elusive. Perhaps this is where we need the most dramatic break with the past.

Discovery is firmly rooted in our cataloging descriptions and descriptive bibliography. At first, catalogs were simply lists of what our collections held. As time moved on, and collections grew larger, we began describing a resource in greater and greater detail. Eventually, each description found a home on a catalog card and these cards were collected in large card catalogs. Each card was a surrogate of the resource itself with a link back to the item on the shelf. Standards such as the Paris Principles or Resource Description and Access help us to create similar descriptions internationally, greatly aiding in the reuse of metadata. But two great trends are challenging our current approach to discovery. First, the sheer volume of items to be discovered, and second, the desire to discover data within a resource, not just metadata about a resource.

Let's look first at a discovery solution that is attempting to solve the problem of volume: the International Image Interoperability Framework, or IIIF.

3.2.3.1 Discovery of Data - IIIF

IIIF is defined as "set of open standards for delivering high-quality, attributed digital objects online at scale". And I think the relative phrase here is "scale". There are literally billions of images available to be discovered and more being added every day. To approach this issue, IIIF has set up a Technical Specification Group on discovery. According to their charter:

> Interoperable resources are only useful if they can be found. This has been well-recognized since the early days of the IIIF community, and more recently it has become clear that a concerted effort to standardize patterns that will facilitate discovery, harvesting and synchronization, indexing, and importing of IIIF resources is required.

When I think about the future of discovery, the challenges of metadata structure and capture, of traditional and alternative approaches to metadata creation, I often think about some images taken from social media in 2018, which are only a few of the millions of images that appeared on the Web that day. They capture a meeting between the then President of the United States, Donald Trump, and European leaders, including Angela Merkel. What would a cataloger choose to capture about these images, how do we choose to provide access, from whose perspective? Can there ever be a neutral presentation of these images? Do I highlight their political nature? Their parodies of artistic styles and periods?

Does the true benefit to scholars, how ***they*** wish to discover these images, derive more from how people are making use of them rather than from a privileged metadata surrogate I may create? The possibilities for metadata creation are almost infinite. We must look to the world of AI and machine learning for assistance in discovery. I'm very hopeful that the IIIF Technical Specification Group on discovery will begin to offer us solutions.

The world of discovery and access is in a time of impatient transformation. The ability to control access and discovery through individually crafted metadata surrogates is breaking down by the sheer onslaught of resources to be described. Our patrons are demanding topical access not through traditional subjects that many find offensive, but rather through a variety of domain specific subject thesauri that make use of terms they find familiar and supportive. How libraries choose to respond will either give them a critical voice in the evolution of discovery within this continually morphing data sphere or marginalize them to a shrinking world encompassed by the outdated semantics of catalog cards.

3.2.3.2 Discovery of Data - Know Systemic Racism (KSR)

The issue of discovery for IIIF will be focused on the explosion of digital images and how we provide access to them. But this issue is part of a broader library imperative, Collections as Data. According to Thomas Padilla:

> To see collections as data begins with reframing all digital objects as data. Data are defined as ordered information, stored digitally, that are amenable to computation. Wax cylinders, reel to reel tape, vellum manuscripts, websites, masterworks, musical scores, social media, code, and software in digital collections are brought onto the same field of consideration. The value of such a shift can be explored in part by asking how thinking about an object as data multiplies and/or extends the questions that can be asked. For example, if the notion of a single digitized text is shifted from a surrogate of a bound paper object to consider the possibility latent in a form that is computationally processable at the level of thousands or even millions of texts, a move is made toward [the making of meaning] ... that engages affordances unique to data.

The routine digitization of all of a library's assets and the creation of discovery and exploration of this material will undoubtedly be our greatest challenge in this century. The Stanford Libraries have approached this puzzle through a number of individual projects, but perhaps the most compelling is Know Systemic Racism, or KSR.

The goal of KSR is well defined: to gather and share data that show how the interconnected systems of Racism work against the Black community in the State of California. This goal can be realized by:

- documenting discrimination,
- identifying interconnections,
- and highlighting helpers.

The effort is being spearheaded by our Racial Justice and Social Equity Librarian, Felicia Smith with technical support organized by our Digital Research Architect, Nicole Coleman.

In September of 2020, Felicia began her first ambitious project in response to the global Black Lives Matter movement, an exhibit called *Say Their Names - No More Names*. The banner for the exhibit, *Know Justice, Know Peace*, still hangs in front of our Library. KSR is an outgrowth of this first effort. It is designed to help people

> discover factual data about interconnected systems that pose threats to people of African descent in the United States that have been shaped by racist policies and practices of institutions across centuries. It will also help people act against these threats by knowing their rights and identifying government agencies and community groups that address systemic racism.

The project is both critical and an excellent example of how libraries must change their approach to discovery. It would be possible for the library to catalog each of the individual resources collected for KSR and create detailed metadata surrogates for them, but the power of the collection lies in the interconnection of the data contained within the resources. As Thomas Padilla states, the value of such a shift can be explored in part by asking how thinking about an object as data multiplies and/or extends the questions that can be asked. In discussing the concept behind KSR, Nicole explains that the:

> Graph data structure reveals the relationships (links) between nodes. This linking structure makes it possible to connect different sources and resources that would otherwise be held apart within the library. Stanford Libraries is collecting California legislation records and California law enforcement agency policy manuals. What effect do these laws and policies have on our society? On the lives of individuals? The knowledge graph helps reveal those connections.
>
> "... when laws are written in race-neutral language, they only reify the existing racialized power structure that devalues Black life".

In the following, brief video, we explore the types of questions that can be asked, the connections that can be made, and interconnections that metadata could never expose.

4. Conclusion

As I began my talk today, I reflected on one of my favorite science fiction movies, *Forbidden Planet*. The race that inhabited that planet, the Krell, suffered the fate predicted by one of two paths offered in Harari's *Homo Deus*, Techno-Humanism. As the Krell evolved, they enhanced their bodies and minds with the latest technology had to offer, eventually separating mind from body. By giving that mind, however, access to unlimited power to create whatever they dreamed by thought alone, their civilization destroyed itself overnight. They had forgotten that their minds were composed of many competing desires, some good, some evil... an accretion of organic imperatives gathered over millions of years of evolution.

Harari sees a similar path developing today. With the latest developments in medicine and science, humans are increasing their lifespan and enhancing both their minds and bodies through a combination of mind-altering substances and mechanical enhancements. This evolution, though, will be available only to the super-rich. It is unlikely that it will drive the evolution of our species as a whole.

But Harari explains an alternative, a new religion, called Dataism. Dataism shifts the focus from the individual, as has long been supported by Humanism, to the flow of information across societies as a whole. This uninhibited flow of data, and the intersubjective webs of meaning built upon them, will help us reach the stars.

And as I look at the goals of a masters degree program in Library Science such as the one here at the University of Florence, I see many correlations with Harari's vision, whether it be the provision of the essential historical, methodological, and critical tools to ensure the management of knowledge, or the technical and scientific training for the purpose of the organization and flow of data. As was said by the Association for Information Science and Technology back in 1968 - Information Science is the discipline that investigates the properties and behavior of information, the forces governing the flow of information, and the means of processing information for optimum accessibility and usability

Data and the flow of information is key to all aspects of Library Science, whether it be:

- a focus on access to data through open data and open science and the developing infrastructure needed to support it, privileging the right of information to circulate freely over the rights of humans to own data and to restrict its movement,
- or a focus on the linking of data through the Web, making use of linked data as a means for uniting all knowledge into a single graph of information,
- or a focus on the discovery of data through the exploration of that Web whether it be through the creation of metadata to support the discovery of physical resources or the exploitation of the data itself hidden in digital resources.

Our role as Information professionals in the creation of this next phase of humanity is central, and the needs are many. The rise of Dataism will certainly transform the world, but it does not mandate a single deterministic outcome. The evolution and the laws of this new world will be created by people such as you.

Today, power is not only access to information, but knowing what information is safe to ignore in an overwhelming and infinite Web. Issues of accessibility and discovery will be critical in shaping our future and those who design that world will be our greatest asset.

Homo Deus
Tecno-Umanesimo vs Dataismo e il ruolo dell'informazione nella società

Lectio magistralis in Biblioteconomia

di

Philip E. Schreur

Traduzione di Denise Biagiotti

Firenze
Università degli studi di Firenze
7 marzo 2023

Casalini Libri
2023

Letture Magistrali in Biblioteconomia

Il Dipartimento di Storia, Archeologia, Geografia, Arti e Spettacolo (SAGAS), il Master di secondo livello in Archivistica, Biblioteconomia e Codicologia e il Sistema bibliotecario di ateneo dell'Università di Firenze sono lieti di presentare la serie *Letture Magistrali in Biblioteconomia*, insieme all'Associazione italiana biblioteche, alla Biblioteca nazionale centrale di Firenze, a "Biblioteche oggi", a Casalini Libri, a "JLIS.it" e alla SISBB.

Le *Letture Magistrali* si rivolgono agli studiosi e ai professionisti, impegnati principalmente i primi nell'elaborazione critica e metodologica dei tratti connotativi della disciplina, i secondi nella sperimentazione e verifica sul campo del laboratorio-biblioteca. La necessaria relazione positiva fra studiosi e professionisti crea un circolo virtuoso che favorisce l'evolversi della disciplina.

Le Letture si rivolgono inoltre agli studenti di Biblioteconomia e a quanti hanno interesse a investigare le tematiche proposte.

Mauro Guerrini
Professore di Biblioteconomia,
Catalogazione e Metadatazione
Università di Firenze

Homo Deus
Tecno-Umanesimo vs Dataismo e il ruolo dell'informazione nella società

Philip E. Schreur
Stanford University

Traduzione di Denise Biagiotti

Abstract

Sessant'anni dopo l'uscita di uno dei film di fantascienza di maggior impatto mai realizzato, *Forbidden planet* (*Il pianeta proibito*), Yuval Harari ha pubblicato la traduzione inglese del suo libro *Homo Deus*. Il futuro dell'Homo sapiens, predetto in *Homo Deus*, presenta sorprendenti parallelismi con il destino degli abitanti del pianeta proibito chiamati Krell. Spinte da un desiderio inestinguibile d'immortalità e di potere, entrambe le civiltà hanno abbracciato un movimento che Harari definisce Tecno-Umanesimo, basato sul potenziamento della mente e del corpo mediante sostanze che alterano la mente e protesi meccaniche. Harari propone un'alternativa a questo destino potenzialmente disastroso: una nuova religione chiamata Dataismo. Il Dataismo sposta l'attenzione dall'individuo, come è stato a lungo sostenuto dall'Umanesimo, al flusso d'informazioni nella società nel suo complesso. Le biblioteche e la scienza dell'informazione avranno un ruolo chiave in questo nuovo mondo mediante la promozione dell'accesso ai dati, del loro collegamento e della loro scoperta. Questo flusso incondizionato di dati e le reti intersoggettive di significato costruite su di essi, ci aiuteranno a raggiungere le stelle.

1. Introduzione - *Forbidden planet*

Il 3 marzo 1956, probabilmente uno dei film di fantascienza di maggior impatto mai prodotto, *Forbidden planet*, fu presentato in anteprima a Charlotte, nella Carolina del Nord. La proiezione si svolse nell'ambito della seconda annuale Southeastern Science Fiction Conference, in una piccola sala di proiezione della Metro-Goldwyn-Mayer (MGM) che ospitava circa quaranta persone.

Forbidden planet è stato un film con molte novità: è stato il primo a includere un viaggio più veloce della luce, a essere ambientato interamente su un pianeta diverso dalla Terra, a rappresentare un robot con una personalità, ovvero Robbie il robot, ed è stato il primo film, di qualsiasi tipologia, a includere una colonna sonora interamente generata con mezzi elettronici.

La trama di *Forbidden Planet* risulta ingannevolmente semplice. Secondo l'Internet Movie Database (IMDb), l'equipaggio di un'astronave del XXIII secolo viene inviato a indagare sull'improvvisa interruzione delle comunicazioni da parte di un avamposto, collocato su un pianeta lontano dalla Terra che gira intorno a una stella chiamata Altair IV. Una volta arrivato, il capitano scopre solo due sopravvissuti umani, un potente robot e i resti di una civiltà perduta. Poco dopo l'atterraggio dell'astronave, si scopre che tutti i coloni originari, eccetto due, sono stati orribilmente uccisi e l'equipaggio dell'astronave inizia presto a morire nello stesso modo.

Uno dei due coloni superstiti, il dottor Morbius, ha scoperto i resti di una civiltà estinta chiamata Krell. Apparentemente, all'apice del loro sviluppo scientifico, ovvero la perfetta fusione dell'organico con la macchina, la loro intera civiltà si è sgretolata in una notte di sangue. E ciò che è stato responsabile dell'estinzione dei Krell si è nuovamente liberato sul pianeta, riducendo inesorabilmente in pezzi ogni forma di vita umana. La squadra di soccorso, con tutti i suoi potenti strumenti e la sua tecnologia, non riesce a fermarlo.

2. Homo Deus

Sessant'anni dopo la prima di *Forbidden planet*, Yuval Harari, professore alla Hebrew University di Gerusalemme, ha pubblicato la versione inglese del suo libro *Homo Deus*. Il libro, il cui sottotitolo è "breve storia del domani", potrebbe ugualmente raccontare l'ascesa e la caduta dei Krell. Il parallelismo tra i Krell e l'Homo sapiens è impressionante. I Krell avevano concentrato tutta la loro attenzione sullo sviluppo scientifico. Mentre la loro civiltà si diffondeva sull'intero pianeta, entrambi cercavano una fonte di potere illimitato e la capacità di fondersi con la macchina, abbandonando infine la materia organica. Il loro successo finale li ha condotti alla rovina. Harari vede la civiltà umana avvicinarsi rapidamente a questa stessa congiuntura.

Aziende come Calico, una filiale di Google, combattono aggressivamente l'invecchiamento e alla fine potrebbero progettare l'immortalità. Gli arti protesici stanno diventando potenti sostituti di braccia e gambe mancanti. Meta Corporation, la società madre di Facebook, ha creato un'intelligenza artificiale (AI) chiamata CICERO, in grado di partecipare a un gioco di strategia, detto Diplomacy, e di negoziare altrettanto bene quanto i giocatori umani. Un oggetto chiamato stentrode, la combinazione di uno stent e di un elettrodo, se collocato nel cervello, potrebbe collegare direttamente il pensiero umano e la macchina. L'Homo sapiens sembra insistere nel perseguire il sentiero dei Krell. Nel suo libro Harari illumina chiaramente la strada che stiamo percorrendo e offre due possibili futuri. Cosa ci ha portato a questo punto e qual è il nostro ruolo come professionisti dell'informazione nel negoziare il nostro futuro?

2.1 Le prime fasi della storia dell'uomo

Secondo Harari, nel corso della storia dell'uomo siamo stati spinti da tre forze implacabili: la carestia, la peste e la guerra. E nel corso della nostra evoluzione da cacciatori-raccoglitori a società agricola fino alla cultura globale di oggi, queste tre forze hanno guidato inesorabilmente la nostra storia.

La carestia è stata un flagello sempre presente. La nostra transizione da cacciatori-raccoglitori a società agricola ha offerto maggiore predicibilità per le nostre fonti di cibo, ma l'esplosione demografica e l'imprevedibilità del clima hanno sempre messo a dura prova le nostre risorse. Allo stesso modo, l'incremento della popolazione e la creazione di una cultura veramente globale hanno reso molto più probabili pandemie come l'influenza spagnola, l'HIV o il COVID. L'aggressione dell'uomo ad altri suoi simili, messa in evidenza dall'attuale guerra in Ucraina, appare incessante. Nonostante tutto questo, Harari dimostra che ci siamo lasciati alle spalle questa parte della nostra storia. Statisticamente questi eventi stanno avendo un effetto molto minore sulla popolazione rispetto al passato.

2.2 Che cos'è un Essere Umano?

Prima di volgere lo sguardo dal passato al futuro, Harari riflette su due elementi chiave che dobbiamo esplorare: il primo è: che cosa significa essere umani? E quali sono gli elementi che hanno permesso all'Homo sapiens e ai Krell di dominare i loro rispettivi mondi?

Nel corso del tempo ci sono state molte teorie su ciò che ci rende umani: alcune l'hanno attribuito alla presenza di un'anima, altre alla nostra coscienza, altre ancora alla capacità di comunicare mediante la parola. Tutte queste teorie, però, si sono dimostrate false, come l'esclusiva capacità di comunicare oralmente, o sono state accettate solo per fede, come la presenza di un'anima. Harari preferisce, invece, concentrarsi su due caratteristiche degli esseri umani che ci rendono unici: la prima, la capacità di lavorare insieme in grandi gruppi per una causa comune e, la seconda, l'abilità di creare reti intersoggettive di significato.

È vero che molte altre creature vivono e lavorano in gruppi più ampi: le formiche formano colonie, le api alveari sotto il governo di una regina, i lupi formano branchi e delimitano territori in cui vagare, ma le somiglianze si fermano qui. I branchi di balene negli oceani del mondo non cooperano per produrre il krill come fonte di cibo, i lupi non creano "passaporti" per viaggiare in sicurezza attraverso i territori l'uno dell'altro, le api non esplorano modi alternativi

di gestire i loro alveari. La capacità unica degli esseri umani di lavorare insieme in grandi gruppi per raggiungere obiettivi comuni ci ha permesso di dominare il nostro mondo. Secondo Harari, il nostro meccanismo per raggiungere questo obiettivo è la seconda caratteristica che ci rende unici: la nostra capacità di creare reti intersoggettive di significato.

Che cosa dà valore al pezzo di carta rettangolare che chiamiamo Euro? Ha un valore intrinseco? Molto poco. È la nostra comune convinzione che abbia un valore che ci permette di usarlo come mezzo di baratto per beni tra città, stati e paesi. La Francia esiste davvero? Esiste solo in quanto i francesi credono che esista e anche altri nel mondo, allo stesso modo, accettano questa convinzione. Riponiamo così tanta fiducia in queste entità intersoggettive che diamo loro persino dei diritti legali. Entità intersoggettive come i governi possono possedere proprietà; entità commerciali come Google possono essere citate in giudizio; entità accademiche come le biblioteche possono imporre sistemi di classificazione. Queste reti intersoggettive di significato danno uno scopo alle nostre vite, permettono alle nostre civiltà di crescere e ci consentono persino di costruire astronavi per raggiungere pianeti proibiti.

E nel nostro mondo della scienza dell'informazione, cosa conferisce ai *Principi di Parigi* la loro autorità? L'*IFLA Library Reference Model* mantiene l'ordine per diritto divino? *Resource Description and Access* (RDA), nella sua forma attuale o precedente, ha un significato assoluto? È la ferma convinzione d'individui come noi, dei benefici derivanti dall'adesione a principi comuni, che conferisce loro valore.

2.3 Umanesimo

Un secondo elemento chiave dell'ascesa degli umani e dell'era dell'informazione è l'attenzione all'individuo, cioè l'Umanesimo. L'Umanesimo, semplificando, attribuisce un valore intrinseco all'umano rispetto al divino. Il movimento filosofico del XVIII secolo, noto come Illuminismo, ha rivolto questa attenzione umanistica alla scienza, alla ragione e all'intellettualismo.

L'Umanesimo stesso ha preso tre strade principali nel XX secolo. La prima e più familiare è l'Umanesimo liberale. L'Umanesimo liberale dà valore all'individuo e all'esperienza individuale. I tuoi sentimenti interiori e le opinioni sono importanti quanto quelli degli altri e i tuoi pensieri e le tue esperienze personali convalidano i tuoi sentimenti e le tue opinioni. Gli umanisti sociali danno più importanza all'esperienza umana collettiva che a quella individuale. Gli umanisti evoluzionisti direbbero che non è importante né l'esperienza individuale né quella collettiva. Essi ritengono che i pensieri e i sentimenti di alcuni individui, o razze, abbiano più peso di altri perché più evoluti.

Sembra però che l'Umanesimo e l'attenzione per l'individuo abbiano fatto il loro corso. Sebbene ci venga ancora detto che la nostra esperienza personale è importante, la superiorità economica si crea mediante la meccanizzazione dell'industria grazie all'intelligenza artificiale (AI). La guerra è sempre più guidata da mezzi meccanici come i droni rispetto alle truppe sul terreno. I social media come Twitter pongono il collettivo in contrasto con l'individuale.

Siamo giunti a un punto critico della nostra storia; il nostro percorso sta divergendo. Secondo Harari, il percorso forgiato dagli esseri umani e dall'Umanesimo ha raggiunto il suo apice e stanno emergendo due nuovi percorsi: il Tecno-Umanesimo e il Dataismo. La gestione delle informazioni e della conoscenza sarà fondamentale per entrambi.

2.4 Tecno-Umanesimo

Il Tecno-Umanesimo, in un certo senso, è la massima espressione dell'Umanesimo, poiché vede ancora l'umanità come l'apice della creazione. Gli esseri umani stanno rapidamente realizzando il loro destino presagito conquistando e sfruttando l'universo. Eppure, allo stesso tempo, abbiamo raggiunto i limiti di questo percorso. Con l'attuale svalutazione dell'individuo e l'esplosione vertiginosa della tecnologia, ci stiamo avvicinando a un punto di rottura nella nostra evoluzione. Il ritmo accelerato dello sviluppo dell'intelligenza artificiale (AI) e dell'apprendimento automatico (ML, Machine Learning) ha prodotto qualcosa che Harari descrive come la

separazione dell'intelligenza dalla coscienza. Questa separazione ci lascia una sola possibilità: forzare la mano della natura e prendere il controllo della nostra evoluzione.

Come parte di questa evoluzione, Harari vede gli esseri umani sostituire le tre paure di carestia, peste e guerra con i tre nuovi obiettivi d'immortalità, felicità e divinità. E questi tre obiettivi possono essere realizzati, proprio come i Krell di *Forbidden planet* realizzano i loro, ovvero mediante la fusione tra l'umano e la macchina. Secondo Harari,

> l'*Homo deus* manterrà alcune caratteristiche umane essenziali, ma godrà anche di abilità fisiche e mentali potenziate che gli permetteranno di tenere testa ai più sofisticati algoritmi non coscienti. Dal momento che l'intelligenza si sta separando dalla coscienza e che l'intelligenza non cosciente si sta sviluppando a gran velocità, gli esseri umani devono potenziare attivamente le loro menti se vogliono rimanere in gioco.

L'open science, gli open data e l'open technology saranno la chiave per realizzare questa rivoluzione.

Sebbene, tuttavia, essa sembri inevitabile, l'evoluzione avrà due svantaggi distinti. In primo luogo, sarà costosa. L'aggiornamento genetico e meccanico dei nostri corpi, la fusione delle nostre menti con i supercomputer, avranno costi proibitivi. Questa sarà un'evoluzione per ricchi; i poveri non ne possono far parte. Il secondo svantaggio può essere appreso da *Forbidden planet* e dalla caduta della super-razza chiamata Krell. Per i Krell, la perdita della loro forma corporea, insieme alla fusione definitiva delle loro menti con il potere e l'intelligenza illimitata della macchina, hanno permesso sia alla loro coscienza evoluta sia alla loro subcoscienza primitiva di avere uguale accesso a un potere illimitato. E in una sola notte di sangue, i Krell hanno distrutto loro stessi. Harari vede la mente umana come altrettanto complessa, una miscela di desideri contrastanti tenuti a malapena sotto controllo. Consentire a tutte le parti di questa coscienza frammentata uguale accesso al potere illimitato si rivelerebbe disastroso.

Fortunatamente, Harari vede un percorso alternativo per noi, il Dataismo.

3. Dataismo

Il Dataismo è definito come una nuova religione che non deifica né gli dei né gli uomini, ma deifica i dati, in tutte le loro forme: dati economici, dati medici, dati personali... E il Dataismo ha un solo e unico comandamento: privilegiare il diritto dell'informazione di circolare liberamente rispetto al diritto dell'uomo di possedere i dati e di limitarne la circolazione. In altre parole, non esistono dati privati. Tutti i dati devono essere aperti e condivisi. Harari presuppone che questi dati non siano necessariamente dati raccolti e resi anonimi. Ciò include i vostri dati personali. È tramite questo flusso illimitato d'informazioni che raggiungiamo il nostro nuovo obiettivo: l'internet di tutte le cose. L'individuo perde la propria posizione di centro dell'universo, come l'Umanesimo avrebbe voluto farci credere, e assume, invece, un ruolo di raccolta di punti di riferimento nella vasta rete globale dell'umanità. Qui, perciò, cominciamo a vedere il nostro ruolo di scienziati dell'informazione in questa rete globale di dati. Diventiamo operatori essenziali nel flusso delle informazioni, assicurandoci che siano aperte, comprensibili e ugualmente disponibili per tutti.

3.1 L'espressione del Dataismo nelle biblioteche

La scienza dell'informazione è diventata un punto chiave in quasi tutti i corsi di laurea in Biblioteconomia. Ciò che sorprende è che questo termine è nato quasi cinquant'anni fa. Negli Stati Uniti, la Association for Information Science and Technology (ASIS&T) è stata fondata nel 1937 come American Documentation Institute. Una delle prime definizioni di scienza dell'informazione risale al 1968:

> la scienza dell'informazione è la disciplina che studia le proprietà e il comportamento delle informazioni, le forze che ne regolano il flusso e i mezzi per elaborarle, in modo da ottimizzarne l'accessibilità e la fruibilità. Si occupa dell'insieme delle conoscenze relative all'origine, alla raccolta, all'organizzazione, all'archiviazione, al recupero, all'interpretazione, alla trasmissione e all'utilizzo delle informazioni.

Una definizione del 2009 mostra modifiche minime:

la scienza dell'informazione è la scienza e la pratica che si occupa della raccolta, dell'archiviazione, del recupero e dell'uso efficace delle informazioni. Si occupa delle informazioni e conoscenze registrabili e delle tecnologie e dei servizi correlati che ne facilitano la gestione e l'uso. Più specificamente, la scienza dell'informazione è un campo di pratica professionale e d'indagine scientifica che si occupa della comunicazione efficace d'informazioni e di oggetti informativi, in particolare di conoscenza registrata tra gli esseri umani nel contesto delle esigenze sociali, organizzative e individuali e dell'uso delle informazioni. Il dominio della scienza dell'informazione è la trasmissione dell'universo della conoscenza umana in forma registrata, incentrata sulla manipolazione (rappresentazione, organizzazione e recupero) delle informazioni, piuttosto che sulla loro conoscenza.

È interessante notare che entrambe le definizioni si concentrano sul flusso delle informazioni, e non sulla comprensione di ciò che le informazioni contengono.

3.2 Il Dataismo a Stanford

Vorrei dedicare la seconda parte del mio intervento a un discorso più pratico sul Dataismo nel nostro ambiente attuale. Come sapete dal mio background, ho trascorso gran parte della mia carriera presso le Stanford University Libraries. Si è rivelato un ambiente meraviglioso sia per gli ambiti più tradizionali delle biblioteche, come la catalogazione e lo sviluppo delle collezioni, sia per la sperimentazione dinamica sul futuro delle biblioteche. Considerare il modo in cui supportiamo il flusso delle informazioni all'interno dell'Università è fondamentale per la missione della biblioteca e i nostri sforzi possono essere suddivisi in tre ampie strategie: accesso, collegamento e scoperta dei dati. Nella sezione che segue, vorrei condividere alcuni esempi di come sosterremo ciascuna di queste strategie nei prossimi anni.

3.2.1 Accesso ai dati - OSTP Memo

In primo luogo, l'accesso ai dati. Nell'agosto del 2022, il White House Office of Science & Technology Policy (OSTP) ha emanato una nuova policy, denominata Nelson Memo. Il memorandum inizia facendo riferimento a una precedente policy emessa dall'ufficio nel 2013. La policy precedente, chiamata *Memorandum on increasing access to the results of federally funded research*:

> chiedeva a tutti i dipartimenti e le agenzie federali statunitensi con oltre 100 milioni di dollari di spese annuali per la ricerca e lo sviluppo di elaborare un piano per sostenere un maggiore accesso pubblico ai risultati della ricerca finanziata a livello federale, con particolare attenzione all'accesso alle pubblicazioni accademiche e ai dati digitali derivanti da tale ricerca.

Sebbene sia stata difficile da attuare, la policy ha permesso che oltre otto milioni di pubblicazioni accademiche siano diventate apertamente disponibili al pubblico con oltre tre milioni di persone che hanno letto questi articoli gratuitamente ogni giorno. La policy originaria, tuttavia, prevedeva un embargo di dodici mesi sull'accesso pubblico libero alle pubblicazioni accademiche, consentendo agli editori di sviluppare modelli di finanziamento che, infine, potessero includere la policy di open access imposta dal governo federale.

Secondo il nuovo memorandum emesso quest'anno, entro 180 giorni dalla pubblicazione del memo le agenzie federali con un budget superiore a 100 milioni di dollari, ed entro 360 giorni quelle con un budget pari o inferiore, devono produrre un piano in base al quale le loro ricerche finanziate a livello federale diventino disponibili al pubblico **senza alcun periodo di embargo** o ritardo dopo la pubblicazione. A causa del margine di manovra nell'attuazione di questi piani, ci sarà una data di attuazione variabile tra il 2024 e il 2028, ma alla fine di questo periodo, tutta la ricerca finanziata a livello federale di qualsiasi tipo da qualsiasi agenzia dovrà diventare liberamente disponibile al momento della pubblicazione.

Il memorandum dichiara di essere guidato da sei elementi principali: accesso equo, condivisione rapida, open science, fede pubblica, ritorno per i contribuenti e integrità della ricerca. Per raggiungere questi obiettivi, le pubblicazioni, quindi anche i dati che le supportano, i metadati per gli articoli e i dati che, a loro volta, li supportano, devono essere resi liberamente disponibili e devono essere utilizzati identificatori persistenti per individui, riconoscimenti, organizzazioni e risultati.

Ecco un grafico che mostra la percentuale di ricerca sostenuta dal governo federale presso i principali istituti di ricerca degli Stati Uniti. Da questo grafico si può notare che Stanford subirà un forte impatto da questo memorandum. Il grafico registra le pubblicazioni totali rispetto a quelle finanziate a livello federale nel periodo 2017-2021. La Stanford University si trova in cima alla curva con 75.672 pubblicazioni totali, di cui 32.493 hanno ricevuto un sostegno federale. In altre parole, il 43% delle pubblicazioni di Stanford riceve finanziamenti federali.

Il modo in cui le Stanford Libraries risponderanno a questo nuovo mandato sarà fondamentale per la sua attuazione. Riteniamo che questa nuova politica influisca sulle Stanford Libraries in quattro aree principali: la nostra Open Access Policy, il nostro Digital Repository, i nostri Research Data Services e i nostri Identifiers Services.

In primo luogo, la nostra Open Access Policy prevede che le biblioteche aiutino i ricercatori della facoltà a comprendere le nuove policy e a orientarsi tra i complessi requisiti degli editori. Abbiamo recentemente assunto un nuovo direttore per l'Office of Scholarly Communication e l'implementazione del Nelson Memo sarà la sua priorità principale. Presumiamo che ci saranno:

- maggiore urgenza nella definizione di un approccio coerente alle licenze aperte e al sostegno di Stanford per le spese di pubblicazione dei ricercatori;
- maggiore necessità di tracciare e segnalare le pubblicazioni e di inserirle nelle pagine web della facoltà;
- e un'accresciuta necessità di studiare modelli di costo per i servizi dei dati della ricerca e del repository.

Questi tre punti focali guideranno gran parte dei suoi sforzi iniziali di sensibilizzazione.

In secondo luogo, il nostro Digital Repository sarà potenziato per supportare i requisiti del memorandum, fornendo ai ricercatori:

- identificatori unici e persistenti;
- metadati;
- libero accesso ai contenuti della Stanford University e non solo;
- autenticazione durante l'accesso;
- sostenibilità a lungo termine della ricerca e dei dati che la supportano;
- riutilizzo dei contenuti;
- sicurezza;
- un formato comune;
- e una politica di conservazione dei dati.

In terzo luogo, la nostra nuova divisione bibliotecaria, Research Data Services, fornirà:

- pianificazione della gestione dei dati;
- servizi di consulenza per la curatela dei dati di ricerca;
- servizi di descrizione dei metadati e dei dati;
- e supporto per l'accesso e la manipolazione dei dati.

Infine, i nostri Identifiers Services dovranno essere migliorati. Forniamo già gli identificatori Open Researcher and Contributor Identifiers (ORCID) e i Digital Object Identifiers (DOI), ma dovremo estendere questi servizi agli identificatori di riconoscimenti, organizzazioni e risultati.

Sebbene il Nelson Memo non imponga l'open access e il flusso di dati per tutti i dati, garantirà l'accesso aperto alla maggior parte dei dati della ricerca prodotti a Stanford e negli Stati Uniti, e i servizi sviluppati dalle Stanford Libraries per fornire questo accesso potrebbero diventare la spina dorsale di una futura rete di tutti i dati.

3.2.2 Collegamento dei dati - Linked Data for Production (LD4P)

Il secondo grande punto nell'approccio della Biblioteca al Dataismo è il collegamento di questo nuovo mondo globale di dati a cui avremo accesso, e secondo me, che ha sempre implicato l'uso di linked data. Da anni perseguo la transizione dei dati tradizionalmente gestiti dalle biblioteche e raccolti in MARC (Machine Readable Cataloging) verso i linked open data. Questo interesse è culminato in una serie di proposte di sovvenzione alla Mellon Foundation denominate Linked Data for Production, o LD4P. Ma LD4P ha radici più profonde a Stanford.

Tutto è iniziato nel 2011, quando le Stanford University Libraries e il Council on Libraries and Information Resources (CLIR) hanno organizzato una conferenza di una settimana sulle prospettive di un prototipo su larga scala, multinazionale e multi-istituzionale, di un ambiente di linked data incentrato sulle biblioteche e sulle risorse a loro più utili. Fin dall'inizio, il nostro interesse si è rivolto a questo ecosistema aperto e globale. Sebbene la conoscenza creata da un'università faccia progredire la nostra comprensione del mondo, questa è solo una parte della rete globale di conoscenza di cui parla Harari e trae la sua importanza finale dalla sua collocazione in quel web aperto.

Uno dei grandi risultati di questa prima conferenza è stato un riassunto conciso dei vantaggi dei linked data:

- in primo luogo, l'approccio linked data offre vantaggi significativi rispetto alle pratiche attuali per la creazione e la fornitura di dati bibliotecari, fornendo al contempo un'estensione naturale ai modelli di condivisione collaborativa storicamente impiegati dalle biblioteche;
- in secondo luogo, i linked data, e in particolare i linked open data, sono **condivisibili**, **estensibili** e **riutilizzabili**. Supportano funzionalità multilingue per i dati e i servizi agli utenti, come l'etichettatura di concetti identificati da URI agnostici rispetto alla lingua.

Subito dopo la conferenza, ho rivolto la mia attenzione a definire come questa transizione possa avvenire concretamente nelle biblioteche. Abbiamo iniziato con l'aiuto della comunità bibliotecaria, sia negli Stati Uniti sia all'estero, e con partner commerciali di valore come SHARE-Virtual Discovery Environment, o SHARE-VDE.

LD4P ha avuto finora tre fasi. La prima è stata chiamata semplicemente *Linked Data for Production* ed è durata dal 2016 al 2018. Questa prima fase si è concentrata sulla valutazione degli strumenti e dell'architettura necessari per passare a un vero sistema di produzione.

La fase successiva, denominata *Pathway to Implementation*, si è svolta nel periodo 2018-2020. La seconda fase si è concentrata sul completamento dei Minimal Viable Products (MVPs) per gli strumenti che stavamo sviluppando e sul chiarire il percorso verso la produzione. Inoltre, ci ha permesso di espandere il progetto a una coorte di biblioteche provenienti dai membri del Program for Cooperative Cataloging (PCC).

La fase attuale di LD4P si chiama *Closing the Loop* (2020-2022). Si concentra sul ciclo di vita dei dati della biblioteca, dalle acquisizioni alla ricerca, per garantire che tutti gli aspetti siano supportati in un ambiente coeso.

Tutte le sovvenzioni si sono concentrate sulla transizione dei flussi di lavoro dei servizi tecnici di base dall'attuale infrastruttura basata sul MARC ai linked open data e al web. Il nostro obiettivo, quindi, era trasformare i dati e l'infrastruttura di una biblioteca in qualcosa che fosse semanticamente articolato, ma anche in grado di collegarsi agli altri dati liberamente aperti che appaiono sul web. LD4P è cresciuto enormemente negli ultimi anni e ha avuto molti partner, sia a livello nazionale sia internazionale, ma la nostra attenzione è sempre rimasta sugli open data e sugli strumenti per l'open access.

Il punto più critico da sviluppare è stato la creazione di un ambiente che permettesse ai catalogatori di catalogare direttamente in linked data. Per noi era essenziale che questo strumento fosse disponibile in cloud, poiché molte biblioteche non hanno il personale tecnico

per configurare l'ambiente a livello locale. Sebbene sia ancora in fase di perfezionamento, l'editor è pronto per l'uso ed è accessibile a chiunque in tutto il mondo per la sperimentazione su sinopia.io.

Da quest'anno Stanford ha iniziato a catalogare la produzione utilizzando BIBFRAME, un'ontologia di linked data, e Sinopia, concentrandosi inizialmente su opere letterarie e conferenze. Nel corso della primavera, stabiliremo l'intero flusso di lavoro e invieremo queste descrizioni in BIBFRAME a OCLC e SHARE-VDE, in modo che possano sviluppare i mezzi per assorbire la nostra produzione catalografica in BIBFRAME.

Questioning Authority (QA) è il nostro strumento open-source progettato per la ricerca e l'inserimento di entità nell'ambiente Sinopia. Lo sviluppo di QA è affidato a team della Cornell e della School of Library and Information Science dell'Università dell'Iowa. I dati sulle entità provenienti da varie autorevoli open source sono standardizzati e memorizzati nella cache, cosicché il catalogatore possa interrogare, esplorare e importare i dati in modo standardizzato.

Il risultato finale di Linked Data for Production, che ha contribuito a realizzare la visione di Harari di una vera e propria rete globale di dati, è stato la creazione di una community aperta chiamata semplicemente LD4. La visione della community è molto semplice:

- il mondo arricchito dai dati delle biblioteche; le biblioteche arricchite dai dati del mondo.

La home page di questa nuova community si trova all'indirizzo LD4.io, dove sono riportate le istruzioni per l'adesione. Oltre a una conferenza annuale, LD4 supporta attualmente una serie di Linked Data Affinity Groups in aree quali Discovery, Ethics, Non-Latin Scripts, Rare Materials, Serials e Wikidata.

Lavorando con la community aperta per incorporare i nostri strumenti e la nostra tecnologia nel contesto globale dei linked data nelle biblioteche, speriamo di aiutare queste ultime a prendere il loro posto nella comunità globale che il Dataismo implica.

3.2.3 Scoperta dei dati

Affinché la visione di Harari del Dataismo si realizzi, abbiamo bisogno di tre elementi: accesso, collegamento e scoperta dei dati. Abbiamo visto come l'accesso ai dati sia stato favorito da mandati come il Nelson Memo, come il collegamento dei dati sia stato realizzato con il passaggio ai linked data, ma la scoperta è sempre sembrata la cosa più sfuggente. Forse è questo il punto in cui abbiamo bisogno di una rottura più drastica con il passato.

La scoperta è saldamente radicata nelle nostre descrizioni catalografiche e nella bibliografia descrittiva. All'inizio i cataloghi erano semplici elenchi di ciò che le nostre collezioni possedevano. Con il passare del tempo e l'ampliarsi delle collezioni, abbiamo iniziato a descrivere una risorsa in modo sempre più dettagliato. Alla fine, ogni descrizione è stata inserita in una scheda di catalogo e queste schede sono state raccolte in grandi cataloghi a schede. Ogni scheda era un surrogato della risorsa stessa, con un collegamento all'item sullo scaffale. Standard come i *Principi di Parigi* o *Resource Description and Access* ci aiutano a creare descrizioni simili a livello internazionale, favorendo notevolmente il riutilizzo dei metadati. Due grandi tendenze, tuttavia, stanno mettendo in discussione il nostro attuale approccio alla scoperta. In primo luogo, l'enorme volume di oggetti da scoprire e, in secondo luogo, il desiderio di scoprire i dati all'interno di una risorsa, non solo i metadati relativi a essa.

Vediamo, innanzitutto, una soluzione per la scoperta che cerca di risolvere il problema del volume: l'International Image Interoperability Framework, o IIIF.

3.2.3.1 Scoperta dei dati – IIIF

IIIF è definito come "un insieme di standard aperti per la fornitura di oggetti digitali di alta qualità e attribuiti online in scala". E credo che l'espressione rilevante sia "in scala". Ci sono, letteralmente, miliardi d'immagini da scoprire e ogni giorno se ne aggiungono altre. Per affrontare questo problema, la IIIF ha istituito un Technical Specification Group per la scoperta. Secondo il loro statuto:

> le risorse interoperabili sono utili solo se possono essere trovate. Questo aspetto è stato ben riconosciuto fin dagli albori della comunità IIIF, e più recentemente è diventato chiaro che è necessario uno sforzo preordinato per standardizzare i modelli che faciliteranno la scoperta, la raccolta e la sincronizzazione, l'indicizzazione e l'importazione delle risorse IIIF.

Quando penso al futuro della scoperta, alle sfide della struttura e della cattura dei metadati, degli approcci tradizionali e alternativi alla creazione dei metadati o metadatazione, penso spesso ad alcune immagini tratte dai social media nel 2018, che sono solo una piccola parte delle milioni di immagini apparse sul web quel giorno. Esse immortalano un incontro tra l'allora Presidente degli Stati Uniti, Donald Trump, e i leader europei, tra cui Angela Merkel. Cosa sceglierebbe un catalogatore per catturare queste immagini, come scegliere di fornire l'accesso, da quale prospettiva? Può mai esistere una presentazione neutrale di queste immagini? Evidenzio la loro natura politica? Le loro parodie di stili e periodi artistici?

Il vero beneficio per gli studiosi, come costoro desiderano scoprire queste immagini, deriva più dal modo in cui le persone ne fanno uso piuttosto che da un surrogato privilegiato di metadati che posso creare? Le possibilità di creazione di metadati sono quasi infinite. Dobbiamo guardare al mondo dell'AI e del Machine Learning per avere assistenza nella scoperta. Spero che l'IIF Technical Specification Group sulla scoperta inizi a offrire delle soluzioni.

Il mondo della scoperta e dell'accesso vive un momento d'impaziente trasformazione. La capacità di controllare l'accesso e la scoperta mediante surrogati di metadati realizzati individualmente sta venendo meno a causa della mole di risorse da descrivere. I nostri clienti chiedono un accesso semantico non tramite i soggetti tradizionali, che molti trovano sconvenienti, bensì mediante una varietà di thesauri di soggetti specifici per il dominio, che fanno uso di termini che riconoscono familiari e di supporto. Il modo in cui le biblioteche decideranno di rispondere darà loro una voce critica nell'evoluzione della scoperta all'interno di questa sfera di dati in continua evoluzione, oppure le emarginerà in un mondo ristretto, racchiuso dalla semantica obsoleta delle schede di catalogo.

3.2.3.2 Scoperta dei dati - Know Systemic Racism (KSR)

Il tema della scoperta per IIIF sarà incentrato sull'esplosione delle immagini digitali e su come fornirne l'accesso. Questo tema fa parte di un imperativo biblioteconomico più ampio: le collezioni come dati. Secondo Thomas Padilla:

> la visione delle collezioni come dati inizia con la riformulazione di tutti gli oggetti digitali come dati. I dati sono definiti come informazioni ordinate, memorizzate digitalmente, che possono essere sottoposte a calcolo. I cilindri di cera, i nastri a bobina, i manoscritti su pergamena, i siti web, i capolavori, le partiture musicali, i social media, il codice e il software delle collezioni digitali sono considerati alla stessa maniera. Il valore di questo cambiamento può essere esplorato in parte chiedendosi in che modo pensare a un oggetto come a un dato moltiplichi o estenda le domande che possono essere poste. Per esempio, se il concetto di un singolo testo digitalizzato si modifica da un surrogato di un oggetto cartaceo rilegato per considerare la possibilità latente in una forma elaborabile computazionalmente a livello di migliaia o addirittura milioni di testi, si compie un passo verso [la creazione di significato] ... che contempla le possibilità uniche dei dati.

La digitalizzazione di routine di tutto il patrimonio di una biblioteca e la creazione di un sistema di scoperta ed esplorazione di questo materiale saranno senza dubbio la nostra più grande sfida in questo secolo. Le Stanford Libraries hanno affrontato questo problema mediante una serie di progetti individuali, ma forse il più interessante è Know Systemic Racism, o KSR.

L'obiettivo di KSR è ben definito: raccogliere e condividere dati che mostrino come i sistemi interconnessi di razzismo operino contro la comunità nera nello Stato della California. L'obiettivo può essere raggiunto:

- documentando la discriminazione;
- identificando le interconnessioni;
- ed evidenziando gli aiutanti.

Lo sforzo è guidato dalla nostra Racial Justice and Social Equity Librarian, Felicia Smith, con il supporto tecnico garantito dal nostro Digital Research Architect, Nicole Coleman.

Nel settembre 2020, Felicia ha iniziato il suo primo ambizioso progetto in risposta al movimento globale Black Lives Matter, una mostra intitolata *Say Their Names - No More Names*. Lo striscione della mostra, Know Justice, Know Peace, è ancora appeso davanti alla nostra biblioteca. KSR è il risultato di questo primo sforzo. È stato progettato per aiutare le persone a

> scoprire dati concreti sui sistemi interconnessi, che rappresentano una minaccia per le persone di origine africana negli Stati Uniti, modellati dalle politiche e dalle pratiche razziste delle istituzioni nel corso dei secoli. Aiuterà, inoltre, le persone ad agire contro queste minacce, conoscendo i propri diritti e identificando le agenzie governative e i gruppi comunitari che affrontano il razzismo sistemico.

Il progetto è al tempo stesso cruciale e un ottimo esempio di come le biblioteche debbano cambiare il loro approccio alla scoperta. Sarebbe possibile per la biblioteca catalogare ogni singola risorsa raccolta per KSR e creare surrogati dettagliati di metadati per esse, ma il potere della collezione risiede nell'interconnessione dei dati contenuti nelle risorse. Come afferma Thomas Padilla, il valore di questo cambiamento può essere esplorato in parte chiedendosi in che modo pensare a un oggetto come a un dato moltiplichi o estenda le domande che possono essere poste. Nel discutere il concetto alla base del KSR, Nicole spiega che:

> la struttura a grafo dei dati rivela le relazioni tra i nodi. Questa struttura di collegamento consente di collegare fonti e risorse diverse che altrimenti sarebbero tenute separate all'interno della biblioteca. Le Stanford Libraries stanno raccogliendo i record legislativi e i manuali di politica delle forze dell'ordine della California. Che conseguenze hanno queste leggi e politiche sulla nostra società? Sulla vita degli individui? Il grafico della conoscenza aiuta a rivelare queste connessioni.
>
> "quando le leggi sono scritte con un linguaggio neutro rispetto alla razza, non fanno altro che reificare la struttura di potere razziale esistente che svaluta la vita dei neri".

Nel breve video che segue, esploriamo i tipi di domande che si possono porre, le connessioni che si possono creare e le interconnessioni che i metadati non potrebbero mai rivelare.

4. Conclusione

Quando ho iniziato il mio intervento di oggi, ho riflettuto su uno dei miei film di fantascienza preferiti, *Forbidden planet*. La razza che abitava quel pianeta, i Krell, subì il destino previsto da uno dei due percorsi proposti nell'*Homo Deus* di Harari, il Tecno-Umanesimo. Mentre i Krell si evolvevano, miglioravano i loro corpi e le loro menti con le ultime tecnologie disponibili, arrivando a separare la mente dal corpo. Tuttavia, dando a quella mente l'accesso al potere illimitato di creare qualsiasi cosa sognassero con il solo pensiero, la loro civiltà si è distrutta da un giorno all'altro. Essi avevano dimenticato che le loro menti erano composte da molti desideri contrastanti, alcuni buoni, altri malvagi... un insieme d'imperativi biologici raccolti in milioni di anni di evoluzione.

Harari vede un percorso simile svilupparsi oggi. Grazie alle ultime conquiste della medicina e della scienza, gli esseri umani stanno aumentando la durata della loro vita e stanno migliorando sia la loro mente sia il loro corpo mediante una combinazione di sostanze che alterano la mente e apportano potenziamenti meccanici. Questa evoluzione, però, sarà disponibile solo per i super-ricchi. È improbabile che possa guidare l'evoluzione della nostra specie nel suo complesso.

Harari propone un'alternativa, una nuova religione, chiamata Dataismo. Il Dataismo sposta l'attenzione dall'individuo, come è stato a lungo sostenuto dall'Umanesimo, al flusso delle informazioni che attraversa le società nel loro complesso. Questo flusso incondizionato di dati e le reti intersoggettive di significato costruite su di esso, ci aiuteranno a raggiungere le stelle.

E se guardo agli obiettivi di un Corso di laurea magistrale in Biblioteconomia come quello dell'Università di Firenze, vedo molte correlazioni con la visione di Harari, sia che si tratti di fornire gli strumenti storici, metodologici e critici essenziali per garantire la

gestione della conoscenza, sia che si tratti della formazione tecnica e scientifica ai fini dell'organizzazione e del flusso di dati. Come è stato sostenuto dalla Association for Information Science and Technology nel 1968, la scienza dell'informazione è la disciplina che studia le proprietà e il comportamento delle informazioni, le forze che ne governano il flusso e i mezzi per elaborarle, in modo da ottimizzarne l'accessibilità e la fruibilità.

I dati e il flusso delle informazioni sono fondamentali per tutti gli aspetti della biblioteconomia, che sia:

- l'attenzione all'accesso ai dati mediante open data e open science e lo sviluppo dell'infrastruttura necessaria a sostenerli, privilegiando il diritto delle informazioni a circolare liberamente rispetto al diritto dell'uomo a possedere i dati e a limitarne la circolazione,
- o l'attenzione sul collegamento dei dati tramite il web, facendo uso dei linked data come mezzo per unire tutta la conoscenza in un unico grafo d'informazioni,
- oppure l'attenzione sulla scoperta dei dati tramite l'esplorazione del web, sia mediante la creazione di metadati a supporto della scoperta di risorse fisiche, sia mediante lo sfruttamento dei dati stessi nascosti nelle risorse digitali.

Il nostro ruolo di professionisti dell'informazione nella creazione di questa prossima fase dell'umanità è centrale e le esigenze sono molteplici. L'ascesa del Dataismo trasformerà certamente il mondo, ma non imporrà un unico risultato deterministico. L'evoluzione e le leggi di questo nuovo mondo saranno create da persone come voi.

Oggi, il potere non è solo l'accesso alle informazioni, ma anche sapere quali informazioni è meglio ignorare in un web travolgente e infinito. Le questioni dell'accessibilità e della scoperta saranno fondamentali per plasmare il nostro futuro e coloro che progetteranno quel mondo saranno la nostra più grande risorsa.