Informatio
27(1), 2022, pp. 10-68 ISSN: 2301-1378

DOI: 10.35643/Info.27.1.12

Artículo original

Data and knowledge infrastructures for scientific research on humans and health in Uruguay: an experience report on a critical constructive design process

Infraestructuras de datos y conocimiento para la investigación científica sobre humanos y salud en Uruguay: relato de experiencia sobre un proceso de diseño constructivo crítico

Infraestruturas de dados e conhecimento para pesquisas científicas sobre humanos e saúde no Uruguai: um relato de experiência sobre um processo de projeto construtivo crítico

Horacio Botti[1] **ORCID: 0000-0001-9880-0767**

Hugo Naya[2,3] **ORCID: 0000-0001-6982-4399**

Silvia Méndez[4] **ORCID: 0000-0002-7592-3016**

Camila Simoes[2] **ORCID: 0000-0003-1664-1606**

Valentina Martufi[5] **ORCID: 0000-0001-6120-0629**

[1] Laboratorio de Biofísica Integrativa, Departamento de Biofísica, Facultad de Medicina, Universidad de la República (Uruguay). Correo electrónico: hbotti@fmed.edu.uy

[2] Unidad de Bioinformática, Institut Pasteur de Montevideo (Uruguay). Correos electrónicos: naya@pasteur.edu.uy, csimoes@cup.edu.uy

[3] Genética Cuantitativa en Facultad de Agronomía - Universidad de la República (Uruguay).

[4] Unidad Académica de Bioética, Facultad de Medicina, Universidad de la República (Uruguay). Correo electrónico: simendez20@gmail.com

[5] Center for Data and Knowledge Integration for Health, Oswaldo Cruz Foundation. Parque Tecnológico da Edf. Tecnocentro, R. Mundo, nº 121 - sala 315 - Trobogy, Salvador - BA, 41745-715, Brazil. Correo electrónico: valentina.martufi@gmail.com

Abstract

Data-centrism, precision health, data-science and open science are factors transforming health research, healthcare and health policy. The COVID-19 pandemic interacted with the former and revealed some deficiencies of the latter. This has in turn stimulated disciplinary clashes, interdisciplinary elaborations and political conflicts, concentrated around the techno-scientific, political and societal changes needed to cope with the complex health problems that were and will be challenging our societies (inequity, poverty, pandemics, climate change,

CC BY

reemerging diseases, etc). Herein, as a continuation of an interdisciplinary dialogue initiated in 2018 with the "Data and Information Sciences Applied to Human Health" project (CIDASH for short), we put forward the idea of planning a data-centric, open, health research infrastructure, interoperable with the integral national health system and with national health registries. We aimed to identify and analyze from interdisciplinary and trans-disciplinary perspectives and based on our experiences in Uruguay the opportunities, concerns and challenges regarding the implementation of an open health research infrastructure (OSHRI). Our specific objectives were: i) to provide a brief account of our experience in human health research in Uruguay; ii) to reflect on the perceived opportunities, challenges, issues and potential solutions in human health research from the scientific, methodological, technical, ethical and regulatory point of view; iii) to elaborate on and propose structuring pillars for future open data infrastructures for human health research in Uruguay. The methodology we followed to report our experience was qualitative, interpretative and deliberative. Results include a list of problems and design pillars identified in relation of the creation of an OSHRI. To discuss these results we took the Centre for the Integration of Health Data and Knowledge (CIDACS) in Salvador de Bahía (Brazil) as an example of an installed and running open science health research infrastructure that warrants access to population-size data with proper ethical and scientific standards. After deliberation on the above-mentioned experiences we found that this report could bring us a step closer in the way of engaging with participatory research and change from within the science community.

Keywords: Open Science, Research Infrastructures, Health Research, Bioethics, Biopolitics, Causality.

Resumen

El centrismo de datos, la salud de precisión, la ciencia de datos y la ciencia abierta son factores que transforman la investigación en salud, la atención médica y las políticas de salud. La pandemia de COVID-19 interactuó con los primeros y reveló algunas deficiencias de los segundos. Esto, a su vez, ha estimulado choques disciplinarios, elaboraciones interdisciplinarias y conflictos políticos, concentrados en torno a los cambios tecnocientíficos, políticos y sociales necesarios para hacer frente a los complejos problemas de salud que desafiaron y desafiarán a nuestras sociedades (desigualdad, pobreza, pandemias, cambio climático, enfermedades reemergentes, etc.). Aquí, como continuación de un diálogo interdisciplinario iniciado en 2018 con el proyecto "Ciencias de la Información y los Datos Aplicados a la Salud Humana" (CIDASH por sus siglas en inglés), planteamos la idea de planificar una infraestructura de investigación en salud, abierta, centrada en datos, interoperable con el sistema nacional integral de salud y con los registros nacionales de salud. Nuestro objetivo fue identificar y analizar desde perspectivas interdisciplinarias y transdisciplinarias y con base en nuestras experiencias en Uruguay, las oportunidades, preocupaciones y desafíos relacionados con la implementación de una infraestructura abierta de investigación en salud (OSHRI). Nuestros objetivos específicos fueron: i) brindar un breve relato de nuestra experiencia en investigación en salud humana en Uruguay; ii) reflexionar sobre las oportunidades percibidas, los desafíos, los

problemas y las posibles soluciones en la investigación en salud humana desde el punto de vista científico, metodológico, técnico, ético y normativo; iii) elaborar y proponer pilares estructurantes de futuras infraestructuras de datos abiertos para la investigación en salud humana en Uruguay. La metodología que seguimos para relatar nuestra experiencia fue cualitativa, interpretativa y deliberativa. Los resultados incluyen una lista de problemas y pilares de diseño identificados en relación con la creación de un OSHRI. Para discutir estos resultados, tomamos el Centro para la Integración de Datos y Conocimientos en Salud (CIDACS) en Salvador de Bahía (Brasil) como ejemplo de una infraestructura de investigación en salud de ciencia abierta instalada y en funcionamiento que garantiza el acceso a datos de tamaño casi poblacional con estándares éticos y científicos particulares. Después de deliberar sobre las experiencias mencionadas anteriormente, descubrimos que este informe podría acercarnos un paso más en la forma de comprometernos con la investigación participativa y el cambio desde dentro de la comunidad científica.

Palabras clave: Ciencia Abierta, Infraestructuras de Investigación, Investigación en Salud, Bioética, Biopolítica, Causalidad.

Resumo

Centrismo de dados, saúde de precisão, ciência de dados e ciência aberta são fatores que transformam a pesquisa em saúde, a saúde e a política de saúde. A pandemia de COVID-19 interagiu com os primeiros e revelou algumas deficiências dos segundos. Isso, por sua vez, estimulou confrontos disciplinares, elaborações interdisciplinares e conflitos políticos, concentrados em torno das mudanças tecno-científicas, políticas e sociais necessárias para lidar com os complexos problemas de saúde que desafiaram e desafiarão nossas sociedades (desigualdade, pobreza, pandemias, mudanças climáticas, doenças reemergentes etc.). Aqui, como uma continuação de um diálogo interdisciplinar iniciado em 2018 com o projeto "Ciências de Dados e Informação Aplicadas à Saúde Humana" (CIDASH), apresentamos a ideia de planejar uma infraestrutura de pesquisa em saúde centrada em dados, aberta e interoperável com o sistema nacional de saúde integral e com os registros nacionais de saúde. Nosso objetivo foi identificar e analisar, a partir de perspectivas interdisciplinares e transdisciplinares e com base em nossas experiências no Uruguai, as oportunidades, preocupações e desafios relacionados à implementação de uma infraestrutura aberta de pesquisa em saúde. Nossos objetivos específicos foram: i) apresentar um breve relato da nossa experiência em pesquisa em saúde humana no Uruguai; ii) refletir sobre as oportunidades, desafios, questões e potenciais soluções percebidas na pesquisa em saúde humana do ponto de vista científico, metodológico, técnico, ético e regulatório; iii) elaborar e propor pilares estruturantes para futuras infraestruturas de dados abertos para pesquisa em saúde humana no Uruguai. A metodologia utilizada foi qualitativa, interpretativa e deliberativa. Os resultados incluem uma lista de problemas e pilares de design identificados em relação à criação de uma infraestrutura de dados abertos para pesquisa em saúde humana. Para discutir esses resultados, tomamos o Centro de Integração de Dados e Conhecimentos em Saúde (CIDACS) em Salvador da Bahia (Brasil) como um exemplo de infraestrutura de pesquisa em saúde de

ciência aberta instalada e em execução que provê acesso a dados populacionais garantindo o respeito de padrões éticos e científicos. Após deliberação sobre as experiências acima mencionadas, descobrimos que este relatório poderia nos aproximar um pouco mais na maneira de nos engajarmos com a pesquisa participativa e a mudança dentro da comunidade científica.

Palavras-chave: Ciência Aberta, Infraestruturas de Pesquisa, Pesquisa em Saúde, Bioética, Biopolítica, Causalidade

Fecha de recibido: 05/12/2021

Fecha de aceptado: 18/03/2022

I. Introduction

Data centrism, big data, precision X (X can be almost anything, medicine, management, etc.), data science (including autonomization of algorithms) and open science are a few processes that are currently impacting health sciences and healthcare at a global scale (Leonelli, 2016; Yu, Beam, & Kohane, 2018). These processes interact and generate philosophical, regulatory, technological, economic, political, governance and social problems. While there is an obvious need for critical evaluation of results, as well as for participation on these processes aiming at promoting social control and well balanced impacts, mounting evidence suggests that they can benefit from the creation of health research infrastructures (Hummel & Braun, 2020).

Data-centrism in scientific research brings new challenges and opportunities. One well established data-centric science model is the “consortium” in which data sharing regulation procedures emerge from consensus and confrontation between parties (Leonelli, 2016). Consensus typically concern the modalities of use of technologies that need to be shared within the consortium and then further disseminated, developed and expanded with the active participation of data producers and users over the globe (Leonelli, 2016). Governments require national statistics (an ancient form of data-centrism). National health registries are organizations which are legitimated by law to collect and prepare clinical data on specific health problems with population width repurposing it to provide unbiased epidemiological descriptions and surveillance, in order to inform policy and the public. Although their main commitment is with the national level, they collaborate with similar organizations, frequently constituting international

networks. Nevertheless, different registries within the national level are not necessarily interoperable and the stored data do not travel from one to another jurisdiction or are linked together. Neither, health registries are able to contribute their population´s personal data to international public repositories for obvious national security, economic and political reasons.

Open Science (OS) is a social movement that promises to make scientific research and the benefits derived from it larger and more accessible to all levels of society, amateur or professional (Woelfle, Olliaro, & Todd, 2011). This movement, began in the 1990s as an open access movement, and more recently it has been promoted as OS, in all continents. Above all, this idea points to an improvement in individual and collaborative research processes, in their communication and reproducibility, in order to achieve rapid production, access to data and use of new knowledge (Babini & Rovelli, 2020). In general, a social movement is conceived as a "network of informal interactions between a plurality of individuals, groups and or organizations, engaged in political or cultural conflicts, on the basis of shared collective identities" (Diani, 1992). In addition, social movements interact with interest groups, coalitions and other organizations, of particular relevance with public universities. Hence, the OS agenda and values both overlap and collide with those of Universidad de la República, which has been mandated to increase, spread and defend culture, promote and protect scientific research, to be in charge of the qualification for the exercise of scientific professions and to contribute to the study of problems of general interest and promote their public understanding (Uruguay, 1958). Therefore, OS is creating new problems, challenges and risks, particularly in the biology, sociology, technology and health domains related to research quality evaluation and promotion, intellectual creation and property rights and knowledge dissemination (Holzmeyer, 2019; Randall, 2021; UDELAR, 2020).

Studying health and healthcare problems through the re-use of population size sample data disseminated through massive communication media was a trending mode of research by ad-hoc and already constituted multidisciplinary research teams, by isolated citizens and groups performing citizen-science like activities as well as things in-between, in attempts to better respond to the COVID-19 pandemic (S. Méndez & Botti, 2021). A major problem we observed is that not all

these researchers and citizens engaged in HR during the pandemic were aware and compliant to regulations of research on human subjects, which means that a greater and perhaps different effort from the academy is needed to achieve responsible science in an open and citizen science scenario or in future national emergencies. In addition, we observed that no significant regulatory or procedural changes better responding to new contexts and bioethical issues were elaborated during COVID-19 national emergency (S. Méndez & Botti, 2021). Similar tensions and mismatches were reported to take place in all regions and some authors have proposed that changes in the research evaluation system were needed in underdeveloped countries (Aarons, 2018). In 2018 we started an interdisciplinary dialogue centered around the applications of information and data sciences in human health (CIDASH, 2019). Herein, we continue that interdisciplinary process, that was shaken by the COVID-19 pandemic, proposing to openly start the critical design a Health Research Infrastructure (HRI) in Uruguay, potentially integrating a new health research commons that could bring support for better science, including the promotion of more legitimate, ethical and secure re-use of sensitive personal, health and administrative data (including population size data samples). We wish to consider ideas, problems, etc. related to an HRI that could in principle be compatible with values that we speculate OS and Universidad de la República could share in relation to health research.

II. Objectives

The main objective of this work was to identify and analyze from interdisciplinary and transdisciplinary perspectives and based on our experiences in Uruguay the opportunities, concerns and challenges regarding the implementation of an open health research infrastructure (OSHRI). Our specific objectives are: i) to provide a brief account of our experience in human health research in Uruguay; ii) to reflect on the perceived opportunities, challenges, issues and potential solutions in human health research from the scientific, methodological, technical, ethical and regulatory point of view; iii) to elaborate on and propose structuring pillars for future open data infrastructures for human health research in Uruguay.

III. Methods

Methodology. The methodological approach herein employed is qualitative. An experience report is described by Daltro and de Faria as a postmodern scientific narrative (Daltro & de Faria, 2019). This concept is especially significant to us, because the experience report methodology departs from strict positivist (modern) approaches and can overcome some of their limitations. In agreement with Daltro and de Faria (Daltro & de Faria, 2019), we think that studying complex real systems involving humans requires accepting the possibility that researchers form observer-participator junctures and therefore, the subjective, perspectival and dynamical nature of the knowledge created through the report of scientific experiences must be acknowledged. Therefore, elaborating generalizations is out of focus in experience reports, rather contextual insights and positioning are probably among its valuable products. Communicating the reader about reasons for key choices and reflections that characterize the experience so far is highly valuable as well. Arriving at conclusions in experience reports is probably not a certain yield. In fact, Daltro and de Faria propose that conclusions should not be part of experience reports (Daltro & de Faria, 2019). In that respect we depart from their view. We think conclusions are possible and valuable if correctly communicated and understood as contextual, historical and dynamic.

Communicating and living the experience. The authors shared the stage in several activities which are briefly described in the Results section. In addition, the authors elaborated written materials or performed oral presentations to share their experiences with all authors. Relevant scientific articles and projects authored by the authors were also shared. Relevant literature was shared and discussed. Some authors proposed texts to be integrated in the document and all authors had the opportunity to comment and propose changes. Most of the authors participated in face to face exchanges using the ZOOM platform. Authors worked to arrive to consensus, nevertheless, we felt necessary to communicate among us how strong were the approvals for 42 key statements. To that end, we performed a brief self-report Likert scale-based survey (Likert, 1933). Responses were used in a new round of reflection, improving the writing and increasing consensus.

Review of the literature. The literature was reviewed using Google Scholar, PubMed and Scielo following purposive sampling (Suen, Huang, & Lee, 2014). Languages were limited to English, Spanish and Portuguese.

Definitions used throughout this report. Here are some definitions we agreed on the three entities that will be centering this discussion: health research, research infrastructures and open science. We define Health Research as encompassing: a) scientific research on the health of human beings (Human Health Research, HHR), this research requires access to participants' bodies, stories, biological samples and/or data produced by them (considering all possible ways); HHR includes population, clinical and (wet and dry) laboratory research; its regulation in Uruguay is covered to a large extent by (Uruguay, 2019), which states that research on humans is "research in any area of knowledge, involving human beings" (that of course includes health); and b) research on factors already known to affect or likely/arguably affecting human health without involving human participants. In addition to HR, scientific research on human beings can have diverse forms and objectives, part of it likely may have eventual, more or less direct connections to health. This is somehow expected since health has itself a holistic and systemic definition which we will not analyze here, but we recognize the relevance of the topic for further work.

The European Commission defines RIs as "single-sited, virtual or distributed" ... "facilities, resources, and services that are used by the research communities to conduct research and foster innovation in their fields. Where relevant, they may be used beyond research, e.g. for education or public services. They include: major scientific equipment (or sets of instruments); knowledge-based resources such as collections, archives, or scientific data; e-infrastructures, such as data and computing systems and communication networks" (Directorate-General for Communication, 2013). The word "cyberinfrastructure" was employed by Atkins and cols. from the National Science Foundation in 2003 (Atkins et al., 2003) and has overlapping meanings with RI (Anderson, 2013; Stewart et al., 2010). A Knowledge Infrastructure (KI) is also a related concept, emphasizing the value of knowledge creation and dissemination, which may be more connected to the profile of activities of universities and data-driven research domains (Borgman,

2019). We herein take the European Commission definition of RIs as sufficiently flexible to encompass the others.

Open science has been defined as "the movement to make scientific research (including publications, data, physical samples, and software) and its dissemination accessible to all levels of an inquiring society, amateur or professional" (Woelfle et al., 2011). That said, it may appear that open science is a monolithic creation that can only be adopted as a radicalism in the wide sense of the word. Differently, Fecher and Friesike have defined open science as an "umbrella term" under which several independent things coexist (Fecher & Friesike, 2014). They proposed five schools of thought in relation to open science, focusing on technology infrastructure architecture, public accessibility to knowledge generation (including but not only "citizen science"), alternative science impact measurement, democratic access to knowledge (open publications that result from public research funding) and collaborative research (including new forms of communication) (Fecher & Friesike, 2014).

IV. Results

Brief account of the experience-shaping research and development activities. The authors participated in five activities that can be considered central for the reported experience: 1) An interdisciplinary and inter-institutional research project called "Ciencias de la Información y de Datos Aplicados a la Salud Humana" ("CIDASH", translated as Information and Data Sciences Applied to Human Health); 2) an interdisciplinary and inter-institutional Uruguayan group constellation originally intended to provide a scientific response to the COVID-19 pandemic in Uruguay, where the "Grupo Uruguayo Interdisciplinario de Análisis de Datos de COVID‑19" ("GUIAD-COVID-19", translated as "Interdisciplinary Uruguayan Group for the Analysis of COVID-19 Data") took part; 3) also part of this response was a group of academics and professionals that worked advising at the level of the Uruguayan government, called "Grupo Asesor Científico Honorario" ("GACH", translated as Honorary Scientific Advisory Group); and 4) work done by several groups and institutions in Uruguay to develop diagnostics and human resources needed to respond to rare diseases, particularly in the field

of genomics, where the “Urugenomes” project had a role. The activities occurred in Uruguay. The foundational proposals for CIDASH, GUIAD-COVID-19, and Urugenomes are available upon a well-founded request for research purposes. These activities were partly initiated, coordinated, integrated and/or evaluated by us as shown in Table I. These activities had several common characteristics: multi and/or interdisciplinarity, multi/inter-institutional support, human health/human research-focused and a preoccupation about, or usage of, personal data.

The main objective of CIDASH was to develop interdisciplinary and inter-institutional activities at the national academic and health fields, which could facilitate and channel the responsible development of applications of information and data sciences in human health, anticipating potential great advantages and important risks associated with them. CIDASH (with about 40 members and 6 organizations represented -Academic Unit of the Comisión Sectorial de Investigación Científica, Faculties of Medicine, Engineering, Information and Communication and Psychology (UdelaR) and Institut Pasteur de Montevideo) allowed us to move forward with our feet on the ground (particularly at the academic level). One central activity of CIDASH was the realization of open discussions on several topics that were reflected in the titles of the activities: “Regulations and ethics in research with personal data”, “Health, Information and Data Sciences. Fundamental concepts”, “Data, medical records, National Electronic Medical Records and health information systems,” "Challenges of Data Science applied to Human Health: regional and international experiences” (where Dr. Mauricio Barreto, from Centre for the Integration of Health Data and Knowledge (CIDACS) delivered on open science and CIDACS’ experience) and “Genomic and Precision Medicine” (where HN delivered on genomics and Urugenomes project (see details below). CIDASH-I was designed as an experience-sharing and a project and experience generating atmosphere. One result of CIDASH-I was the elaboration of CIDASH-II (August, 2020) whose main objective was to contribute to the development of applications in human health of computer, information and data sciences which a) effectively integrate the consideration of ethical, human rights, biopolitical, democratic, epistemological as well as scientific, technological, human, social and economic

aspects/values, and b) need to ultimately be purposed at improving the quality of life of individuals, groups and communities. Two specific aims were then delineated: 1) to contribute to the development of human resources and interdisciplinary groups in life, health, computing, information and data sciences; and 2) to contribute to the development of intersectoral infrastructures necessary for quality inter/disciplinary research in computer, data and information sciences applied to human health. This experience report is one way of advancing in the last direction. CIDASH-II focuses also in bridging the humanities-sciences gap, which is reflected in the proposal at many levels, including its institutional support. CIDASH-II was paused because of funding difficulties and the prioritization of working on the COVID-19 pandemic. CIDASH-II will be relaunched in March 2022.

The 16th of March of 2020 an interdisciplinary and intersectoral group of researchers and medical specialists, later named GUIAD-COVID-19, was created to respond to the COVID-19 pandemic aiming at collaborating with the University Hospital, the National Integrated Health System and the national health authority (the Ministry of Public Health) with a strong participation of UdelaR. This group maintained a non-traditional citizen science-like self organizing activity but received support from the institutions represented by its members, all of them active researchers or health professionals. The collaboration with the Ministry of Public Health could not take place directly as originally planned, instead GUIAD-COVID-19 supported the creation (16th of April of 2020) and partly integrated the GACH. The GACH had the mission “to scientifically advise the Presidency of the Uruguayan Republic”... “on the road to "the new normality”” (https://www.gub.uy/presidencia/politicas-y-gestion/mision-del-grupo-asesor-cientifico-honorario-gach). GUIAD-COVID-19 activities can be understood as a mix of. The knowledge creation research activities of GUIAD-COVID-19 (yielding scientific reports that are published in the group's website https://guiad-covid.github.io/), were originally proposed to follow the national regulations for research on humans, which includes external ethical evaluation of research projects whenever needed. However, the GUIAD-COVID-19 had around 50 members including researchers in engineering, mathematics and biology and

health professionals and therefore, a consensus on how to comply with those regulations under the data centrism, open science, data science and mutidisciplinarity could not be attained. Also, participation in GACH gave some members of GUIAD-COVID-19 access to disaggregated personal data at the national level, which was needed for advising activities. This required signing confidentiality agreements with the Ministry of Public Health. Part of the knowledge created under these conditions was published as operative and/or argumentative documents emanating from the GACH. The activities performed by GUIAD-COVID-19 and GACH also included performing public appearances generally aimed at bringing the scientific perspective on the pandemic closer to the general public, a labor that notoriously evolved on digital press and broadcast media as well as on digital social webs.

The development of bioinformatics at the national level in an organized manner reached a milestone in 2009, when the first master's degree in bioinformatics in the country and possibly one of the first in the region was created. In addition, in September 2014, an agreement was signed between all the institutions supporting the URUGENOMES project (Naya, 2021), which obtained non-refundable financing from the Inter-American Development Bank (IDB). Importantly, the project helped to unveil a part of the history of Uruguay, shedding light on Uruguayan ancestry genomics, which is also key for understanding genomic variability in Uruguay and informing clinical genomics (Spangenberg et al., 2016). The URUGENOMES project contributes to science and healthcare in the field of rare diseases (which also serves human resources training). In fact, the creation of national capabilities for the analysis of human genomes was one specific objectives of the project. Under budget constraints and poor long-term policies, Facultad de Medicina has not yet been able to take full advantage of the new human capital that has been created in Uruguay (i.e. Masters and PhDs in Bioinformatics, Biology), restraining medical genomic and precision medicine development. The Urugenomes project serves also to exemplify the difficulties and advancements in the fields of the harmonization of science and social control and of the applications and challenging aspects of genomics research in Uruguay and globally. The project was evaluated in stages by a newly formed Research

Ethics Committee at Institut Pasteur de Montevideo, which was a learning process, at a time (2012-2014) when the Uruguayan regulation was less complete than it is today. This may explain why the original project proposed that "a database of variants and the original files" carrying detailed genomic information of participants would be eventually openly available (see http://urugenomes.org/base-de-datos/). We believe that there is no better way to exemplify the natural tension that exists between genomic open science progress and participants, close contacts, group and third-party protection. Applying the genomic methodology in the study of individuals suffering rare diseases provide a frequent exception to that generality. Although each rare disease has a low frequency, they total more than 7000 diseases, hence a substantial proportion of the Uruguayan population is suffering from them. Whole exome and genome sequencing (WES and WGS) technologies have been demonstrating their diagnostic potential, reducing costs and substantially shortening diagnostic times in rare diseases. 30 genomes of individuals with rare diseases (and in some cases direct relatives) were sequenced to evaluate the possible impact of the systematic application of this technology in the diagnostic protocol in Uruguay as part of the Urugenomes project. The results showed that more than half of the cases analyzed were diagnosed at the molecular level from the complete genome(Raggio et al., 2021; Spangenberg et al., 2021; Spangenberg et al., 2016; Spangenberg et al., 2019). International figures show diagnostic rates of 41% with whole genome sequencing (WGS) and 36% with whole exome sequencing (WES)(Clark et al., 2018) and a study has shown that the use of whole genome sequencing (WGS) can have a diagnostic success of 62.5% (Liu et al., 2019).

Table I. Authors and their participation in the reported experience

	CIDASH-(I&II) (2018-2019, 2020-2022)				**GUIAD-COVID-19 (2020-2022)**				**GACH (2020-2021)**				**Urugenomes (2014-2022)**			
	a	b	c	d	a	b	c	d	a	b	c	d	a	b	c	d
HB	■	■	■		■	■	■				■					
HN			■					■			■		■	■	■	
SM			■				■									
CS			■												■	
VM																

a, initiated; b, coordinated; c, integrated; d, evaluated

Perceived opportunities, challenges, issues and potential solutions in human health research

We have an opportunity to collectively unveil and work on structural limitations that hamper relevant health research. We think that we all have experienced structural limitations when practicing or evaluating health research, and there will always be such constraints. By structural we mean a limitation with roots in existing institutional arrangements or other constraints beyond those of individuals. Adequate social control through, for example external ethical evaluation is not considered a structural limitation (see below). These kinds of issues can self-perpetuate in pernicious contexts putting at risk the quality of research, the performance and behavior of researchers and the health and wellbeing of all, principally of the most vulnerable and those already harmed (vulnerated) (Kottow, 2012). We felt ethically compelled to work on uncovering structural limitations in health research education, social control and research infrastructure.

Issues in social control of health research and health research education. Social control of science is sometimes perceived by researchers as a structural limitation (Wolpe, 2006), but those constraints are purposed at protecting individuals, institutions, collectives and the society as a whole from research-related risks and harms. On the other hand, social control and bioethical evaluation are not monolithic or stationary and must respond to new issues and contexts. We understand that the poor education of researchers is a structural reason for disvaluing the role of social control in health research but also that there are deficits in regulation. These issues are intertwined. The quality-ethics evaluation processes in Uruguay is suffering from several problems in our understanding. On one side, Uruguay has a fragmented regulation covering many facets of health research (MSP, 2008; Uruguay, 2006, 2008a, 2008b, 2009a, 2010, 2012, 2014, 2019, 2021), etc. and the application and control of the actions implemented under these regulations is heterogeneous (with respect to institutional support, accreditation status, integration, etc). It must be stressed that the performances and characteristics of research ethics committees vary widely in Uruguay as well as in other countries, with also varying roles for central structures (Vidal, 2017). This is really important since Research Ethics Committees (RECs)

are cornerstone organizations of the health research evaluation systems. The fulfillment of Decree 158/19 requires bigger institutional support, particularly for the functioning of the RECs and clear messages from the responsible institutions (centrally the Ministerio de Salud Pública). On the other side, scientists are heterogeneous too with respect to their preparation for health and human research. Human health research ethics committees must evaluate the credentials of the researchers based on their CVs only. This is unlike for laboratory animal research were formal accreditation of animal research scientists currently exist (Uruguay, 2009b). An additional cause of educational heterogeneity arises probably because the programs at the graduate level in different careers may not provide today the experience in human research needed. Therefore, the system is highly heterogeneous, both at the researcher/research team and external evaluator/research ethics committee sides. Another aspect that deserves special attention in Uruguay is the lack of an effective research follow-up process on participant safety and data security after the initial ethical approval, which is a established activity of the research and innovation system in more developed countries.

We suggest that harm associated with risks identified in relation to research data and information management, processing and/or analysis must also be more stringently addressed in order to protect participants and ensure that researchers, institutions and funders remain accountable, as the responsible health research principle indicates. However higher stringency can only be effectively achieved without compromising the volume of research after increasing human training opportunities (formal education and proper experiences), better research infrastructures and adequate incentives. Instruments for participant safety and data security monitoring requires public investment, probably in several forms (UNICEF, 2005). The case of the creation and functioning of an Open Science Health Research Infrastructure will surely require a specific law with derogative effects on several regulatory texts, but more importantly, will give to academics and the government the opportunity to co-design and co-run an OSHRI, one of the most relevant pieces of the health knowledge generation machinery, one that has a high structuring potential, but nevertheless, just a piece of a system.

We think that to co-design and co-run an OSHRI, there are risks and risks balances we need to better understand. Particularly, those involving the already recognized personal, group, community and social risks associated with either a) not pursuing health research on humans (Jones et al., 2017), b) performing health research extrapolating Mertonian norms (universalism, communism, disinterestedness and organized skepticism) (Merton, 1973) or c) controlling human and health research in different ways (including but not only, independent bioethical evaluation). These balances can be considered either in the traditional or in the alternative OS context. Critical interdisciplinary elaborations will be very welcomed in these respects, particularly including distant disciplines, which we think would be highly beneficial for arriving at wider agreements and better results. Being aware of cultural gaps (see below) and looking at the complementarities between different methodologies, ontologies, epistemologies, ethics and theories could be useful to that end.

Issues related to the hard-soft science-humanities gaps in the education of researchers interested in human and social affairs. We think there are issues in interdisciplinary research that are related to the hard-soft science-humanities cultures gaps (for example in the infrequent integration of mathematicians and philosophers in health research teams). In this respect, it is a good thing that researchers from other fields are showing increasing interest in HR (recently due to the COVID-19 pandemic). In our experience, the roles of critical reflection on objectives, project writing and bioethical evaluation have been particularly difficult to accept as mandatory for HR by the new coming researchers during the response to the COVID-19 pandemic in Uruguay. We have experienced clashes that are typical of nascent true interdisciplinary work, which we think should be followed from a slower and safer learning curve. The health research and healthcare communities have already acknowledged the problem of biased databases (sampling bias), artificial intelligence/data science algorithms and resulting models (Knight et al., 2021; Panch, Mattie, & Atun, 2019). However, a key understanding we wish to provoke is that a bias-aware Open Science community is not enough. We agree with other authors who had already concluded that the previous “anonymize, release and forget” (ARF) praxis

concerning personal data handling should be abandoned (Lomas, 2019; Nowakowski, 2016; Ohm, 2009). The reason for that conclusion is that it leads to privacy and autonomy violations (including the pervasive self-privacy-autonomy violation engineered by the global data giants like Google and Facebook) and elevating the risks of re-identification, surveillance, control and harm to individuals, collectives and to all of us (i.e. through structural racism, Nazism, sexism, gender harassment) (Carrigan, Green, & Rahman-Davies, 2021; Lecuona Ramírez & Observatori de Bioètica i Dret, 2021). Therefore, we agree with Burtch when she proposes that "ethics needs to be made a cornerstone of any quantitative learning program, including traditional academic degrees, data science boot camps, MOOCs (massive open online courses), and everything in between", cited in (Franks, 2020).

Issues deriving from a low prioritization of research activities in the academic, public and private health sectors and an avid global and/or professional market for engineers and PhDs. Stronger public funding and strategic academic and public research agenda building are critically relevant for a more productive government-academy interaction, particularly when it comes to build new better facilities and to utilize knowledge for health and development. Unfortunately, raising levels of investment in science in Uruguay has proven to be very difficult, and will probably be a limiting factor in the coming years; therefore, global and regional funds could be needed. It is widely acknowledged that a RI proposal has to be evaluated, in part, in terms of its costs, benefits, and sustainability (Florio, 2019), but there are other dimensions that, in our opinion, need the same level of attention. In addition to the identification of medium- and long-term health problems and health objectives, work must also be done to identify the possible rationales and general characteristics of the projects that will be supported by the HRI, including reflection on the multiplicity of health research and action paradigms that are available and most valuable (see below). To this end, it is necessary to consider already advanced experiences as well as relevant characteristics at the local level (i.e. the structure of our health system, our geography, epidemiology, culture, etc.), with internal and external histories. We are convinced that planning a HRI must also explicitly include from the start

the consideration for the health, human development, bioethical and biopolitical issues and challenges that, as a double-edged sword, the HRI will address and generate. Critical to any HRI is the founding human capital and a strong user community, and both may well be scarce in Uruguay (see for instance (L. Méndez, Pellegrino, Robaina, & Vigorito, 2019)). Among related phenomena, we consider that the "Brain Drain" that affects the academy, that is researchers and professionals (including data administrators and experts in new computing processes) either leaving the country, leaving early the academic circuit to the private sector or working for abroad (directly or indirectly), will be a great limitation that should be correctly addressed. Repatriation and/or nationalization of Human capital (called "Brain Gain"), should be considered to complement local capabilities, along with the right incentives, strengthening, and creation of proper educational programs. Educational programs as well incentives will best develop later with a fully functional HRI maintaining regional and global scientific knowledge exchange. The institutional placement of a HRI is also a key issue, because it influences the mission, governance, funding and relation to users and other stakeholders. Nevertheless, we believe that Universidad de la República has a strong nucleating capacity, and could be the institution giving the first step ahead. We also suggest that the responsibility for planning HRI needs to be shared with the Ministry of Public Health, which could also be the hosting institution, if different alternatives (i.e. a national health research institute) turn out to be inappropriate. The support/participation of other government divisions, academic institutions and national or health committees and institutions will be needed, as well as a consult to all the democratic national political parties. In the case of CIDACS, for example, the Centre relies on - aside from the continued support of FIOCRUZ and the Ministry of Health for its mere existence - the collaboration of the government institutions that hold the original databases, to be able to maintain its own databases up to date, and thus continue to contribute in a significant manner to the development of scientific knowledge with the aim of benefiting the society.

Issues associated with the reuse of genomic participant and patient data in current global and local scenarios. We consider prominent changes that

challenge the role of research ethics committees in project evaluation in developed countries that influences the world in several ways (import of technologies for example), particularly those that are related to genomics. First, driven by the explosive surge of artificial intelligence and big data projects and tools, new institutional ethics review boards are being established in private and public sector institutions and a myriad of ethical codes and guidelines are being produced by professional associations and companies (particularly in the USA), which generates a new diversity of contexts in which to remain independent as well as a new set of difficult tasks for research and clinical ethics review boards that, in many occasions are ad-hoc or not experienced as a collective (Ferretti, Ienca, Hurst, & Vayena, 2020; Friesen et al., 2021; Hagendorff, 2020). The consequences of these issues spread easily in a globalized economy. Second, precision medicine, which strongly relies on the study of human genomic variability (as well as in specialized laboratory and informatics infrastructures), at the same time that it is reducing some inequities, it is generating new ones and new risks for participants, patients and third parties (non-participants and other patients) derived from the management, analysis and accumulation of data, which may be needed for the employment of high price medications and procedures, often under pressures from the industry and health insurance companies (Cutler, 2020; Gronde, Uyl-de Groot, & Pieters, 2017). Third, Open Data Science (ODS), meaning science made with the re-use of open data (data that was collected for a previous purpose and has then been published, which can be re-used without the need of new permissions, without costs, by anyone and for any purpose) (Murray-Rust, 2008) is proposed to be key for development in all fields. Open DS is characterized by sharing and disseminating all data, the desideratum of OD citizen science movements (van de Sandt, Dallmeier-Tiessen, Lavasa, & Petras, 2019). Fourth, in our experience, the birth of a high performance society at the turn of a millennium, where “data is the new oil” fueling artificial intelligence, displacing traditional technologies, revolutionizing the industry (producing the so called Fourth Industrial Revolution, 4IR) but also changing science, healthcare and every aspect of our lives is a view on the future provoking a broad range of emotions, from excitement to fear. A key ingredient of the 4IR, is the automatic generation of knowledge, decisions and actions within cyber physical systems yielding goods

and prices with personalization (Gleason; Moavenzadeh, 2015; Rekettye & Pranjić, 2020). The ubiquitous data driven-automatic decision making of the 4IR has to be engineered, tested and adopted (mostly in unintended or forced ways) by end consumers. One aspect of this prominent change might be the imposition of Open Data Science itself. We note that some reports dedicated to criteria that should be observed prior to opening government data do not consider bioethics and biopolitical issues at all (Yang, Lo, & Shiang, 2015). Under these conditions, researchers, stakeholders, academic institution leaders and government authorities should ask: how is open genomics being targeted by the 4IR to respond to capital and not the health and human development imperatives?

Genetics, genomics and bioinformatics for human health is an area of current interest but still lagging in Uruguay. The discipline "genetics," as traditionally defined (the study of heredity related to mono or oligogenic traits), represents now just one approach within the wider genetic (including quantitative genetics of polygenic diseases) and genomic methodologies. Genomics has a strong connection with bioinformatics, and both of them are at the root of precision medicine. Despite the long history of genetics in Uruguay (Garzón, 2010), clinical genetics is a rare option for postgraduates in medicine, therefore, medical services related to the discipline are in shortage in Uruguay. This not only complicates the stronger development of clinical genetics itself but also constitutes a structural limitation for better healthcare and for the advancement of genomics and molecular medicine art and science as a whole. In fact, unlike Brazil, Argentina, Colombia, Venezuela, Chile and Peru, Uruguay has not yet integrated high throughput genomic studies as part of "normal" healthcare (i.e. cancer diagnosis and treatment) (Phillips, Douglas, Wordsworth, Buchanan, & Marshall, 2021), with the recent exception of industry-dependent pipelines at Fundación Pérez Scremini (ROCHE, 2021). In this respect, it is worth noting that the transnational direct-to-consumer clinical genomics testing business model promoted by the industry has strong drawbacks, including little contribution to science and social development and lower rates of success for individuals and groups (Rodrigues, 2020). A problem we expect to aggravate with health OS, which is not new for the clinical research and bioethics community, will be associated with the (ill-

intentioned) control of science and patient organizations through funding, networking and infiltration (Lundh, Lexchin, Mintzes, Schroll, & Bero, 2018; Moynihan & Bero, 2017; Taylor & Denegri, 2017). This may be facilitated by weak regulations, uncoordinated responses from the government and academia and poor public investments in health and healthcare science.

Issues in project evaluation preventing accountability and a proper evaluation of the risks of harm to third parties. Some risks of health research are off-focus topics in most current bioethical evaluation contexts (consider for example basic research, reviews of the literature, the study of open health or behavior datasets). Some already recognized risks to third parties include bystander risks, inductive risks, the dual-use dilemma (as mentioned in the case of the HBP above) and risks in particular subject areas (Herington & Tanona, 2020). Traditionally, independent evaluation of proposals for health research on humans follows independent ethical reflection and deliberation by members of institutional research ethics review boards (RECs), in which accepted general principles and rules of wisdom are used for tailored case-specific decision making. Under this framework, the "first do no harm principle" with roots in the *Corpus hippocraticum* ("to do good or at least to do no harm") (Craik, 2015) means that the risk of (physical, moral, psychological, social, legal, and financial) harm for participating in a traditional research project is only acceptable i) if not clearly inacceptable and ii) if consensus is reached that it is balanced or outweighed by potential benefits to the same and/or each person in the group. The analysis can be justified to be extended to the participant's most strong contacts (i.e. direct relatives) and to the groups the participant belongs to (i.e. patients with the same disease). Only potential benefits to the participants matter when evaluating balances related to those directly involved in the study (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research Department of Health, 1979). Therefore, potential benefits to society related to OS adoption and other practices belong to a lower priority tier when planning and conducting traditional human health research (biomedical, behavioral, epidemiological, etc.). Moreover, the conduct of researchers must follow a given path: first, participant risks must be minimized, then participant

benefits must be maximized and later, if a negative net risk-benefit balance for participants persists, the net risk of harm to participants must be significantly outweighed by the social benefits of the research, only if the risks and the effects are acceptable by participants and the REC (UNESCO, 2005). While this and other principles are widely accepted in the health research on humans community, a turn to the "do no net harm" is being processed in biological and social sciences and in evidence-based social action only in recent years (Davies & Bawa, 2012; Putzel, 2010). What some authors propose is that extending the traditional bioethical evaluation by RECs to projects into the new contexts of research on humans (i.e. big data, complex problems) should be done but is and will be increasingly less encompassing (Herington & Tanona, 2020; Vayena & Blasimme, 2021). Harm to third parties due to the findings of research are also of concern (Hausman, 2007), but less appreciated in the literature. It could be the case of a review of the literature or studies using animal models. This might be a relevant problem deserving research given that as scientists we are currently exposed to and misguided by wrong incentives (i.e. the "publish or perish" normative). The current practice conditions that these kinds of research risks can only be evaluated by peers before publication, an evaluation process which concentrates on the scientific validity of the work already done and does not take explicitly into account the possibility of harm to bystanders, social groups or the society as a whole before and after publication. Researchers, reviewers, editors and scientific journals are currently unaccountable for harm to non-participants due to research findings (which may be even more important if using biased datasets and/or black box automatic methods). Some authors have agreed that institutional research review boards are not prepared for the evaluation of third party risks (Hausman, 2007; S. K. Shah et al., 2018). Also, the do no net harm principle in some instances requires that the risks to which participants are exposed should be outweighed by social benefits (the social value of a research). It has been proposed that the social value of a research project is something often difficult to weight and well beyond the capabilities of most institutional review boards (S. K. Shah et al., 2018). Social control of science is not stationary as we mentioned above. Consider for example that until recently, epidemiological observational research was widely regarded as not raising significant ethical

issues and was commonly carried out without approval of an ethical review committee, but that has changed already (CIOMS, 2009). So looking ahead, expanding the scope and improving the systematization of project evaluation and monitoring in a few more dimensions is an eventually relevant direction worth following for the sake of health research and open science".

Issues in the access to the benefits of health research. The inequity in access to the benefits of health research will likely continue to increase. The majority of large-scale genomic research projects have been conducted in Europe, the USA and Canada. Therefore, we predict that a large inequity in access to this type of analyses and knowledge will continue to develop in Uruguay, which compels action. In fact, it has been concluded that the benefits of genomic studies hardly reach patients around the world equally. Even in developed countries, low-income families' access to genetic and genomic analyses is severely hampered. Public programs at the national level to facilitate access to this type of diagnosis for patients with rare diseases and other conditions of possible genetic origin have been implemented in a few countries (Senier, Tan, Smollin, & Lee, 2019; Turro et al., 2020). While human genomics has shown results and potential, clinically and scientifically (as briefly exemplified above), we acknowledge that evidence have been provided that the promises of genomics observed at the discourse level have had a great influence on science and technology, funding decisions in developed countries (which presumably would not be fair there and here) (Lemay, 2020), implying that proposed impacts Genomic RIs have not been met yet even in developed countries. The same reasoning applies to areas such as precision health in the USA and Europe (Hoeyer, 2019; P. D. Shah, 2021). Therefore, we ask, what could be the rational, legitimate and ethical plan for developing genomics, precision health and the needed research infrastructures? How can genomics (in its wider sense) data be linked to the causes of the causes to obtain multilevel health causation models that could show us new and more effective collective health interventions? Earlier proposals to national authorities to enter into what could be called an era of "genomics commons" for Uruguay did not succeed, which may have many explanations. It might be time already, we think, to promote a national ethical and scientific reflection on those issues.

Can the uncertainties and risks related to open science health research be justified on the basis of the promise of a greater common good for a greater number of people, despite being potentially harmful to some people or groups? In the bioethical analysis of randomized controlled trials, depending on the ethical theory we stand from, there are certain risks that can be justified when the result contributes to protecting a larger number of people or communities (this applies to consequentialism) (E. J. Emanuel et al., 2008). Since a single and certain answer to this question may not exist and the evaluations involved might be difficult (Herington & Tanona, 2020) particularly in resource-limited contexts, the issue requires collective compromise and reflexive discussion, rather than avoidance. The privileged place the health research community gives to ethics and the fact that OS has been considered a necessity for confronting with professional ethics issues (i.e. in relation to climate change data and knowledge) (Cai, Judd, & Lontzek, 2012), may be indications of a larger than expected compatibility between OS, health research and health policy. Recent experiences related to the COVID-19 pandemic give us a lot think about in these respects.

Structuring pillars for the development of open human health research and health research infrastructures in Uruguay. Experiencing health research and open science in these last five years or so has been challenging to us. We have made mistakes, had self-reassuring moments, read, received advice, we have elaborated new projects and ideas, etc. Hence, we think we have learned something we wish to share. We highlight some of these insights, in the hope that they might be found useful for the development of fruitful and respectful open health research and health research infrastructures, therefore we score them as "pillars".

Pillar 1: The final objective and characteristics of the HRI and the health research it facilitates must be defined in advance and critically revised. As we mentioned above, HRIs must be conceived as double-edged artifacts, with potential for benefits and harms, as the same occurs, of course, with scientific research in general. Therefore, an HRI can be considered a Dual-Use Innovation of Concern (DUIC, see below). Health itself is an abstract concept for which many definitions exist. If we accept that health as a concept and value is dynamic,

contextual and local, we may also be prepared to embrace its complexity. Therefore, defining the final objective of health research is not an easy task. The important message here is that despite its complexities, it is worth to make relevant choices. The global and national scientific response to COVID-19 has plenty of examples of misguiding and secrecy, where scientists played various roles, but in general our participation was aimed at engaging either with research-informed policy-making and researching to produce potentially useful knowledge (useful for policy-making and/or policy evaluation). When national authorities were asked about the objective of the work requested from scientists in the research-informed policy-making boundary the answer has been "to scientifically advise the Presidency of the Uruguayan Republic"... on the road to "the new normality""
(https://medios.presidencia.gub.uy/tav_portal/2020/noticias/AG_280/Informe_GACH.pdf). Of course, this is not an example of good guidance based on clear objectives. A major secret has been the cost-effectiveness formula the national authorities have used to guide actions worldwide, a topic we will not treat herein.

Pillar 2. The academic community (particularly the RI staff) should perfect on their matters, which includes to remain open to explore, debate, choose, and be critical about the health research paradigms that impregnate their work. Quoting Edgar Morin "El conocimiento es navegar en un océano de incertidumbres a través de archipiélagos de certezas" ("Knowledge is navigating an ocean of uncertainties through archipelagos of certainties") (Morin, 1999). Philosophical (ontological, epistemological and ethical), theoretical, methodological, ideological, religious, gender and other relevant biases as well as conflicts of interest should be distinguished and critically addressed (if possible) when performing and evaluating health research. Reflecting on what we should expect from a given design for an OSHRI in terms of better health and healthcare requires revisiting our assumptions about how we think about and act on health. For example, we find at least five radical perspectives (by radical perspectives we mean that they are often treated as independently acting factors in generalizable hierarchies) proposing that health is caused, "determined" or dependent on "determinants" by: 1) socio-economic factors, defined as "the material, social,

political, and cultural conditions that shape our lives and our behaviors" (Michael Marmot & Allen, 2014) or as "the conditions in the environments in which people live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks. For the purposes of this report, the social determinants of health are education; employment; health systems and services; housing; income and wealth; the physical environment; public safety; the social environment; and transportation and cause variable exposures and inequity" (National Academies of Sciences & Medicine, 2017); equity, a central concept that deserves attention, has been defined variably, as "the absence of avoidable, unfair, or remediable differences –inequalities- among groups of people, whether those groups are defined socially, economically, demographically, or geographically or by other means of stratification" (Hill-Briggs et al., 2021), "the state in which everyone has the opportunity to attain full health potential and no one is disadvantaged from achieving this potential because of social position or any other socially defined circumstance" (National Academies of Sciences & Medicine, 2017); 2) factors related to human agency, individual choice, will, self-determination and control (Deci & Ryan, 2008; Tsey, 2008); 3) biological factors from within the individual or in relation to host-parasite relations, mostly genes and genomics that determine general or differential susceptibility (or virulence); 4) physical and biological environmental factors (National Research Council, 2006) or 5) health care systems factors. Some authors treat the different factors as constituting ethico-political dilemmas (i.e. social vs will determinisms), trilemmas, etc. Rather, in agreement with other authors, we acknowledge that all these factors combined, and others too, cause complexity and health effects at the individual as well as at all organization levels, requiring the inference of models on the basis of knowledge and appropriate data, capturing the dynamic behavior of the systems.

How we understand causation deserves attention because the concept is critical for evaluating the reasonableness of belief levels and expectations implicit in the statements above. If we say "Health is socio-economically determined." we mean "Health is only socio-economically determined" (periods in the sentences are used to emphasize completeness), and we are thus neglecting all other influences,

which in addition we must know completely. In other words, in general, scientific research can create information (evidence) that a given model is inferior to another on the basis of new/more data, new/more knowledge (including methodologies as technologies and theories) and/or new/untested hypotheses. These models can be deterministic, stochastic or mixed. "X determines Y" is sometimes used in place of "X causes Y", but the former should be read today as "X causes Y deterministically" (not probabilistically or with probability of 1). The explanation for that can be traced back to Hume who only considered deterministic causality, hence our inherited notion generally assumes deterministic relationships, homogeneity and then law-like behaviors (Dupré & Cartwright, 1988; Neumayer & Plümper, 2017). Causal stochastic models may not rely on explicit detailed chains of events and cause and effect relationships and cannot predict at the individual (single event) level (Pearl, 2009).

Pillar 3. Provide bridges for the health data repository archipelago and provide a "hub" for integrative knowledge creation. The advancement of genomics (regionally as well as globally) requires a quote of scientist modesty, because besides a few cases (most of them already taken up by genetics), the complexity of the causal genome-phenome relations is not currently well understood (Hebbring, 2019; Pellman & Zhang, 2021; Riolo, 2021; Sakaue et al., 2021). Considering that open genomic data science advancement will continue to be a condition for generating the extensive and accurate datasets that causal genome-phenome relations inference require (Song, Huang, Zhang, Bates, & Wright, 2018), that participant-nonparticipant harm prevention demands continued adjustments, we propose that a public health research infrastructure integrating omics and phenomics data could be relevant. In addition, it may increase the genomic sovereignty both at the technical and knowledge levels, making analyses and clinical results more probable and of better quality. A national genomic data harvesting system would most likely involve interoperability between clinical and research infrastructures. In these directions, it is worth considering the experiences of Japan and Uruguay. The potential of electronic health records (EHRs) for containing genomic data and fostering genomic and precision health had been identified long ago (Martín-Sanchez,

Maojo, & Lopez-Campos, 2002; Masui & Takada, 2003), a potential that, in order to be realized needs public national biobanks and proper regulations (Crawford & Sedor, 2021; Linder, Bastarache, Hughey, & Peterson, 2021; Yamamoto et al., 2018). Notoriously, Japan included genomic data under two categories: the traditional category of personal information (name and genomic data together) and genomic data alone (without identifiers) under the category of "individual identification code information" -a category reserved for identification information, such as passport numbers, fingerprint scan data, social security numbers, etc.- (Yamamoto et al., 2018). Interestingly, direct-to-consumer genetic testing services were becoming more common in Japan at that time, and some researchers and medical professionals were worried about how personal information including genomic data accumulated at those companies would be handled (Yamamoto et al., 2018), a process which is happening now in Uruguay. In parallel to regulatory changes securing personal genomes, Japan was also building biobanking infrastructures promoting genomics and precision health at the bench, computer and bedside (Nagai et al., 2017; Takayama et al., 2021). We think personal genomic data is identificatory *per se* even after anonymization, and therefore it shouldn't be treated as open data or be freely shared by researchers. The path followed by Japan is of course not the only possibility; there are different flavors of biobanks and genomebanks (O'Doherty et al., 2021). In our opinion, the advantage of integrating genomes and EHRs is huge if the information system is properly developed, because it helps bridging the genome-phenome gap (i.e. the identification of potentially causal relations) (Linder et al., 2021). Recommendations on how to integrate genomic (and in general omics) data into EHRs have been produced and some implementations following them are available (Grebe et al., 2020; Lau-Min et al., 2021). Multiple, small (institutions) to bigger (public/private sector) jurisdictions within a country will clearly generate difficulties. Interoperability constrains are typically the higher barriers in home grown EHRs (EHRs developed and tailored locally to each health institution, a characteristic of our EHR system). Using genome and phenome data has two extreme solutions; either i) participants in a disease-specific cohort consent and renounce to their right to privacy, and data is shared by selected researchers with clear purposes (project evaluation and monitoring

needed) within a national or international system (Chan et al., 2017; Network, 2021); or ii) data is protected from researchers but used within a federated secure edge data sharing and computing research system, where participants are not asked to consent any rights renouncement (at bigger implementation costs, but lower harm risks) (Rehm et al., 2021; Voisin et al., 2021). Participants could in the latter case, dynamically consent on their participation in one, a few, or all research projects associated with the database. Highly qualified human intervention would be needed in both cases at several points. Intertwining of the healthcare and health research systems would be higher in the last case. The Uruguayan EHR information system –"Historia Clínica Electrónica Nacional" (HCEN)- evolved within the framework of the Salud.uy program at Agencia de Gobierno Electrónico y Sociedad de la Información (AGESIC), and became operative in 2018 with the purpose of bringing together all public and private health actors, around strategic medical informatics definitions with a user-centered approach (Pidre et al., 2018). HCEN represents a clear example of organizational interoperability at the level of a health system, where the main complexity is centered in the multiple actors and centers involved and the related technical specificities (many small jurisdictions plus home-grown local information systems). The main function of HCEN is to acquire patient-related clinical data for use and analysis by the healthcare team, being the core of the health information system (Chá Ghiglia, 2019). Even though connecting the HCEN with an OSHRI is a considerable technical and financial challenge, the payoff in terms of quality and quantity of data for health research and knowledge creation will likely be significant. Future implementations may use distributed ledger systems, blockchain or a better technology to afford reasonably secure and efficient communication paths in order to integrate genomics, clinical and administrative data, preserving participant data security (Rehm et al., 2021; Sanmarchi, Toscano, Fattorini, Bucci, & Golinelli, 2021). The OSHRI could provide the data integration, processing and analysis environment needed for secure/safe and ethical open health science without personal data dissemination.

Pillar 4. Researchers and health authorities should be aware and responsible by design of the biopolitical consequences of pursuing and not pursuing

health research. Recognition of the existence of different personal, group, community and social risks associated with: a) not pursuing health research on humans (Jones et al., 2017), b) extrapolating pure Mertonian norms (universalism, communism, organized skepticism and disinterestedness) to human health research, or c) controlling science on humans in different ways (i.e. through independent bioethical evaluation). These possibilities can be considered as taking place either in the traditional or in the alternative OS context. First, we ask: which is this desired common good in health research and open science? Does it arise from consensus or does it respond to the interests of particular groups? Is it ethically reasonable to accept all types of risks assuming they are known and have been significantly minimized? In order to answer these questions, it is necessary to think about who defines that "pursued good" and what interests and ethical values it responds to. To identify possible problems and moral conflicts, we propose to briefly contextualize this problematization from three conceptual lines: biopolitics and bio-power, open science as a movement, and bioethics. From the perspectives of biopolitics and biopower, we wish to consider concepts resignified by Michel Foucault as part of his theory of power (Foucault, 2001, 2007, 2008). In his works, biopolitics appears as a concept that refers to the way in which the analysis of man's biological life becomes the object of politics, in an image where the power of knowledge is constituted as an agent of transformation of human life (Foucault, 2000, 2001, 2007, 2008). Surveillance and control of birth, disease, life expectancy, mortality, migration, and today we would add human mobility become part, according to the author, of a form of exercise of power, a "bio-power," and a fundamental pillar for the development of the capitalist production model (Foucault, 2001). The value of the body and health, as well as the distribution of the productive force were central elements for this development, reconfiguring the political struggles based on the affirmation of the right to life and the body. These struggles embody moral evaluations that change in different historical moments, since life, the determination of its value, illness, and death, are not simple facts. It is enlightening to think of historical examples that account for this process and that were developed by the author in his genealogy (Foucault, 2001). In the XX century some ideologies supported genetics as the basis of the purity of a race and "quality" as criterion for descent took the stage. Under the

supremacy of one race over another, crimes have been perpetrated, in the name of that sought-after good. In this sense, we understand that, depending on the definition of health and criteria of normality from which these parameters are established, as well as the configuration of actors in the power network, different actions will be produced for the administration of the life of the population or some strata by governments and institutions, as well as implications in the way of producing socially legitimized scientific knowledge. Hence, health research and human actions in general should be considered fragile, open to committing injustice and producing harm, explaining why ethical control and biopolitical analysis are so relevant for mature scientific communities, particularly if considering the planning of a HRI or the future development of genomic and precision medicine. In a globalized world, the health crisis caused by SARS-CoV-2 (and the causes that caused it) catalyzed and starkly showed the paradigms supporting health decision-making in different countries, as well as the industry's race for patents for the generation of vaccines and associated knowledge, to which access is widely known to be dramatically inequitable and globally dangerous (Herzog, Norheim, Emanuel, & McCoy, 2021). Likewise, this scenario promoted a so-called open science policy, putting it at the center of the debate (Molldrem, Hussain, & Smith, 2021). We can identify in the OS proposals the principles that guide it: the conception of knowledge as a public good, openness of access to publications and data managed by academia in conjunction with government institutions, and the absence of profit (Babini & Rovelli, 2020). Citizen participation and inclusion appear as driving ideas in this concept. To us, the proposal implies, on the one hand, a way of producing knowledge, and on the other, a series of consequences in the ways in which the assets destined for the financing of research are distributed and managed. In our understanding, this is a key point, since it would only be a discursive change if new forms of participation were not processed in the definition of these criteria. In other words, if current research priorities and resource allocations are defined by funding agencies with public or private funds, citizen participation would be essential, under the risk of replicating biases in the definition of priorities. This aspect can be understood as the value of an investigation, a requirement that together with scientific validity, fair subject selection, favorable risk-benefit ratio, independent evaluation,

informed consent and respect for enrolled subjects were systematized by Emanuel (E. Emanuel, 1999) and are part of a rational framework for the ethical analysis of biomedical research with human beings. According to the author's proposal, they are all universally applicable and are ordered in a logical sequence, constituting declarations of value that must be specified in each case and particular context. We will focus especially on two of them that may shed some light on aspects that seem central to this reflection: value or relevance and the independent evaluation of research proposals.

The value of an investigation represents a judgment on its social, scientific or clinical importance where the contribution to the health, well-being or knowledge of the population should be valued from the impact of its results and the collectivization of its benefits, a requirement that stems from the need of a responsible use of finite resources and the imperative to avoid exploitation (E. Emanuel, 1999). In other words, the estimation of the value of an investigation does not correspond to scientific validity, a necessary but not sufficient requirement. Good research from a methodological point of view is not necessarily ethical, relevant or good quality health research from this perspective. According to Emanuel, independent evaluation is a requirement that is based on the fact that researchers have the potential for conflict of interest and therefore they must remain socially accountable (E. Emanuel, 1999). These conflicts of interests not only relate to lucrative aims, but can be associated with academic prestige, struggle for funding and positions of power or respond to corporate pressure (E. Emanuel, 1999). Although many of them may be legitimate, they can affect various aspects of the development of an investigation. In order to minimize these biases, the independent evaluation seeks to ensure that participants are treated ethically, preventing them from being used as mere means (human exploitation) and preventing damage or the disclosure of sensitive data and information which should remain confidential (E. Emanuel, 1999). Of particular importance in the context of this reflection, Emanuel suggests that in addition to scientific review boards and research ethics committees there should be data and safety monitoring boards (E. Emanuel, 1999). These instruments that should give support to health research have not been implemented in Uruguay yet. It is, in our

opinion, an ethico-political imperative to build these competencies, something the Universidad de la República, the national government, and the research system as a whole should acknowledge.

Science has been internationally and nationally agreed to be a Universal Human Right. It is included in the Universal Declaration of Human Rights in its article 27 (Nations, 1948) and the International Covenant on Economic, Social and Cultural Rights (Nations, 1966). The UNESCO Universal Declaration of Bioethics and Human Rights(UNESCO, 2005) recognizes the benefits for the human species of scientific and technological advances, provided that they seek to promote the greater well-being and quality of life of individuals, families, groups or communities, always guided by the recognition of the dignity of the human person and respect for human rights. Likewise, health is understood in a broad and multifactorial sense, so that it not only depends on the progress of scientific and technological research, which clearly generates impacts that must be recognized. It is argued that moral sensitivity and ethical reflection should be part of this scientific and technological development, and bioethical reflection plays a fundamental role in decision-making in the face of the possible problems posed by this development.

According to this moral minimum of respect for human rights (Cortina, 2000), new developments and ways of doing science should not only pursue the technological imperative, but also pose a challenge of new approaches to think about the values that are at stake in this social practice. We understand that these values are present in Open Science as an instituting proposal, however, as we have tried to show, various tensions and procedural aspects arise that, as a society, we will have to rethink in order to achieve the desired objective. In a similar direction, the recent UNESCO document recommends to consider the dialogue with other systems of knowledge and ethics as a pillar for OS development, specially including the respect for the human rights of indigenous people (UNESCO, 2021), as declared in (Nations, 2007). Developments such as an OSHRI including infrastructures for storage, preparation and analysis of genomic data entail dilemmas and particular ethical problems for the different areas –see Art 3. in (UNESCO, 2005). Finally, under the bioethical principle of

responsibility (Jonas, 1985), the human being must act with precaution when deciding which technological objectives to pursue, avoiding decisions based on the technological imperative alone, in other words: not everything possible must be done. Of course, this requires profound reflection. Nevertheless, it might be agreed that this imperative of responsibility might be more prevalent now that there is ample evidence of an endangered future for humanity, as recognized in (UNESCO, 2021).

V. Discussion

No practice takes place in a philosophical vacuum, and all artifacts have politics. This markedly applies to science and scientific technologies. This work makes some emphasis at analyzing human genomics as a techno-science in connection to health that is both increasing and closing inequity gaps. We agree with the view that we cannot expect to decrease inequity significantly if not acting on the causes of the causes (M. Marmot, 2018). Socio-economic factors are widely and reasonably considered the most relevant factors causing health variability, but by no means the only factors that matter. In addition, we expect that socio-economic factors cause both deterministic (for which we would assume homogeneity and known mechanism), stochastic effects (which we may assume are too complex to be explained and predicted) and perhaps also mixed effects. We should ask: does it matter? We think yes, but of course not knowing the nature of the effect and the mechanism of production does not mean we have our hands tied, we should seek to expose and remove the well-recognized damaging cause, as Breilh claims (Breilh, 2013). In addition, as researchers we can and should be on alert, sensitive to unexpected or counterintuitive results (both positive and negative) at one or many levels, which would give evidence of how uncertain we should feel about projecting patterns and regularities learned in the past or in different geographies, which, in turn should incentivize research on causal mechanisms at local levels, avoiding too Humean approaches (Pearl, 2021).

One thing we wish to consider is that big data and artificial intelligence have been expanding exponentially in the last 20 years, nevertheless, the complexity of real (i.e. biological, sociological, health) problems is still perplexing. Should

processed, organized and curated data waiting for accumulation and later explanation then be the main output of data-driven and OS genomic research? One assumption behind data accumulation in huge electronic stores is that some patterns in data can only be recognized and extracted if more data is available, but what happens if new patterns accumulate too with more data (linearly or not) as the context is also rapidly changing or if, alternatively, more data brings even stronger biases? Therefore, critical research in careful data accumulation and preparation for research and state-of-the-art analytical methodologies are both critical shared necessities of the health research community now and in the future. Traditional researchers that think critically know that the theoretical lenses we have been using and developing to look at data are imperfect. Skepticism about the scientific method and mainstream programs is not new to epistemologists. The reader may wish to consider that they might be reasonably skeptical about data-driven "artificial research" accomplishments and promises.

National regulation adjustments, ethical guidelines and the participative design of responsible research and innovation mechanisms will surely be needed in the near future more than ever, because risks, plurality of actors, education gaps and social, technological and political change are increasing rapidly. It might be the right time for our OS community to undergo an open process of reflexive synthesis, and eventually, adopt bioethical and bio-socio-political reflection and principles, as well as other participatory responsible conducts looking to improve approaches of ethical dilemmas in OS project proposal and execution, because they can strengthen and preserve OS, society and humanity in the long run (Beauvais, Knoppers, & Illes, 2021). The need for a convergence of ethical standards in different disciplines (i.e. social, medical and engineering ethics) stems out of the rapid convergence of data and communication technologies, multi/interdisciplinary research and globalization of science and problems (including loss of privacy), as has been evidenced in the still ongoing COVID-19 pandemic crisis (S. Méndez & Botti, 2021). "Rede CoVida" from Brazil and GUIAD-COVID-19 and GACH experiences from the Uruguay were different. Rede CoVida was launched by a group of researchers and communication professionals from CIDACS and the Federal University of Bahia (Universidade

Federal da Bahia - UFBA), a group that grew past the state boundaries to form an ample multidisciplinary team which aimed to support decision-makers in the fight against the pandemic in a timely manner, as well as maintaining the public informed. This was done by monitoring the evolution of the pandemic, creating mathematical models to predict its developments, and synthesizing scientific findings for dissemination (RedeCoVida, 2020). CIDACS' ample experience in both the management of administrative datasets and dissemination of scientific knowledge certainly facilitated this experience. While we consider that the comparative study of GACH, GUIAD-COVID-19 and RedeCoVida iniciatives should be done, including the study of their effectiveness, we observed that independent ethical evaluations were absent in all of them. We think that the comparative study of GACH, GUIAD-COVID-19 and RedeCoVida iniciatives should be done, including the study of their relative effectiveness. In addition, we observe that independent ethical evaluations were absent in all of them. GUIAD-COVID-19 included an internal group committed to provide ethical problematization, reflection and deliberation, which we think was poorly effective.

The Uruguayan health data research infrastructures are currently heterogeneous and fragmented, therefore weak. Nevertheless, they the point on which we must leverage the needed improvements. Screening for ideas when planning means to filter out the evidentiary incorrect from the still promising. Defining some dimensions of this analysis as critical is mandatory from the start of planning. In this respect we provide arguments that HRIs are important artifacts in current health and human research systems. The FAIR (Findable, Accessible, Interoperable and Reusable) data principles were originally proposed with the main objective of providing an ample agreement on minimal characteristics for storing, communicating, sharing and reusing data needed to foster data-driven open research and open government, and strongly recommend to publish data in complete Open Access wherever possible (Force11, 2014). The fairness of the FAIR data principles and related implementations are a matter of strong scrutiny. FAIR principles cannot be considered fair from the perspective of bioethical and biopolitical analyses (with the main objective of protecting and promoting growth

in people's life, future and dignity while conducting research). Adapting the original FAIR principles, making them fairer and harmonic with local regulatory developments, is especially relevant for an OSHRI. FAIR data principles are notoriously implemented in the Centre for the Integration of Health Data and Knowledge (CIDACS). CIDACS is part of the Gonçalo Moniz Institute of the Oswaldo Cruz Foundation (FIOCRUZ), a public institution attached to the Brazilian Ministry of Health. CIDACS represents a multifaceted initiative centred around the utilization of administrative data focused on the evaluation of the impact of social protection policies and programmes and other relevant health problems in the affecting the Brazilian society. Administrative data is not collected with the purpose of carrying out research, but rather it is gathered routinely for administrative purposes. Nonetheless, it represents a great potential for answering complex research questions (M. L. Barreto et al., 2019). In fact, the linkage of various administrative datasets allows carrying out comprehensive analyses that each individual institution contributing their databases would not manage to realize independently. The differential brought by CIDACS has been the utilization of administrative data to study the impact of governmental policies targeting the social determinants of health (e.g., social protection policies) on health outcomes, using individual-level data. In fact, previous data linkage initiatives in Brazil generally used this type of data to study the effects of common exposures on health outcomes (M. L. Barreto et al., 2019). The employment of administrative data for research can hold significant advantages (including large sample sizes, longitudinal structures, high population coverage and high data quality). Nevertheless, it faces frequent limitations (mostly related to data quality and missing data) (Harron et al., 2017) and a number of challenges. First of all, different datasets must be obtained from different government bodies on a routine basis, in order to maintain the linked datasets up to date. Secondly, an appropriate computational infrastructure needs to be in place, with the capacity to execute data preparation and linkage procedures on databases with millions of entries. The establishment of such an infrastructure requires a significant financial investment in both human and non-human assets. Finally, protocols must be established to ensure that the data remains secure and the privacy of the individuals whose information is being handled is preserved, in accordance with the legislation

related to data protection (M. L. Barreto et al., 2019). In Brazil, with Federal Law 12.527/2011 (the Information Access Law), access to administrative data containing personal information for research purposes began to require ethical approval by institutional committees, followed by the formal authorization from the relevant government data system administrators. As research involving personal data requires the explicit approval of institutional ethical committees, whenever individual consent is not feasible, ethical assessments are focused on risk-benefit analysis, participants' rights, measures to prevent harm to and discrimination of individuals or groups, researcher responsibilities and the monitoring of approved research (M. L. Barreto et al., 2019). However, in 2018 the Brazilian General Data Protection Law was passed (Law 13.709/2018), and became effective in 2020. This legislation, which was largely inspired by the European Union General Data Protection Regulation, established a series of principles and privacy protection measures. One of these determines that any individual whose data is stored in a database can request from the data guardian, at any time, information regarding the sharing of his/her personal data, the criteria and procedures used in the treatment of personal data and request the removal (opt-out) of his/her data from the database in question (Maurício Lima Barreto, Almeida, & Doneda, 2019). One of the strategies employed by CIDACS to preserve the privacy of identifiable data and curb the possibility of re-identification of de-identified data is to separate linkage and analysis processes (Harron et al., 2017). The Centre, in fact, has two separate environments for managing the data. On the one hand, the Data Production Centre is a secure room housing the computational infrastructure for treating and linking original identified databases. In this space are produced the de-identified/anonymized datasets that are then accessible to researchers in the Data Analysis Environment (DAE), under the terms and conditions established by the Centre (M. L. Barreto et al., 2019). When it was established in 2016, by decree 32-2016, the organization of CIDACS was designed to achieve the objective of creating the 100 Million Brazilians Cohort. Since then, its initial scope has further developed to incorporate new studies and additional partnerships, which required a reconfiguration of the Centre's organizational arrangements (M. L. Barreto et al., 2019; Mauricio L. Barreto et al., 2021). CIDACS databases are currently being mapped to align with

The Observational Medical Outcomes Partnership Common Data Model (OMOP CDM), so as to allow the Brazilian data held by the centre to be explored within the scope of international comparison studies (Stang et al., 2010).

Clearly, CIDACS implements its own version of the FAIR data principles, one that we think is FAIRer than the original when considering the re-identification and manipulation risks, we as society have been building up and must face today. First: CIDACS requires that the "Persons who wish to receive authorisation must: 1) Be affiliated to the institution or be identified as collaborators; 2) Present a detailed research project together with ethical approval by an appropriate Brazilian institutional review board; 3) Provide a clear data plan restricted to the objectives of the proposed study, and a summary of the analysis plan intended to guide the linkage and/or extraction of a relevant set of records and variables; 4) Sign terms of responsibility regarding the access and use of data; 5) Perform the analysis of datasets provided using the CIDACS data analysis environment, a safe and secure infrastructure that provides remote access to deidentified/anonymized datasets and analysis tools" (M. L. Barreto et al., 2019). Many of the problems we experienced in Uruguay (see results section) are considered. Importantly, the researchers are not allowed to share, have and much less to "own" the analyzed datasets, which are managed exclusively within CIDACS platforms. Therefore, CIDACS is strongly preventing data misuse and person re-identification. In addition, ethic approval, strong justification is needed and the data analysis processes actually performed can be monitored, ensuring that they are in agreement with the proposed objectives.

A HRI able to integrate administrative, clinical and genomic data is probably necessary for testing complex health research hypotheses (i.e. expressed in causal models, connecting entities at several levels of social and material organization) and to produce some science advancements that can be translated into healthcare. The already achieved interoperability within the Uruguayan health research system could be expanded to national registries and a future HRI. If that was the plan, there would be ample space for integrative and interdisciplinary work, a direction towards which CIDACS is advancing. This could be more relevant if effective interventions are more probable as a consequence of embracing and

understanding part of that complexity, both at the academic and government levels, eventually producing better health, more happiness and less inequity.

The European scientific system has organized itself to deliver policies and strategies for the creation, development and governance of a network of research infrastructures since the turn of the century. This process has been fueled since 2006 by the European Strategy Forum on Research Infrastructures (ESFRI). Some ESFRI-ERA RI network-supported health research projects involving human participants are particularly notorious because of their social control. Human neurobiological recordings using electrodes inserted in the brains of patients who awaited surgery for other causes, human brain functional MRI images and artificial intelligence are being combined in the Human Brain Project (HBP) (Savage, 2019). Some of the key questions the HBP asks are how human brain-artificial intelligence interaction can help us to understand what consciousness is and to build intelligent robots to assist people with disabilities. HBP researchers are aware that knowledge and technology produced in these respects are dual-use research of concern (DURC): can be used for right or wrong, including dangerous applications (Savage, 2019). Notoriously, the ethical, political and societal implications of the possible new understandings of the brain and the mind are acknowledged by the researchers, who have given the project a particular regulatory frame of reference: the Responsible Research and Innovation framework (Owen, Macnaghten, & Stilgoe, 2012). One aspect of this "responsibility by design" is that researchers must be transparent about research progress and any one can raise and register an ethical concern against the project on that basis, a concern with which researchers will deal with transparency in a process that involves an ombudsperson and the research ethics committee in charge of the project evaluation (Project, 2021). Beyond this particular framework, we believe that these and other strategies that allow "responsibility by design" and participation from the community should be considered when planning and running an HRI.

The global picture of the open science affairs and their connection to health research trends can be glimpsed to read the "temperature" and some potential directions of change. The United Nations Educational, Scientific and Cultural

Organization (UNESCO) has been promoting open science for decades. UNESCO released a global recommendation on open science (UNESCO, 2021), which was approved by its 41st assembly for adoption as a normative instrument. The document contains relevant statements we would like to mention here. UNESCO recognizes "the urgency of addressing complex and interconnected environmental, social and economic challenges for the people and the planet, including poverty, health issues, access to education, rising inequalities and disparities of opportunity, increasing science, technology and innovation gaps, natural resource depletion, loss of biodiversity, land degradation, climate change, natural and human-made disasters, spiraling conflicts and related humanitarian crises" and UNESCO also acknowledges "the vital importance of science, technology and innovation (STI) to respond to these challenges by providing solutions to improve human well-being, advance environmental sustainability and respect for the planet's biological and cultural diversity, foster sustainable social and economic development and promote democracy and peace" (UNESCO, 2021). Therefore, UNESCO establishes a research agenda for a global open science, focusing the research objectives on problems that pose "complex and interconnected environmental, social and economic challenges" (i.e. health problems) (UNESCO, 2021). This recommendation might influence financing for research projects in Uruguay and in underdeveloped regions, but will also require a shift away from the politically uninvolved science (scientists, organizations and institutions), and to be open to the study and reflection about the always present science-policy boundary (Turnhout, Stuiver, Klostermann, Harms, & Leeuwis, 2013; Wesselink & Hoppe, 2020). The document also urges countries and communities to debate on a series of topics, including the infrastructural changes needed for fostering OS (UNESCO, 2021). Several authors and boards conclude that Open Science is a multiplicity in itself, a complex globalized program still under preparation (Babini & Rovelli, 2020; Fecher & Friesike, 2014; E. Méndez, 2021; Mendez et al., 2020). Beyond open science, health research and science in general are undergoing a digital transformation. The research agenda proposed by UNESCO for Open Science should not be considered completely dependent on the progress or only proper for Open Science. We reason that solving ethical, social justice, and technological problems related to traditional health research that are being

aggravated in a context of fast changes will also make the path to Open Science easier and safer.

A first question we feel compelled to ask is: In which ways can open science and health research be combined to learn from each other? We have aimed to problematize the combination of open science and health research, chiefly considering their axiology and objectives but also the problems they bring to our society and the opportunities for social control that could be created there. It is generally agreed that bioethical reflection on moral problems and the values linked with health research with human participants is essential and this could and should be extended to other types of developments and social modalities of response to problems (i.e. OS, RIs). Two questions that we wish to consider first are from which perspective (paradigm and theoretical assumptions) can we approach the discussion and which are the hegemonic positions that are translated from the evaluative point of view? A partial look would be problematic, as it could lead us to focus only on aspects related to risks, for example, which must be included but should not be the only aspect of concern. According to some authors, open science necessarily brings increased risks to health research, to the point that we have to choose between OS and privacy (Dennis et al., 2019). Two particularly relevant risks, already discussed, are those of re-identification due to data linkage that allows databases to be combined (a process that is not prevented by anonymization) (Ohm, 2009) and that of manipulation by ill-intentioned agents with the potential to harm vulnerable people and/or groups (Wood, 2014). Considering the just mentioned possibilities, it has been concluded that research ethics regulations as well as research infrastructures available in different European countries have proven insufficient or ineffective (Lecuona Ramírez & Observatori de Bioètica i Dret, 2021).

Uruguay is lacking RIs with a capacity comparable to that of CIDACS'. The Centro Nacional de Supercomputación (ClusterUY) provides the highest computing power available in Uruguay and is a relevant antecedent for an HRI. ClusterUY must be differentiated from an OSHRI. ClusterUY is a high performance computing platform, which emerged as an academic initiative to provide services to solve complex problems in Uruguay (Nesmachnow &

Iturriaga, 2019). The main precedent in Uruguay, and motivation for the development of ClusterUY, is ClusterFING, a project that successfully initiated the development of the paradigm of use and centralized management of computing resources. ClusterFING was a High-Performance Computing (HPC) infrastructure of the Facultad de Ingeniería (Universidad de la República). Its initial infrastructure was acquired in 2008, funded by the Comisión Sectorial de Investigación Científica (CSIC), Universidad de la República, Uruguay. ClusterUY then arose, scaling the dimension of the computing infrastructure and extending the service to the national level (Nesmachnow & Iturriaga, 2019). The initial investment for the acquisition of the computational equipment was financed by the Agencia Nacional de Investigación e Innovación, through a call launched in 2012, and the counterpart has been provided by CSIC. The project is currently operating in a technically and financially self-sustaining manner. In this way, ClusterUY positions itself competitively with similar infrastructures in Latin America, whose services are nowadays accessible to researchers, scientists and technicians of the country, with remote access from anywhere in Uruguay, operating at different levels according to the needs of each problem (Gitler, Gomes, & Nesmachnow, 2020). While ClusterUY is excellently serving a big research and innovation community, it is centered around basic science and technological problems, and adequate solutions to problems like critical and responsible research, person/group respect and protection, integral project independent evaluation, high personal data security, data linkage, interdisciplinary health research environment are not an integral part of it, which are needs for health knowledge creation, as performed at CIDACS.

There have always existed tensions between social order, scientific inquiry and development (Atkinson, 1978; Nafziger, 2007). Social control (different definitions exist) is a social behavior identified with “how people define and respond to deviant behavior”, and how we eventually convert ourselves to conformity (Black, 2014; Janowitz, 1975). Social control of research aims at preventing issues related to legitimacy, relevance, responsibility, validity, fairness in participant selection, fair risk exposure, respect of the individual, etc. and at promoting wise individual and collective behavior, tending to the conservation of

the social order or to beneficial changes in it. Bioethics is "a discipline of applied ethics and a particular way of ethical reasoning" (Gordon, 2021). Wikipedia defines it as "the study of the ethical issues emerging from advances in biology, medicine and technologies." Hottois explains that "Bioethics covers a set of research, discourses and practices, generally multidisciplinary and pluralistic, which aim to clarify and, if possible, solve ethical questions raised by biomedical and biotechnological research and development within the societies characterized, to varying degrees, by being individualistic, multicultural, and evolutionary" (Hottois, 2004). Bioethics certainly contributes to improved social control of scientific research, healthcare and public health. The ethical evaluation and monitoring of research projects by institutional review boards is a key aspect in HR. This last aspect (project execution monitoring) is only partly defined in the Uruguayan regulation and has implementation problems. In addition, data privacy, safety and security evaluation boards do not take part in the evaluation and monitoring of most health research projects. Therefore, to improve science, OS not only requires new research infrastructures (RIs) but also revised structures for social and technological control of research.

VI. Concluding remarks

This work, rooted at and inextricable from the reported authors' experiences can be understood as part of a trans-disciplinary process or methodology with similarity to critical and constructive design research, were the "researchers becomes a change agent who is collaboratively developing structures intended to critique and support the transformation of the communities being studied" and where "deep relationships between researchers and research participants" are highly valuable (Barab, Thomas, Dodge, Squire, & Newell, 2004; Bardzell, Bardzell, Forlizzi, Zimmerman, & Antanitis). Accordingly, although a new health research infrastructure (HRI) for Uruguay has not yet been proposed as far as we know, we conclude that the need, the characteristics of it and the problems it will produce can and should be thoroughly and openly evaluated. We think that Uruguay would benefit from critically designing a HRI, addressing current and projected (i.e. in relation to OS) necessary conditions as well as requirements under future uncertain scientific human and health research scenarios. We also

conclude that further deliberations are needed before commitment agreements are signed within or between institutions on the way of creating an OSHRI. This article hopefully will help us share with other researchers, academics and stakeholders our current humble perspectives and experiences on this subject.

Considering the preceding statements, we think one key principle we can all accept is to promote and expand external and independent *ex-ante* evaluation by peers and non-peers of strongly justified health research proposals. Non-peer evaluation of research and technological projects is something that needs to be taken more seriously by the academic and non-academic community and by the national government. Strong justification would mean that data-centric projects must express the proposed or expected causal connections between research procedures and outputs with potential health improvements as well as with participant's and non-participants dignity, privacy and safety, because that is the general multiple objective of health research. Respecting the dignity, preserving the privacy and protecting the safety at person, group and whole society level requires actions both individual and collective. Individual action for health researchers includes assuming bioethics and biopolitics as a general basis for research and science evaluation and education. Collective actions include, in our opinion, a) planning the best possible OSHRI (one that: is oriented towards closing the inequity gaps concerning access to the benefits of health research and healthcare, ensures the best possible data quality, best possible technologies, gives strong justification for actions, promotes peer plus non-peer evaluation and bioethical and biopolitical reflection, practices socially acceptable and responsible by design health research); b) planning the best possible preparation for responding to a future pandemic or any other health or human crisis (one that includes bioethics and biopolitics among other aspects that needs to be addressed and are not addressed herein); and c) planning the best possible research evaluation system (one that: gives a better support for research ethics committees, is more transparent, effectively respond to health research requirements in times of emergencies, that improves education of scientists and professionals, and that ensures the best possible ex-ante evaluation, participant safety and data security monitoring).

During this work we held difficult deliberations that are still in progress. In them, the role of democracy, participation and respect for the individual and the relevance of the many forms of trust (in persons, institutions, law, democracy itself, etc.) needed for fruitful work and social life were really apparent. In this sense, it is important to recognize that the best design of a HRI will be useless at preventing its degradation and misuse if democracy is broken, institutions corrupted and ethics minimized.

VII. Acknowledgements

We thank Espacio Interdisciplinario (Universidad de la República) for sponsoring CIDASH (https://ei.udelar.edu.uy/grupos-financiados/sigla-acronimo/ciencias-de-la-informacion-y-datos-aplicadas-la-salud-humana). We acknowledge helpfull discussions with Susana Lorenzo Erro (Odontología social, Facultad de Odontología, Universidad de la República), Silvia Rivera (Facultad de Ciencias Sociales, UBA, Argentina), Alejandro Margetic and other members of PIDAE Project (Argentina), Mauricio Barreto (CIDACS, Brazil). We also acknowledge the collaboration and work of other CIDASH members, particularly of Cecilia Tomassini (Unidad Académica de CSIC) and Ignacio Ramírez (Instituto de Ingeniería Eléctrica, Facultad de Ingeniería). We appreciate the suggestions on English grammar made by Patricia Miranda (Department of Biophysics, Faculty of Medicine, UdelaR).

References

Aarons, D. (2018). Research in epidemic and emergency situations: A model for collaboration and expediting ethics review in two Caribbean countries. *Dev World Bioeth, 18*(4), 375-384. doi: 10.1111/dewb.12157

Anderson, Sheila. (2013). What are Research Infrastructures? *International Journal of Humanities and Arts Computing, 7*(1-2), 4-23. doi: 10.3366/ijhac.2013.0078

Atkins, Daniel E, Droegemeier, Kelvin K, Feldman, Stuart I, Garcia-Molina, Hector, Klein, Michael L, Messerschmitt, David G, . . . Wright, Margaret H. (2003). Revolutionizing science and engineering through cyberinfrastructure. *Report of the National Science Foundation blue-ribbon advisory panel on cyberinfrastructure*, *1*.

Atkinson, Richard C. (1978). Rights and responsibilities in scientific research. *Bulletin of the Atomic Scientists, 34*(10), 10-14.

Babini, Dominique, & Rovelli, Laura. (2020). *Tendencias recientes en las políticas científicas de ciencia abierta y acceso abierto en Iberoamérica*. CLACSO-Fundación Carolina.

Barab, Sasha A., Thomas, Michael K., Dodge, Tyler, Squire, Kurt, & Newell, Markeda. (2004). Critical design ethnography: Designing for change. *Anthropology & Education Quarterly, 35*(2), 254-268.

Bardzell, Shaowen, Bardzell, Jeffrey, Forlizzi, Jodi, Zimmerman, John, & Antanitis, John. (2012). *Critical design and critical theory: the challenge of designing for provocation*.

Barreto, M. L., Ichihara, M. Y., Almeida, B. A., Barreto, M. E., Cabral, L., Fiaccone, R. L., . . . Smeeth, L. (2019). The Centre for Data and Knowledge Integration for Health (CIDACS): Linking Health and Social Data in Brazil. *Int J Popul Data Sci, 4*(2), 1140. doi: 10.23889/ijpds.v4i2.1140

Barreto, Mauricio L., Ichihara, Maria Yury, Pescarini, Julia M., Ali, M. Sanni, Borges, Gabriela L., Fiaccone, Rosemeire L., . . . Sena, Samila. (2021). Cohort profile: The 100 Million Brazilian Cohort. *International journal of epidemiology*.

Barreto, Maurício Lima, Almeida, Bethania de Araujo, & Doneda, Danilo. (2019). *Uso e proteção de dados pessoais na pesquisa científica*.

Beauvais, M. J. S., Knoppers, B. M., & Illes, J. (2021). A marathon, not a sprint - neuroimaging, Open Science and ethics. *Neuroimage, 236*, 118041. doi: 10.1016/j.neuroimage.2021.118041

Black, Donald J. (2014). *The social structure of right and wrong* (Rev. ed.). San Diego: Academic Press.

Borgman, C. L. (2019). Knowledge infrastructures in past, present, and future tense. UCLA Knowledge Infrastructure Workshop. Los Angeles, CA.: University of California - Los Angeles.

Breilh, Jaime. (2013). *Addressing the Challenges of Social Determination of Health: a View from the South*.

Cai, Yongyang, Judd, Kenneth L, & Lontzek, Thomas S. (2012). Open science is necessary. *Nature Climate Change, 2*(5), 299-299.

Carrigan, Coleen, Green, Madison W, & Rahman-Davies, Abibat. (2021). "The revolution will not be supervised": Consent and open secrets in data science. *Big Data & Society*, *8*(2), 20539517211035673.

CIDASH. (2019). *Ciencias de la Información y Datos Aplicadas a la Salud Humana*. Recuperado de https://ei.udelar.edu.uy/grupos-financiados/sigla-acronimo/ciencias-de-la-informacion-y-datos-aplicadas-la-salud-humana

CIOMS. (2009). *International ethical guidelines for epidemiological studies*. COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES.

Clark, Michelle M, Stark, Zornitza, Farnaes, Lauge, Tan, Tiong Y, White, Susan M, Dimmock, David, & Kingsmore, Stephen F. (2018). Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ genomic medicine, 3*(1), 1-10.

Cortina, Adela. (2000). *Ética mínima*. Tecnos Madrid.

Craik, E. M. (2015). *The 'Hippocratic' corpus : content and context*.

Crawford, D. C., & Sedor, J. R. (2021). Biobanks Linked to Electronic Health Records Accelerate Genomic Discovery. *J Am Soc Nephrol, 32*(8), 1828-1829. doi: 10.1681/ASN.2021060836

Cutler, David M. (2020). Early Returns from the Era of Precision Medicine. *JAMA, 323*(2), 109-110. doi: 10.1001/jama.2019.20659

Chá Ghiglia, María Mercedes. (2019). Historia clínica electrónica: herramienta para la continuidad de asistencia. *Revista Médica del Uruguay, 35*(3), 107-123.

Chan, Y., Tung, M., Garruss, A. S., Zaranek, S. W., Chan, Y. K., Lunshof, J. E., . . . Church, G. M. (2017). An unbiased index to quantify participant's phenotypic contribution to an open-access cohort. *Sci Rep, 7*, 46148. doi: 10.1038/srep46148

Daltro, Mônica Ramos, & de Faria, Anna Amélia. (2019). Relato de experiência: Uma narrativa científica na pós-modernidade. *Estudos e pesquisas em psicologia, 19*(1), 223-237.

Davies, Tim G, & Bawa, Zainab Ashraf. (2012). The promises and perils of open government data (OGD). *The Journal of Community Informatics, 8*(2), 1-8.

Deci, Edward L, & Ryan, Richard M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne, 49*(3), 182.

Dennis, Simon, Garrett, Paul, Yim, Hyungwook, Hamm, Jihun, Osth, Adam F, Sreekumar, Vishnu, & Stone, Ben. (2019). Privacy versus open science. *Behavior research methods, 51*(4), 1839-1848.

Diani, Mario. (1992). The Concept of Social Movement. *The Sociological Review, 40*(1), 1-25. doi: 10.1111/j.1467-954X.1992.tb02943.x

Directorate-General for Communication, European Commission. (2013). *Research Infrastructures, including e-Infrastructures - Horizon 2020 - European Commission*. Recuperado de

https://ec.europa.eu/programmes/horizon2020/en/h2020-section/research-infrastructures-including-e-infrastructures

Dupré, John, & Cartwright, Nancy. (1988). Probability and Causality: Why Hume and Indeterminism Don't Mix. *Noûs, 22*(4), 521-536. doi: 10.2307/2215455

Emanuel, Ezekiel. (1999). *¿Qué hace que la investigación clínica sea ética? Siete requisitos éticos. Investigación en sujetos humanos: experiencia internacional*. Santiago de Chile: Programa Regional de Bioética OPS/OMS, 1999, 33-46.

Emanuel, Ezekiel J, Grady, Christine C, Crouch, Robert A, Lie, Reidar K, Miller, Franklin G, & Wendler, David D. (2008). *The Oxford textbook of clinical research ethics*. Oxford University Press.

Fecher, Benedikt, & Friesike, Sascha. (2014). Open Science: One Term, Five Schools of Thought. In Sönke Bartling & Sascha Friesike (Eds.), *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing* (pp. 17-47). Cham: Springer International Publishing.

Ferretti, A., Ienca, M., Hurst, S., & Vayena, E. (2020). Big Data, Biomedical Research, and Ethics Review: New Challenges for IRBs. *Ethics Hum Res, 42*(5), 17-28. doi: 10.1002/eahr.500065

Florio, Massimo. (2019). *Investing in science: social cost-benefit analysis of research infrastructures*.

Force11. (2014). Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing version b1. 0. FORCE11.

Foucault, Michel. (2000). *Vigilar y castigar: nacimiento de la prisión*. Siglo xxi.

Foucault, Michel. (2001). *Historia de la sexualidad (Vol. 3)*. siglo XXI.

Foucault, Michel. (2007). *El nacimiento de la biopolítica. Curso en el Collège de France de 1977-1978* (Trad. H. Pons). Buenos Aires: Fondo de Cultura Económica.

Foucault, Michel. (2008). *Seguridad, territorio, población* (Vol. 265). Ediciones Akal.

Franks, Bill. (2020). *97 Things About Ethics Everyone in Data Science Should Know*. O'Reilly Media.

Friesen, P., Douglas-Jones, R., Marks, M., Pierce, R., Fletcher, K., Mishra, A., . . . Sallamuddin, T. (2021). Governing AI-Driven Health Research: Are IRBs Up to the Task? *Ethics Hum Res, 43*(2), 35-42. doi: 10.1002/eahr.500085

Garzón, Fernando Mañé. (2010). *Historia de la genética clínica en el Uruguay*. Departamento de Historia de la Medicina, Facultad de Medicina.

Gitler, Isidoro, Gomes, Antônio Tadeu A, & Nesmachnow, Sergio. (2020). The Latin American supercomputing ecosystem for science. *Communications of the ACM, 63*(11), 66-71.

Gleason, Nancy W. *Higher education in the era of the fourth industrial revolution* (1st edition. ed.).

Gordon, J-S. (2021). *Bioethics*. Recuperado de https://iep.utm.edu/bioethic/

Grebe, T. A., Khushf, G., Chen, M., Bailey, D., Brenman, L. M., Williams, M. S., . . . Legal Issues, Committee. (2020). The interface of genomic information with the electronic health record: a point to consider statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med, 22*(9), 1431-1436. doi: 10.1038/s41436-020-0841-2

Gronde, T. V., Uyl-de Groot, C. A., & Pieters, T. (2017). Addressing the challenge of high-priced prescription drugs in the era of precision medicine: A systematic review of drug life cycles, therapeutic drug markets and regulatory frameworks. *PLoS One, 12*(8), e0182613. doi: 10.1371/journal.pone.0182613

Hagendorff, Thilo. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines, 30*(1), 99-120. doi: 10.1007/s11023-020-09517-8

Harron, K., Dibben, C., Boyd, J., Hjern, A., Azimaee, M., Barreto, M. L., & Goldstein, H. (2017). Challenges in administrative data linkage for research. *Big Data Soc, 4*(2), 2053951717745678. doi: 10.1177/2053951717745678

Hausman, Daniel M. (2007). Third-party risks in research: should IRBs address them? *IRB: Ethics & Human Research, 29*(3), 1-5.

Hebbring, S. (2019). Genomic and Phenomic Research in the 21st Century. *Trends Genet, 35*(1), 29-41. doi: 10.1016/j.tig.2018.09.007

Herington, J., & Tanona, S. (2020). The Social Risks of Science. *Hastings Cent Rep, 50*(6), 27-38. doi: 10.1002/hast.1196

Herzog, L. M., Norheim, O. F., Emanuel, E. J., & McCoy, M. S. (2021). Covax must go beyond proportional allocation of covid vaccines to ensure fair and equitable access. *BMJ, 372*, m4853. doi: 10.1136/bmj.m4853

Hill-Briggs, Felicia, Adler, Nancy E, Berkowitz, Seth A, Chin, Marshall H, Gary-Webb, Tiffany L, Navas-Acien, Ana, . . . Haire-Joshu, Debra. (2021). Social determinants of health and diabetes: a scientific review. *Diabetes Care, 44*(1), 258-279.

Hoeyer, Klaus. (2019). Data as promise: Reconfiguring Danish public health through personalized medicine. *Social studies of science, 49*(4), 531-555.

Holzmeyer, C. (2019). Open science initiatives: challenges for public health promotion. *Health Promot Int, 34*(3), 624-633. doi: 10.1093/heapro/day002

Hottois, Gilbert. (2004). *Qu'est-Ce Que la Bioéthique?*. Vrin.

Hummel, P., & Braun, M. (2020). Just data? Solidarity and justice in data-driven medicine. *Life Sci Soc Policy, 16*(1), 8. doi: 10.1186/s40504-020-00101-7

Janowitz, Morris. (1975). Sociological theory and social control. *American Journal of sociology, 81*(1), 82-108.

Jonas, Hans. (1985). *The imperative of responsibility: In search of an ethics for the technological age*. University of Chicago press.

Jones, K. H., Laurie, G., Stevens, L., Dobbs, C., Ford, D. V., & Lea, N. (2017). The other side of the coin: Harm due to the non-use of health-related data. *Int J Med Inform, 97*, 43-51. doi: 10.1016/j.ijmedinf.2016.09.010

Knight, H. E., Deeny, S. R., Dreyer, K., Engmann, J., Mackintosh, M., Raza, S., . . . Steventon, A. (2021). Challenging racism in the use of health data. *Lancet Digit Health, 3*(3), e144-e146. doi: 10.1016/S2589-7500(21)00019-4

Kottow, Miguel. (2012). Vulnerabilidad entre derechos humanos y bioética. Relaciones tormentosas, conflictos insolutos. *Derecho PUCP*(69), 25-44.

Lau-Min, Kelsey S., Asher, Stephanie Byers, Chen, Jessica, Domchek, Susan M., Feldman, Michael, Joffe, Steven, . . . Nathanson, Katherine L. (2021). Real-world integration of genomic data into the electronic health record: the PennChart Genomics Initiative. *Genetics in Medicine, 23*(4), 603-605. doi: 10.1038/s41436-020-01056-y

Lecuona Ramírez, Itziar de, & Observatori de Bioètica i Dret. (2021). *Pautes per avaluar projectes de recerca i innovació en salut que utilitzin tecnologies emergents i dades personals*. Recuperado de http://diposit.ub.edu/dspace/handle/2445/178612

Lemay, Margaret A. (2020). The Role of Expectations of Science in Shaping Research Policy: A Discursive Analysis of the Creation of Genome Canada. *Minerva, 58*(2), 235-260.

Leonelli, Sabina. (2016). *Data-centric biology: a philosophical study*.

Linder, J. E., Bastarache, L., Hughey, J. J., & Peterson, J. F. (2021). The Role of Electronic Health Records in Advancing Genomic Medicine. *Annu Rev Genomics Hum Genet, 22*, 219-238. doi: 10.1146/annurev-genom-121120-125204

Liu, Hong-Yan, Zhou, Liyuan, Zheng, Meng-Yue, Huang, Jia, Wan, Shu, Zhu, Aiying, . . . Li, Jia. (2019). Diagnostic and clinical utility of whole genome sequencing in a cohort of undiagnosed Chinese families with rare diseases. *Scientific reports, 9*(1), 1-11.

Lomas, Natasha. (2019). *Researchers spotlight the lie of 'anonymous' data*. TechCrunch.

Lundh, A., Lexchin, J., Mintzes, B., Schroll, J. B., & Bero, L. (2018). Industry sponsorship and research outcome: systematic review with meta-analysis. *Intensive Care Med, 44*(10), 1603-1612. doi: 10.1007/s00134-018-5293-7

Marmot, M. (2018). Inclusion health: addressing the causes of the causes. *Lancet, 391*(10117), 186-188. doi: 10.1016/S0140-6736(17)32848-9

Marmot, Michael, & Allen, Jessica J. (2014). *Social determinants of health equity*. American Public Health Association.

Martín-Sanchez, Fernando, Maojo, Victor, & Lopez-Campos, Guillermo. (2002). Integrating genomics into health information systems. *Methods of information in medicine, 41*(01), 25-30.

Masui, T., & Takada, Y. (2003). [Ethical, legal, and social issues of genome research--new phase of genome research desperately requires social understanding and safeguards on the use of medical records and other personal information]. *Yakugaku Zasshi, 123*(3), 107-119. doi: 10.1248/yakushi.123.107

Méndez, Eva. (2021). Open Science por defecto. *La nueva normalidad para la investigación. Arbor, 197*(799), a587-a587.

Mendez, Eva, Lawrence, Rebecca, MacCallum, Catriona J, Moar, Eva, Lossau, Norbert, Deketelaere, Kurt, . . . Garfinkel, Michele. (2020). *Progress on Open Science: Towards a Shared Research Knowledge System*. Final Report of the Open Science Policy Platform: European commission.

Méndez, Luciana, Pellegrino, Adela, Robaina, Sofía, & Vigorito, Andrea. (2019). *Primer censo de personas uruguayas e inmigrantes con título de doctorado: informe de resultados*. Documento de Trabajo. FCS-UM. PP; 3.

Méndez, S., & Botti, H. (2021). Desafíos de la datificación de la salud en Uruguay durante la pandemia de COVID-19. *Revista Médica del Uruguay, 37*(2).

Merton, Robert K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Moavenzadeh, John. (2015). *The 4th industrial revolution: Reshaping the future of production*. Paper presented at the World Economic Forum.

Molldrem, S., Hussain, M. I., & Smith, A. K. J. (2021). Open science, COVID-19, and the news: Exploring controversies in the circulation of early SARS-CoV-2 genomic epidemiology research. *Glob Public Health, 16*(8-9), 1468-1481. doi: 10.1080/17441692.2021.1896766

Morin, Edgar. (1999). Los siete saberes necesarios para la educación del futuro.

Moynihan, R., & Bero, L. (2017). Toward a Healthier Patient Voice: More Independence, Less Industry Funding. *JAMA Intern Med, 177*(3), 350-351. doi: 10.1001/jamainternmed.2016.9179

Ordenanza N° 648/008 La reglamentación sobre Investigación Clínica en Seres Humanos (2008).

Murray-Rust, Peter. (2008). Open Data in Science. *Nature Precedings*. doi: 10.1038/npre.2008.1526.1

Nafziger, E Wayne. (2007). From Seers to Sen: the meaning of economic development Advancing development (pp. 50-62): Springer.

Nagai, A., Hirata, M., Kamatani, Y., Muto, K., Matsuda, K., Kiyohara, Y., . . . Kubo, M. (2017). Overview of the BioBank Japan Project: Study design and profile. *J Epidemiol, 27*(3S), S2-S8. doi: 10.1016/j.je.2016.12.005

National Academies of Sciences, Engineering, & Medicine. (2017). *Communities in Action: Pathways to Health Equity*. Washington, DC: The National Academies Press.

Universal Declaration of Human Rights, General Assembly 27A C.F.R. (1948).

International Covenant on Economic, Social and Cultural Rights, 2200A (XXI) C.F.R. (1966).

United Nations Declaration on the Rights of Indigenous Peoples, A/RES/61/295 C.F.R. (2007).

Naya, H. (2021). *Fortalecimiento de las capacidades técnicas y humanas para las exportaciones de servicios genómicos*. Recuperado de http://urugenomes.org/

Nesmachnow, Sergio, & Iturriaga, Santiago. (2019). Cluster-UY: collaborative scientific high-performance computing in Uruguay. Paper presented at the International Conference on Supercomputing in Mexico.

Network, The Personal Genome Project: Global. (2021). *The Personal Genome Project*. Recuperado de https://www.personalgenomes.org/us

Neumayer, E., & Plümper, T. (2017). *Causal Complexity and the Limits to Inferential Validity Robustness Tests for Quantitative Research*. Cambridge University Press.

Nowakowski, M. (2016). Data Privacy: Anonymization and Re-identification. FRIDAY, MARCH 11, 2016 — 9:00 AM TO 9:45 AM EST: Information Systems & Technology.

O'Doherty, K. C., Shabani, M., Dove, E. S., Bentzen, H. B., Borry, P., Burgess, M. M., . . . Burke, W. (2021). Toward better governance of human genomic data. *Nat Genet, 53*(1), 2-8. doi: 10.1038/s41588-020-00742-6

Ohm, Paul. (2009). Broken promises of privacy: Responding to the surprising failure of anonymization. *Ucla L. Rev.*, 57, 1701.

Owen, Richard, Macnaghten, Phil, & Stilgoe, Jack. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy, 39*(6), 751-760. doi: 10.1093/scipol/scs093

Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. *J Glob Health, 9*(2), 010318. doi: 10.7189/jogh.09.020318

Pearl, Judea. (2009). *Causality*. Cambridge university press.

Pearl, Judea. (2021). Radical empiricism and machine learning research. *Journal of Causal Inference, 9*(1), 78-82.

Pellman, David, & Zhang, Cheng-Zhong. (2021). Decoding complex patterns of oncogene amplification. *Nature Genetics*. doi: 10.1038/s41588-021-00974-0

Phillips, K. A., Douglas, M. P., Wordsworth, S., Buchanan, J., & Marshall, D. A. (2021). Availability and funding of clinical genomic sequencing globally. *BMJ Glob Health, 6*(2). doi: 10.1136/bmjgh-2020-004415

Pidre, Silvana, González, Laura, Mendoza, Rodrigo, Piñatares, Mauricio, Granja, Nicolás, Serra, Flavia, & Ruggia, Raúl. (2018). Data quality management in e-health integration platforms: the case of Uruguay. *CLEI Electronic Journal, 21*(2), 8.

Project, Human Brain. (2021). Register an Ethical Concern. 2021. Recuperado de https://www.humanbrainproject.eu/en/social-ethical-reflective/about/register-ethical-concern/

Putzel, James. (2010). *Do no harm: international support for statebuilding*. Organisation for Economic Co-opertation and Development.

Raggio, V., Dell'Oca, N., Simoes, C., Tapie, A., Medici, C., Costa, G., . . . Spangenberg, L. (2021). Whole genome sequencing reveals a frameshift mutation and a large deletion in YY1AP1 in a girl with a panvascular artery disease. *Hum Genomics, 15*(1), 28. doi: 10.1186/s40246-021-00328-1

Randall, Gregory. (2021). To foster Open Science we need a new system to protect intellectual creation. *IAU Horizons, 26*(1), 44-45.

RedeCoVida. (2020). Rede CoVida – Ciência, Informação e Solidariedade. Diretrizes Organizacionais Funcionograma-Rede-CoVida. Brazil: https://redecovida.org/.

Rehm, Heidi L., Page, Ângela J. H., Smith, Lindsay, Adams, Jeremy B., Alterovitz, Gil, Babb, Lawrence J., . . . Birney, Ewan. (2021). GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell Genomics, 1*(2), 100029. doi: 10.1016/j.xgen.2021.100029

Rekettye, Gábor, & Pranjić, Goran. (2020). Price personalization in the Big Data and GDPR context. *Marketing & Menedzsment, 54*(3), 5-14.

Riolo, J. S. (2021). A Comparative Analysis of Genome Complexity and Manufacturability with Engineering Benchmarks. (Master's thesis), University of Cincinnati, OhioLINK Electronic Theses and Dissertations Center. Recuperado de http://rave.ohiolink.edu/etdc/view?acc_num=ucin1627663196868138

ROCHE. (2021). *Niños y adolescentes con cáncer acceden a tecnología de vanguardia en diagnóstico gracias a un acuerdo de la Fundación Pérez Scremini y Roche*.

Rodrigues, Mariana Vitti. (2020). Genomics, Big Data and Privacy: Reflections upon the implications of direct-to-consumer genetic testing. *Natureza Humana-Revista Internacional de Filosofia e Psicanálise, 22*(1), 21-21.

Sakaue, S., Kanai, M., Tanigawa, Y., Karjalainen, J., Kurki, M., Koshiba, S., . . . Okada, Y. (2021). A cross-population atlas of genetic associations for 220 human phenotypes. *Nat Genet, 53*(10), 1415-1424. doi: 10.1038/s41588-021-00931-x

Sanmarchi, Francesco, Toscano, Fabrizio, Fattorini, Mattia, Bucci, Andrea, & Golinelli, Davide. (2021). Distributed solutions for a reliable data-driven transformation of healthcare management and research. *Frontiers in Public Health, 9*, 944.

Savage, N. (2019). How AI and neuroscience drive each other forwards. *Nature, 571*(7766), S15-S17. doi: 10.1038/d41586-019-02212-4

Senier, Laura, Tan, Catherine, Smollin, Leandra, & Lee, Rachael. (2019). Understanding the potential of state-based public health genomics programs to mitigate disparities in access to clinical genetic services. *Genetics in Medicine, 21*(2), 373-381. doi: 10.1038/s41436-018-0056-y

Shah, P. D. (2021). Polygenic Risk Scores for Breast Cancer-Can They Deliver on the Promise of Precision Medicine? *JAMA Netw Open, 4*(8), e2119333. doi: 10.1001/jamanetworkopen.2021.19333

Shah, S. K., Kimmelman, J., Lyerly, A. D., Lynch, H. F., Miller, F. G., Palacios, R., . . . Zorrilla, C. (2018). Bystander risk, social value, and ethics of human research. *Science, 360*(6385), 158-159. doi: 10.1126/science.aaq0917

Song, Wenyu, Huang, Hailiang, Zhang, Cheng-Zhong, Bates, David W., & Wright, Adam. (2018). Using whole genome scores to compare three clinical phenotyping methods in complex diseases. *Scientific reports, 8*(1), 1-10.

Spangenberg, L., Fariello, M. I., Arce, D., Illanes, G., Greif, G., Shin, J. Y., . . . Naya, H. (2021). Indigenous Ancestry and Admixture in the Uruguayan Population. *Front Genet, 12*, 733195. doi: 10.3389/fgene.2021.733195

Spangenberg, L., Grana, M., Greif, G., Suarez-Rivero, J. M., Krysztal, K., Tapie, A., . . . Naya, H. (2016). 3697G>A in MT-ND1 is a causative mutation in mitochondrial disease. *Mitochondrion, 28*, 54-59. doi: 10.1016/j.mito.2016.03.006

Spangenberg, L., Grana, M., Mansilla, S., Martínez, J., Tapie, A., Greif, G., . . . Naya, H. (2019). Deep sequencing discovery of causal mtDNA mutations in a patient with unspecific neurological disease. *Mitochondrion, 46*, 337-344. doi: 10.1016/j.mito.2018.09.004

Stang, P. E., Ryan, P. B., Racoosin, J. A., Overhage, J. M., Hartzema, A. G., Reich, C., . . . Woodcock, J. (2010). Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. *Ann Intern Med, 153*(9), 600-606. doi: 10.7326/0003-4819-153-9-201011020-00010

Stewart, Craig A., Simms, Stephen, Plale, Beth, Link, Matthew, Hancock, David Y., & Fox, Geoffrey C. (2010). What is cyberinfrastructure. Paper presented at the Proceedings of the 38th annual ACM SIGUCCS fall conference: navigation and discovery, Norfolk, Virginia, USA. doi: 10.1145/1878335.1878347

Suen, Lee-Jen Wu, Huang, Hui-Man, & Lee, Hao-Hsien. (2014). A comparison of convenience sampling and purposive sampling. *Hu Li Za Zhi, 61*(3), 105.

Takayama, J., Tadaka, S., Yano, K., Katsuoka, F., Gocho, C., Funayama, T., . . . Tamiya, G. (2021). Construction and integration of three de novo Japanese human genome assemblies toward a population-specific reference. *Nat Commun, 12*(1), 226. doi: 10.1038/s41467-020-20146-8

Taylor, J., & Denegri, S. (2017). Industry links with patient organisations. *BMJ, 356*, j1251. doi: 10.1136/bmj.j1251

The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research Department of Health, Education, and Welfare. (1979). *THE BELMONT REPORT. Ethical Principles and Guidelines for the Protection of Human Subjects of Research*.

Tsey, K. (2008). The control factor: a neglected social determinant of health. *Lancet, 372*(9650), 1629. doi: 10.1016/S0140-6736(08)61678-5

Turnhout, Esther, Stuiver, Marian, Klostermann, Judith, Harms, Bette, & Leeuwis, Cees. (2013). New roles of science in society: Different repertoires of knowledge brokering. *Science and Public Policy, 40*(3), 354-365. doi: 10.1093/scipol/scs114

Turro, Ernest, Astle, William J., Megy, Karyn, Gräf, Stefan, Greene, Daniel, Shamardina, Olga, . . . Nihr BioResource for the, Genomes Project. (2020). Whole-genome sequencing of patients with rare diseases in a national health system. *Nature, 583*(7814), 96-102. doi: 10.1038/s41586-020-2434-2

UDELAR. (2020). *Una mirada a la política de Ciencia Abierta en América Latina*. Recuperado de https://udelar.edu.uy/portal/2020/10/una-mirada-a-la-politica-de-ciencia-abierta-en-america-latina/

UNESCO. (2005). *Universal Declaration on Bioethics and Human Rights*. Recuperado de http://portal.unesco.org/en/ev.php-URL_ID=31058&URL_DO=DO_TOPIC&URL_SECTION=201.html

UNESCO. (2021). *UNESCO Recommendation on Open Science* [34]. Recuperado de https://unesdoc.unesco.org/ark:/48223/pf0000379949?4=null&queryId=eae62776-60c4-4693-a40e-0651d5e311d2

UNICEF. (2005). *Operational guidelines for the establishment and functioning of data and safety monitoring boards*. World Health Organization.

Uruguay. (1958, octubre 29). *Ley 12549. Título: LEY ORGÁNICA DE LA UNIVERSIDAD DE LA REPÚBLICA*. Recuperado de http://www.impo.com.uy/bases/leyes/12549-1958

Uruguay. (2006, junio 2). *Decreto S/N/006. Título: NORMAS DE CONTROL, CALIDAD Y SEGURIDAD PARA EL TRANSPLANTE DE LAS CÉLULAS Y TEJIDOS HUMANOS*. Recuperado de https://www.gub.uy/presidencia/institucional/normativa/decreto-sn006-57

Uruguay. (2008a, agosto 18). *Ley 18331. Título: LEY DE PROTECCION DE DATOS PERSONALES*. Recuperado de https://www.impo.com.uy/bases/leyes/18331-2008

Uruguay. (2008b, agosto 26). *Ley 18335. Título: DERECHOS Y OBLIGACIONES DE PACIENTES Y USUARIOS DE LOS SERVICIOS DE SALUD*. Recuperado de https://www.impo.com.uy/bases/leyes/18335-2008

Uruguay. (2009a, setiembre 15). *Decreto 414/009. Título: REGLAMENTACION DE LA LEY 18.331, RELATIVO A LA PROTECCION DE DATOS PERSONALES*. Recuperado de https://www.impo.com.uy/bases/decretos/414-2009

Uruguay. (2009b, octubre 21). *Ley 18611. Título: PROCEDIMIENTOS PARA LA UTILIZACIÓN DE ANIMALES EN ACTIVIDADES DE EXPERIMENTACIÓN, DOCENCIA E INVESTIGACIÓN CIENTÍFICA*. Recuperado de https://www.impo.com.uy/bases/leyes/18611-2009/5

Uruguay. (2010, setiembre 16). *Decreto 274/010. Título: REGLAMENTACIÓN DE LA LEY N° 18.335 SOBRE DERECHOS Y OBLIGACIONES DE PACIENTES Y USUARIOS DE LOS SERVICIOS DE SALUD*. Recuperado de https://www.impo.com.uy/bases/decretos/274-2010

Uruguay. (2012, setiembre 21). *Ley 18968. Título: MODIFICACIÓN DE DISPOSICIONES RELATIVAS A LA DONACIÓN Y TRANSPLANTE DE*

CÉLULAS, ÓRGANOS Y TEJIDOS. Recuperado de https://www.impo.com.uy/bases/leyes/18968-2012/3

Uruguay. (2014, octubre 17). *Ley 19286. Título: APROBACIÓN DEL CÓDIGO DE ÉTICA MÉDICA*. Recuperado de https://www.impo.com.uy/bases/leyes/19286-2014

Uruguay. (2019, junio 12). *Decreto 158/019. Título: APROBACIÓN DEL PROYECTO ELABORADO POR LA COMISIÓN NACIONAL DE ÉTICA EN INVESTIGACIÓN VINCULADA A LA DIRECCIÓN GENERAL DE LA SALUD, REFERIDA A LA INVESTIGACIÓN EN SERES HUMANOS*. Recuperado de https://www.impo.com.uy/bases/decretos/158-2019

Uruguay. (2021, abril 27). Ley 19948. Título: APROBACIÓN DEL PROTOCOLO DE ENMIENDA DEL CONVENIO PARA LA PROTECCIÓN DE LAS PERSONAS CON RESPECTO AL TRATAMIENTO DE DATOS PERSONALES, SUSCRITO EN ESTRASBURGO. Recuperado de https://www.impo.com.uy/bases/leyes/19948-2021

van de Sandt, Stephanie, Dallmeier-Tiessen, Sünje, Lavasa, Artemis, & Petras, Vivien. (2019). The Definition of Reuse. *Data Science Journal, 18*(1).

Vayena, Effy, & Blasimme, Alessandro. (2021). Health Research with Big Data: Time for Systemic Oversight. *Journal of Law, Medicine & Ethics, 46*(1), 119-129. doi: 10.1177/1073110518766026

Vidal, S. (2017). Comités de ética de la investigación. Ministerio de Salud, Presidencia de la Nación (Argentina), 6.

Voisin, C., Linden, M., Dyke, S. O. M., Bowers, S. R., Alper, P., Barkley, M. P., . . . Nyronen, T. H. (2021). GA4GH Passport standard for digital identity and access permissions. *Cell Genom, 1*(2), None. doi: 10.1016/j.xgen.2021.100030

Wesselink, Anna, & Hoppe, Robert. (2020). Boundary Organizations: Intermediaries in Science–Policy Interactions: Oxford University Press.

Woelfle, Michael, Olliaro, Piero, & Todd, Matthew H. (2011). Open science is a research accelerator. *Nature Chemistry, 3*(10), 745-748. doi: 10.1038/nchem.1149

Wolpe, Paul Root. (2006). Reasons Scientists Avoid Thinking about Ethics. *Cell, 125*(6), 1023-1025. doi: https://doi.org/10.1016/j.cell.2006.06.001

Wood, Allen W. (2014). Coercion, manipulation, exploitation Manipulation: Theory and practice (pp. 17-50).

Yamamoto, N., Fujita, T., Kawashima, M., Wittig, J., Suzuki, M., & Kato, K. (2018). The inclusion of genomic data in the 2015 revision of Japan's Protection of Personal

Information Act: protection of wider range of genomic data as our next challenge. *J Hum Genet*, *63*(4), 537-538. doi: 10.1038/s10038-017-0409-4

Yang, Tung-Mou, Lo, Jin, & Shiang, Jing. (2015). To open or not to open? Determinants of open government data. *Journal of Information Science*, *41*(5), 596-612.

Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nat Biomed Eng*, *2*(10), 719-731. doi: 10.1038/s41551-018-0305-z

Declaration of conflicts of interests

The authors declare no conflicts of interest.

Editor's Note

This article was approved by Gregory Randall.

Authors' contributions

HB designed, assembled and wrote the several manuscript versions the authors worked on. HB interviewed all authors and designed an internal survey. HB, HN, SM, CS and VM contributed initial written texts, participated in direct group virtual exchanges and corrected the manuscript.