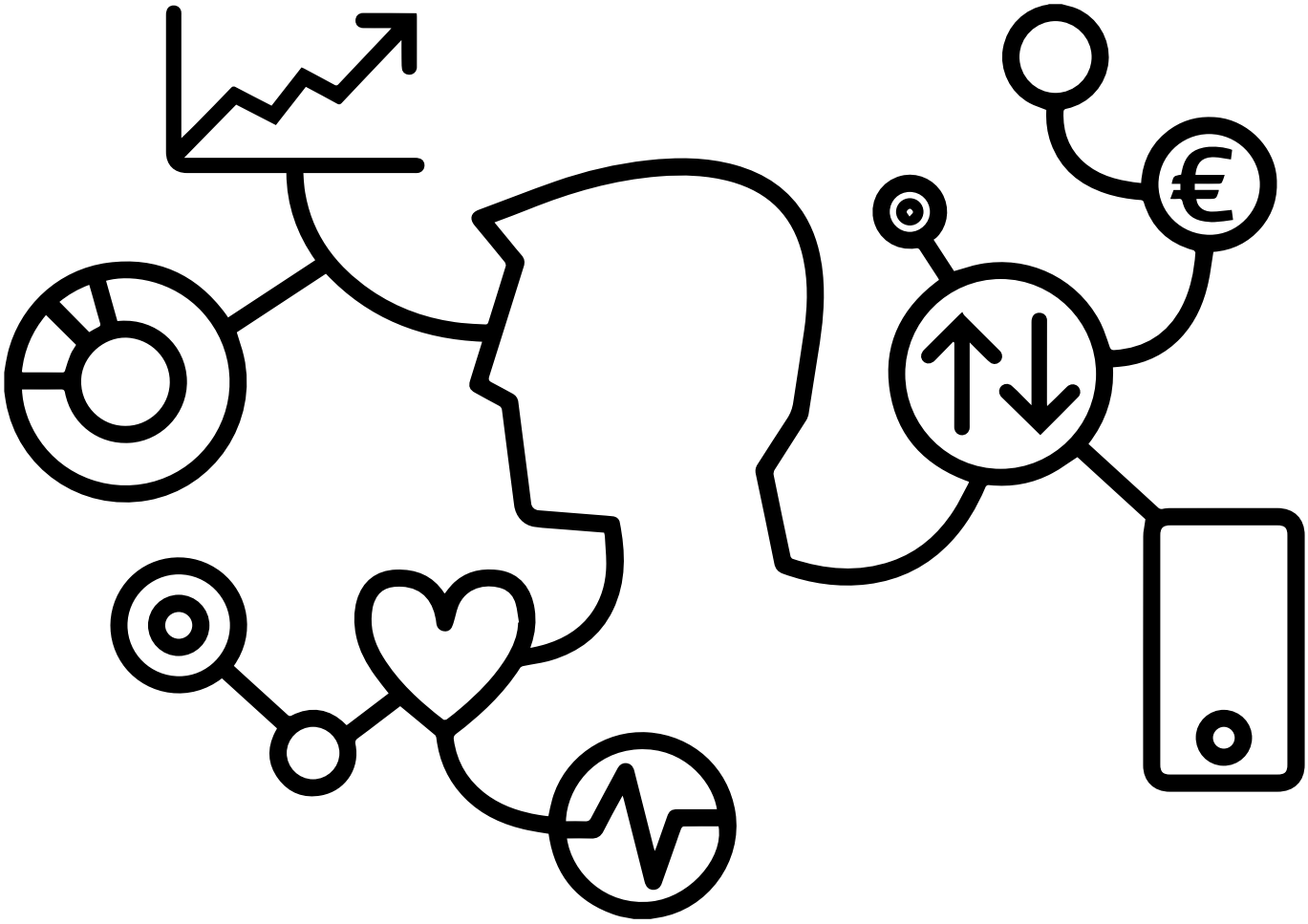
Fundación Innovación Bankinter

23 English version

- **In few years our production data has grown exponentially.** We need tools to analyze and process the relevant information: *Big Data.*
- **How are the implications of *Big Data* for us as individuals and relationships with our environment.**
- **How to improve our human performance though large volume of data process.**

Big data
The power of data

"The Big Data technology already is here, now to learn how to use it."



Fundación
Innovación
Bankinter

Big data
The power of data

Acknowledgements

Our gratitude to **Esther Paniagua,** journalist and author of this report. Without their collaboration were not possible the launch of this publication.

Our thanks to all of the Members of the XXIII Future Trends Forum (FTF) who made the forum a success, and a special thanks to those who made contributions to this publication:

For their invaluable support in the composition of this publication:
Chris Meyer
Lita Sands
Vicki Sayfert-Margolis
Stefan Bungart
J.C. Herz

For their invaluable role in the methodology and organization of the Future Trends Forum:
Chris Meyer

Garrick Jones
Clemens Hackl

Raquel Durán

Finally we would like to thank the members of the Foundacion Innovación Bankinter team for their commitment and follow through in the development of the content of this publication:

Sergio Martínez-Cava
María Teresa Jiménez
Marce Cancho
Lara García de Vinuesa
Dorsey Lockhart
Pablo Lancry

Trends Forum meeting

Ana María Arboleda
General Manager Atlantis Healthcare Spain

Miguel Arias
COO at CartoDB

Puneet Batra
Co Founder of LevelTrigger

Jessica Bland
Senior Researcher in Technology Futures of Nesta

Carla Brodley
Professor and Dean of the College of Computer and Information Science at Northeastern University

Stefan Bungart
Leader of GE Software Europe at GE Global Research

Antonio Damasio
David Dornsife Professor of Neuroscience, and Director of Brain and Creativity Institute at University of Southern California

Hanna Damasio
Dana Dornsife Professor of Neuroscience and Director and Professor of Psychology and Neurology University of Southern California

Maarten den Braber
Co-Founder of SHIFT

Julie Freeman
Artist, TED Senior Fellow, Head of Data as Culture Art Programme at the Open Data Institute (ODI)

J.C. Herz
Product designer and strategist engaged with Big Data

Eduardo Hruschka
Chief Data Scientist at Big Data Brazil

Richard Kivel
Entrepreneur & Investor: Technology, Bio/Pharma, Diagnostics and Healthcare & Bankinter Foundation Trustee

Joseph Kvedar
Director of Center for Connected Health

Gregory La Blanc
Lecturer Haas Economic Analysis at Berkeley University

Philip Lader
Non Executive Chairman of WPP Group & Bankinter Foundation Trustee

Julia Li
Founder & CEO of HCD Global, Chief Advisor for MIT Media Lab

Juan Carlos López
Head of Academic Relations & Collaboration at Roche Innovation Center

Chris Meyer
CEO of Nerve LLC and Trustee of the Bankinter Innovation Foundation

Emilio Méndez
Director, Center for Functional Nanomaterials at the U.S. Department of Energy's, Brookhaven National Laboratory, 1998 Prince of Asturias Prize & Bankinter Foundation Trustee

Esteban Moro
Researcher at Universidad Carlos III de Madrid

Iyad Rahwan
Associate Professor of Computing & Information Science at Masdar Institute

Pablo Rodríguez
Researcher and Director of Innovation at Telefónica

Lita Sands
Global Head of Digital Transformation at Novartis

Vicki Seyfert-Margolis
Founder & CEO of My Own Med, Inc

Eden Shochat
Founder of Aleph & Bankinter Foundation Trustee

Taryn Sullivan
CEO of Efficiency Exchange

Steve Trachtenberg
President Emeritus of George Washington University & Bankinter Foundation Trustee

Wilfried Vanhonacker
Coca Cola Professor of Marketing, Olayan School of Business, AUB & Bankinter Foundation Trustee

Andreu Veà
President, Internet Society & Member of Advisory Board to the Internet Hall of Fame

Thank you so much
Fundación Innovación Bankinter

The power of data

Our ability to produce information has evolved at the speed of light over the last few years. Because of the massive amount of available data, new tools to analyze and process data and identify relevant information have emerged. We have just started to seize the opportunities created by current technologies.

"If you stacked the new books being published next to each other you would have to move at 90 miles an hour just to keep up with the end of the line", wrote Stephen Hawking in The Universe in a Nutshell thirteen years ago already. "If the exponential growth continued, there would be ten papers a second in my kind of theoretical physics, and no time to read them", said the acclaimed physicist.

In fact, that is exactly what has happened. According to the professor of bioinformatics at the Harvard School of Public Health, Winston Hide, **in the last five years, more scientific data have been generated than in the entire history of mankind.** The sheer volume of existing data is so that if they were stored physically, they would take up more space than a galaxy, assures Hide.

This new, accelerated speed at which we generate contents has contributed to the emergence of a new science: big data. The name is self-explanatory. However, this concept in and of itself does not suffice to explain the science, which goes beyond the sheer concept of quantity. This is more than just massive data; it involves value creation through processing and analyzing data—that is **truly revolutionary of**

big data and the reason why thirty experts from around the world met at the XIII Future Trends Forum organized by the Bankinter Foundation of Innovation, the inspiration for this publication.

In the process leading up to this point, numerous devices and applications that measure, structure, process and analyze huge amounts of data have been developed thanks to new technological infrastructure designed to support petabytes of information. There are new, more polished technologies emerging constantly to improve organizational efficiency, accurately identify customers' tastes and needs, connect to governments and citizens, reinforce security and smart management in cities, carry out all sorts of forecasts and form useful **behavior patterns** for companies to gain deeper knowledge of their target public and raise awareness among individuals regarding their lifestyle, even recommending certain changes to habits in order to improve some aspect of their life & health.

This is just the beginning, we've barely started to harness the potential unleashed by big data analysis. Big data has opened the door to a better informed, more efficient society, capable of achievements that not so long ago belonged to the realm of science fiction.

Hawking may not have time to read all scientific papers written, but he will not need to. Some big data technology will do it for him, and he will just interpret the results. In fact, he will not have to wait much: thirteen years have gone by since his forecast, and this is part of the present. Data-driven **informed decision making** is a reality today, and ignoring it equals lagging behind.

The challenge now is to find the needle in the haystack, identify what is **relevant** and apply it while engaging all stakeholders (companies, organizations, governments, professionals and citizens), and overcoming legal, structural and human barriers along the way. No-one said it would be easy, but it sure is worth trying.

In the last five years, more scientific data have been generated than in the entire history of mankind.

Analyze huge amounts of data have been developed thanks to new technological infrastructure designed to support petabytes of information.

Find the needle in the haystack, identify what is relevant and apply it while engaging all stakeholders.

The Six Ws of Big Data in Business

What is the point of big data in business? What can it do for my company? How can I harness it to its utmost potential?

All news piece needs to answer at least six questions. Five of them start with W (Who, What, Where, When and Why) and the last one, How, also contains letter W. Those are the six Ws of journalism, although they are not exclusive to that field. We wanted to use the six Ws here to explain the vision of data science from a business perspective and answer these six questions:

'Who' is actually quite clear: **big data,** an industry worth $3.2 billion in 2010 and bound to hike up to $17 billion by 2015, according to IBM. It might be a controversial business at times, especially reputation-wise, when companies dig and trade in data generated by our personal lives online. But that is also why it is profitable: emerging start-ups like Splunk or companies like Acxiom have made $300 million and $1 billion in annual revenues each.

'What' is the actual use: **quantify, predict and change.** Quantify in order to tap into the knowledge stemming from the analysis of (otherwise immeasurable) massive data sets; ultimately to discover important facts, leaks we were missing, or behavior patterns of people linked to the organization (employees, clients, consumers...). And predict purchase or vote decisions based on them, or modify inefficient processes.

'Where?' We can apply it **comprehensively and across all departments** (HR, operations, finance, marketing...) within a company to attain several objectives: process and performance efficiency, leadership and talent training, business development, product & services research and innovation...

'How?' Innovating. By applying a broad, open vision that requires a shift in the corporate mindset. Embedding big data into an organization's DNA. Devising (and implementing) new approaches to consumers, new ways of generating value while developing a product or providing a service and creating new lines of business for this purpose; sharing the problems of the organization and its challenges to find crowdsourced solutions; going from the traditional iterative innovation to disruptive innovation.

'Why' and **'what for'** are also quite clear: to **optimize and transform** Data from the Internet, social media, phones and devices accurately measure our performance and use this knowledge to predict future performance. These measurements are useful to make more efficient business processes, but they often lead you a step further, to the transformation

of the organizational structure, making it leaner and more easily adaptable to market demands.

However, big data science cannot do this by itself. The 6 Ws are missing the **human factor,** which needs to be combined with the technology. On one hand, financial management, predictive and causal analysis and business intelligence tools are increasingly sophisticated and can be integrated into one single platform. But these applications will be useless if change is not implemented gradually and starts by actively listening to employees; in

Missing the human factor, which needs to be combined with the technology.

a constant process of measuring their needs and continuously strengthening the values and drivers of the transformation in progress.

Together in data

Big data connects us to the people around us and drives interaction at many levels —ranging from social to business—and in many fields—from health care to security.

Let's move from companies to individuals, to the **implications of big data for each and every one of us as individuals and our web of relationships:** between public officers and citizens, doctors and patients, executives and employees, coaches and players or law enforcement agents vs. criminals.

"Technology is accelerating the ability to listen to concerns. This is quite remarkable when applied to **the workings of cities.** The emerging participation platforms started by many governments around the world—be them local, regional or national platforms—enable citizens to express their needs regarding the place where they live and also notify incidences. In return, the administration is expected to react promptly. The problem does not lie in the ability to respond to the question, but rather, in its priority. The pressure of citizens might lead to inappropriate decisions and leave aside solutions to larger-scope problems that require more careful attention and meticulous, prolonged action.

The health care system is another vertical experiencing the arrival of big data. For the time being, it is merely an observer, as if this had nothing to do with it. But as it stands still, everything around it is in motion, and **relationships between patients, caregivers and healthcare professionals are**

Implications of big data for each and every one of us as individuals and our web of relationships.

changing inside and outside the system, in physical or online communities and within families. At an individual level, self-care is changing too, as we become aware of our lifestyle and how to improve it through applications and devices that measure our activity and healthcare markers. You may see it or not, but the healthcare system already is in a process of reinvention.

And so are corporations, where big data has arrived as the final incentive needed to renew mature hierarchical structures that do not satisfy the current needs of employees. This is the beginning of a **transition towards a new organization and relationship paradigm** between employees of an organization. Transparency is the key value, thanks to open data policies.

Employees demand more empowerment to make decisions, and big data gives it to them, which in turn leads to greater awareness of their responsibilities. The role of the executive power is being redefined, from bosses to leaders, from authority to facilitators, from controlling to nurturing talent.

The role of sport coaches is also changing. Computational science could be of use if you conceive the game not as a group of individuals but as a network where the better the connections (passes) between the points (players), the better the performance. Therefore, new strategies could be defined for the playing field by combining data analysis and the theory of networks. In fact, it is a reality already. The barrier is human, more so than technological: it requires an open mindset towards data and a **leveled relationship between players and coaches.** The work of the coaches will now be to "apply" results obtained from the data analysis into a strategy shared with the players. And most of the previous work will be undertaken by a new position: the data scientist, who will compile and clean the data.

This profile is currently scarce: it is an emerging profession and demand is not yet adjusted to the actual needs. The case of law enforcement agencies is quite obvious. Out of 49,000 policemen in New York, there are only 880 trained in data. This leaves them in a less than favorable position to use the latest tools available to prevent and detect crimes (online and on the street) and find the guilty party. Even if they have state-of-the-art technology, they lack skilled staff, which renders the technology useless.

This shortage of resources is in contrast with the economic abundance of criminal organizations, which have access to the same tools as the police. Add to this that they are not subject to legal barriers on privacy, data protection or exchange of information. This means that they can use big data without boundaries to develop smart, criminal super-organizations and feed their network and supply chain endlessly. Therefore, **the relationship between law enforcement agents and criminals is one of unfair disadvantage.**

Remaining challenges

Polishing, understanding and applying data while guaranteeing privacy are the three greatest challenges of big data.

Big data needs to retake three tests, and it might fail again if it doesn't act accordingly. The first test has to do with cleaning out subprime data. The second, with reading comprehension applied in practice. And the third, with ethics and accountability. In order to make sure these three subjects do not evolve from challenge to barrier and make us lag behind, we need to do our homework.

Contextualize and clean. That is the first data treatment, necessary to obtain a high-quality contribution instead of useless, simply declarative data. Algorithms must feed from human interaction, because a high heart rate while exercising is not the same as a high heart rate while chilling on the sofa. Nevertheless, useless, unstructured, wrong data will always make it through the first filter. They need to be detected and eliminated. The better the initial filter and selection, the easier the second run will be—while avoiding over-using the technical feasibility of unlimited data compilation that we actually don't need.

Understand and apply. That is the second task. Good reading comprehension of data is the key to appropriate communication. To this end, skilled data scientists and analysts that convey the key messages become essential. They may use existing visualization tools to facilitate big data to business professionals. The next step is applying the knowledge, what is to be done with the data. It is not as easy as it looks. Healthcare professionals, for one, face daily doubts regarding the use made of patient data (improved approach, learn about what works and what doesn't in the system...).

Balance. When compiling medical data, another question arises: Where to stop? This is linked to the third pending subject: the fine balance between property (and its limits), privacy and security. Who owns the personal data collected by governments, law enforcement agencies and other organizations? Where does their right to use them start and the right to protect one's information end? On one side of the scale there are the privacy of individuals and the security of their data, and on the other, their physical security when a law enforcement action could benefit from those data.

This is an obvious conflict of interest that we must solve in order to harness the potential of big data. There is no choice: we either do the homework or fail the test again.

The Promised Land

Past times were not always better. The future will be better, provided that the aspirations of big data come to fruition.

Facing the challenges and overcoming the barriers to big data can lead us to a fantasy world. A fully connected world ruled by transparency; with open data and free, collaborative tools; where companies, organizations and institutions, governments and individuals work together to fight poverty, healthcare, food and environmental problems; where collaborative work and open innovation lead to create and implement what today are considered science-fiction technologies.

This golden age could be attained if the **legal, human, market, format, access and skill-related barriers** stopping it today are overcome. The first page written in the history of the future includes a great global alliance to share data in the fight against the great scourges of humanity, enabled by global privacy regulation to standardize access to data. Compliance will be checked via machine learning algorithms.

In the medium term, data management tools will become widespread, and years later they will become available via mobile apps, gamified to the point that even kids will master them. At the same time, data will be open under a single international standard, and formats will be standardized—in medical records, for instance. But there will be a catastrophe amidst all this: a pirate cyber attack from a criminal super-organization will destroy the world's IT network and the Internet as we know it.

This will force a reconsideration of the protection and privacy paradigms and how to fight fraud and identity theft. We will move back to a conservative stance of social resistance that will ease up as transparency and systems that guarantee data protection evolve. Further, increasingly more people will seek training in data science or analysis, and **new academic disciplines will be created** inside and outside of the official education system, offline and online, and no college student will graduate without having studied big data science. Employment will be guaranteed—at least for a few years, until the excess supply leads the market to saturation.
Operations in cities will be self-managed through an algorithm, and companies will have a marketplace to buy and sell data. As the sensor technology progresses and the quality of data improves; the connected, driverless car might become a reality—something that will change how we see cars: as a service, rather than property. Additionally, there will be quantum computers at the service of data. And the volume issue will disappear, since all the data in the world will fit in a hard drive the size of a spoon, just as Hawking's universe fits in a nutshell.

This is the future before us, provided that the experts making predictions are right on track. Now it is time to go back to reality, to 2015, and hit the ground running so that in a few years we can revel in the golden age of big data.

The future of Big Data

The future of *Big Data* Time line

2015

01.Jan. > 01.Jan.
Big Data Analysis Competencies.
Competencies required for quality big data analysis are defined.

01.Jan. > 01.Jan.
"Fast-lane" Training for Big Data Analysis.
A few leading universities and business schools start fast-lane training to build talent for big data analytics.

01.Jan. > 01.Jan.
Data Protection Legislation Passed.
Data protection policies/ legislation implemented in all kinds of business.

2016

01.Jan. > 01.Jan.
New federated identity system.
A new federated identity system augments the Facebook graph.

01.Jan. > 01.Jan.
Agreement for use of big data to fight poverty, crime and other problems.
Governments, big companies, and UN agencies sign agreement to share their data to fight poverty, food supply problems, epidemics, and organised crime throughout the world.

01.Jan. > 01.Jan.
iTunes reveals what user info it collects.
Apple requires that every app in iTunes reveal what user information it collects and how it monetises that data, on a standard "about" screen on the iTunes Store Page.

2017

01.Jan. > 01.Jan.
Global Adoption of Open Data Certificates.
Open data certificates are globally adopted.

01.Jan. > 01.Jan.
Europe enacts data protection regulation.
Data protection regulation in Europe harmonies data access and privacy legislation which becomes adopted as the global gold standard.

01.Jan. > 01.Jan.
Leading university gains name as trainer of big data scientists.
One of the world's leading universities becomes well known for producing data analysis experts for e-commerce and gasification.

01.Ene. > 01.Ene.
Increased open-mindedness and ecosystem.
Individuals become more open-minded while cities and corporations build ecosystem.

2018

01.Jan. > 01.Jan.
Machine learning algorithms provide for privacy.
Machine learning algorithms become accountable for/ provide for privacy concerns.

01.Jan. > 01.Jan.
Realisation that big data is not what it is cracked up to be.
Realisation that big data is not what it is cracked up to be.

01.Jan. > 01.Jan.
Marketplace to buy and sell data.
A marketplace to buy and sell data is created that assesses the privacy impact of a trade and assesses algorithm lift of a trade.

01.Jan. > 01.Feb.
Human-data interaction design becomes part of curriculum.
Human-data interaction design education is integrated into every discipline of study, enabling more individuals to figure out how to create more value with data.

2019

01.Jan. > 01.Jan.
Application that identifies useful data.
IBM announces an application that allows for the extraction of 1% of useful information in any data set and automatically disposes of the useless data.

14.Mar. > 19.Apr.
Equal connectivity / internet access worldwide.

29.May. > 24.Jul.
Universal patient data / history format.
Universal patient data / history format - with tracking controls and interactive care (Alert / event driven)

11.Jul. > 17.Aug.
All college students have exposure to data science.

2020

24.Jan. > 27.Mar.
Congress fearful of the direction of big data.
Congress fearful of the direction that big data is going and seeing orwellian implications will regulate the use of information with an eye on application and privacy.

03.Feb. > 9.Mar.
Machine learning tools for personal use.
Machine learning tools for personal use with easy interface and accessible for everyone - closing the gap between data scientists and mainstream.

03.Feb. > 09.Mar.
Digital natives are in government leadership.

18.Feb. > 24.Mar.
Massive hacking terrorism brings down the web.
Massive hacking terrorism brings down the web and creates isolated cells of data / connectivity.

2021

09.Jan. > 15.Feb.
One single standard.
(as it happens with TCP).

18.Feb. > 27.Mar.
International agreement on healthcare data standards.
International agreement on the standards that must be observed to report healthcare / patient data.

14.Jun. > 20.Jul.
New education disciplines.
New education disciplines / paradigm across High schools and Universities.

14.Jul. > 20.Aug.
Remote learning credentials.

2022

13.May. > 19.Jun.
Legal guidance on what patient data must be shared by corporations.

25.Jul. > 31.Ago.
Joint big data education campaign by US/China/Europe.
Joint big data education campaign by US/China/Europe with open policy of work VISAs to enable exchange of talent & human resources.

2023

23.May. > 19.Jun.
Worldwide data privacy design principles.
Worldwide data privacy design principles are now embedded in laws in every country.

27.May. > 02.Jul.
Shift of Citizens and Compnaies.

08.Aug. > 09.Sep.
Big data predictions.
Big data predictions for climate change meet reality.

2024

11.Jan. > 16.Feb.
Connected Cars.
Driverless sensors; owned by service provider, not people; no need for insurance or parking.

2025

08.Jan. > 13.Feb.
Jobless data analysts.

09.Jan. > 24.Feb.
Quantum computer for data analysis.

10.Jan. > 15.Feb.
1st fully automated algorithmic municipal government.
1st fully automated algorithmic municipal government successfully runs a city for a whole month.

10.Jan. > 15.Feb.
New global strict regulations approved.
New global strict regulations on fraud & identity theft are approved.

18.Jan. > 24.Feb.
Mobile data mining and open collaborative creation tools.

29.Jan. > 03.Mar.
Tools enable everyone to analyse data.

14.Feb. > 17.Mar.
Data art is accepted by traditional art world and commands $$$.

Año

- **Undeveloped Markets milestones**
- **Data formats, Quality & Standards milestones**
- **Human factor & Resistance milestones**
- **Legislation & Legal Framework milestones**
- **Skills & Resources milestones**
- **Silos (Access, API) milestones**
- **Other milestones:**

Fundación Innovación Bankinter