



HACIA UNA NUEVA OLA EN LA REVOLUCIÓN DE LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES (TIC)



**Presidencia
de la Nación**

Ministerio de
Ciencia, Tecnología
e Innovación Productiva



Secretaría de
Planeamiento y Políticas

HACIA UNA NUEVA OLA EN LA REVOLUCIÓN DE LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES (TIC)

Consultor experto: Gabriel Baum.

El contenido de la presente publicación es responsabilidad de sus autores y no representa la posición u opinión del Ministerio de Ciencia, Tecnología e Innovación Productiva.

Ciudad Autónoma de Buenos Aires, diciembre de 2014.

Baum, Gabriel Alfredo

Hacia una nueva ola en la revolución de las tecnologías de la información y las comunicaciones / Gabriel Alfredo Baum ; con prólogo de Ruth Ladenheim. - 1a ed. - Buenos Aires : Ministerio de Ciencia, Tecnología e Innovación Productiva, 2015.

E-Book.

ISBN 978-987-1632-51-0

1. Nuevas tecnologías. 2. Informática. I. Ladenheim, Ruth, prolog. II. Título
CDD 005.3

Fecha de catalogación: 03/06/2015

AUTORIDADES

- Presidenta de la Nación
Dra. Cristina Fernández de Kirchner
- Ministro de Ciencia, Tecnología e Innovación Productiva
Dr. Lino Barañao
- Secretaria de Planeamiento y Políticas en Ciencia, Tecnología e Innovación Productiva
Dra. Ruth Ladenheim
- Subsecretario de Estudios y Prospectiva
Lic. Jorge Robbio
- Director Nacional de Estudios
Dr. Ing. Martín Villanueva

RECONOCIMIENTOS

La supervisión y revisión del trabajo estuvo a cargo del equipo técnico del Programa Nacional de Prospectiva Tecnológica (Programa Nacional PRONAPTEC) perteneciente a la Dirección Nacional de Estudios del Ministerio de Ciencia, Tecnología e Innovación Productiva:

- Lic. Alicia Recalde.
- Lic. Manuel Marí.
- Lic. Ricardo Carri.
- A.E. Adriana Sánchez Rico.

Por consultas y/o sugerencias, por favor dirigirse a pronaptec@mincyt.gob.ar

ÍNDICE

1. Prólogo	5
2. Introducción	6
3. Evaluación crítica del Libro blanco	8
4. Discusión del diagnóstico y prospectiva elaborado en el Proyecto MinCyT 2012-2014	10
4.1. Diagnóstico	10
4.2. El futuro de las TIC en las áreas de aplicación seleccionadas: desafíos	11
5. Hacia una sociedad del conocimiento en Argentina	18
6. Nuevas tendencias y disrupciones en las TIC	19
6.1. Computación móvil	19
6.2. Internet de las cosas	20
6.3. <i>Big Data</i>	20
6.4. <i>Cloud Computing</i>	20
6.5. Hacia una integración de estas tendencias y disrupciones	21
6.6. La era de los grandes datos: esperanzas, problemas y recomendaciones	23
7. Conclusiones y recomendaciones	31

1. PRÓLOGO

El documento que presentamos aquí constituye el último de una serie de estudios encargados por el Ministerio de Ciencia, Tecnología e Innovación Productiva en torno a este sector fundamental para el desarrollo científico y tecnológico del país y pretende alertar sobre los grandes desafíos que plantean los últimos avances registrados en el mismo. Ha sido elaborado por el Dr. Gabriel Baum, Director del Laboratorio de Investigación y Formación en Informática Avanzada (LIFIA) de la Universidad Nacional de La Plata, quien había dirigido los anteriores trabajos del Ministerio en la materia, el conocido Libro Blanco de la Prospectiva TIC, publicado en 2008, y el estudio prospectivo para un proyecto más reciente sobre el sector, llevado a cabo en el período 2012-2014.

En dichos estudios se había profundizado el análisis de lo que el autor llama las primeras olas de la revolución TIC: en primer lugar, la creciente potencia y proliferación de dispositivos (“tecnología portátil pequeña y poderosa al servicio de las personas”), y el desarrollo concomitante de la industria del *software*, que nuestro país vivió con gran intensidad en los primeros años del presente siglo; en segundo lugar, lo que el último estudio mencionado constató, la contribución del sector TIC observada en los últimos años a la creación de una sociedad del conocimiento en nuestro país.

En ese sentido, cabe destacar el desarrollo de una infraestructura de informatización y comunicaciones emprendida a partir de programas como Argentina Digital, Argentina conectada y EducAR, así como con las posibilidades de aplicación intensiva de las TIC que aparecen en el agro, en salud, en industria y en energía; en este último caso, especialmente tras la recuperación de YPF y las exploraciones iniciadas de yacimientos no convencionales de petróleo, grandes demandantes de desarrollos informáticos.

Ya en el segundo estudio de 2012-2014 se había empezado a constatar la aparición de nuevas tendencias en el universo TIC en todo el mundo, las que seguramente transformarán las formas actuales de hacer ciencia y de producir, lo que aquí se llama la nueva ola en la revolución TIC: estas ten-

dencias son las que el presente documento pretende profundizar.

Esta segunda ola estará centrada sobre todo en lo que se llama *Internet of Things* (la Internet de las cosas) y las *Big Data* (o ciencia de los grandes datos), ambos desarrollos estrechamente interconectados. Como dice el documento, “las *Big Data* entran en la escena para extraer conocimiento y valor de la inmensa masa de datos generada por las personas, pero también por una variadísima cantidad de dispositivos como sensores, satélites, radares, etc. (*Internet of Things*). La capacidad de las empresas, los Estados, los científicos y tecnólogos, y también la gente de a pie, para explotar estos datos y transformarlos en servicios útiles será una de las variables centrales del crecimiento y la competitividad en la próxima década”.

En el documento se describen las características de esta segunda ola de la revolución de las TIC y se recomiendan posibles cursos de acción para que el país pueda entrar con fuerza y capacidad de decisión en ella, a partir de objetivos de largo plazo y la creación de condiciones para la generación de nuevos modelos de negocios que promuevan el bienestar general y una sociedad basada en el conocimiento y la inclusión.

Este trabajo continúa la línea de focalización del Plan Nacional de Ciencia, Tecnología e Innovación, Argentina Innovadora 2020, que prioriza 35 núcleos socio productivos estratégicos en el marco de 6 sectores: agroindustria; ambiente y desarrollo sustentable; desarrollo social; energía; industria y salud. La focalización orientará las intervenciones en aquellos tópicos donde la ciencia y la tecnología han abierto nuevas oportunidades y donde las políticas públicas puedan generar transformaciones de mayor impacto al tiempo que garantizan la federalización de los resultados de la innovación.

Dra. Ruth Ladenheim

Secretaría de Planeamiento y Políticas
en Ciencia, Tecnología e Innovación Productiva

2. INTRODUCCIÓN

El presente estudio sobre el futuro de las Tecnologías de la Información y las Comunicaciones (TIC) tiene por objeto profundizar y enriquecer el proyecto desarrollado por el Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación (en adelante, el Ministerio) entre 2012 y 2014, denominado “Estudio de Consultoría Sector TIC” (en adelante, “Proyecto MinCyT 2012-2014”, realizado por la empresa Trends Consulting S.A.”). Este proyecto, a su vez, había tenido como uno de sus objetivos profundizar y ampliar las visiones que habían surgido del Libro Blanco de la Prospectiva TIC¹, elaborado también por el Ministerio en 2007-2008.

Entre otras cosas, el Proyecto MinCyT 2012-2014 se había propuesto además analizar la situación actual del sector y el futuro del mismo en sus tecnologías centrales y en cuatro áreas de aplicación (agroTIC, salud, energía e industria).

También inició el tratamiento de algunas nuevas tendencias, que en el último año han ido cobrando más fuerza en el mundo y también en Argentina. El presente estudio, pues, además de resumir y actualizar los resultados de los dos anteriores, se centrará en forma particular en el análisis más profundizado de estas nuevas tendencias, principalmente los temas de *Internet of Things* y las *Big Data*.

El Libro Blanco, publicado en 2009, había estado centrado principalmente en la industria del *software*, analizando sus tendencias y las necesidades de acciones futuras para su consolidación en Argentina.

En la revisión de los resultados del estudio realizado en el Proyecto MinCyT 2012-2014, se pudo constatar que se confirman sus previsiones generales. Pero en los años transcurridos desde entonces se han empezado a observar algunos cambios importantes, tanto en el orden global, como regional y nacional, que ameritaban una nueva discusión acerca del futuro de las TIC, y consecuentemente, de las estrategias, planes e iniciativas para continuar promoviendo su crecimiento en el país.

Asimismo, el Plan Argentina Innovadora del Ministerio -así como otras políticas públicas desarrolladas desde entonces- brindan un marco más claro

para pensar el futuro. Esto es lo que se comenzó a analizar en el Proyecto MinCyT 2012-2014 y se pone a discusión pública en el presente documento.

Dos temáticas centrales conforman, pues, el núcleo de las reflexiones de este documento:

1. Actualmente, la “revolución de las TIC” aparece en la superficie a través de la creciente potencia y proliferación de dispositivos (decíamos en el Libro Blanco, 2009: “tecnología portátil, pequeña y poderosa al servicio de las personas”), solamente las primeras olas de servicios digitales que simplificarán la vida de las personas y empresas han alcanzado la costa.

La “segunda ola” de esta nueva revolución de las TIC ya está a la vista y comenzando su etapa de difusión masiva. Los *Big Data* entran en la escena para extraer conocimiento y valor de la inmensa masa de datos generada por las personas, pero también por una variadísima cantidad de dispositivos como sensores, satélites, radares, etc.

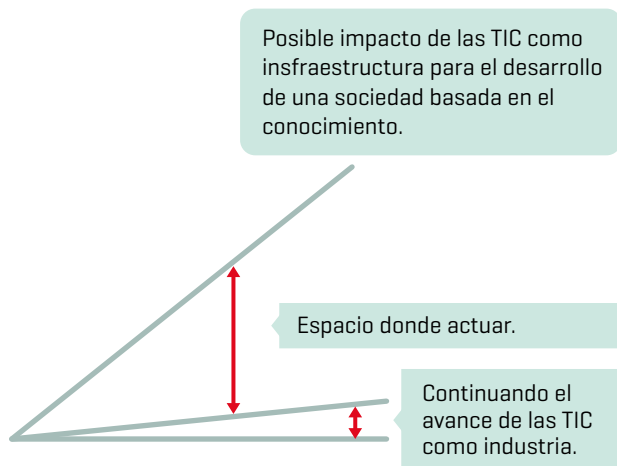
La capacidad de las empresas, los Estados, los científicos y tecnólogos, y también la gente de a pie, para explotar estos datos y transformarlos en servicios útiles será una de las variables centrales del crecimiento y la competitividad en la próxima década.

2. En función de esta visión, la estrategia de base para las TIC debería consistir en la generación de nuevos “negocios digitales” -en un sentido amplio, que abarca tanto al sector privado como público- a través de un programa que mejore las pre-condiciones para orientar estos nuevos negocios en diversos rumbos.

Las TIC, antes que una industria en sí misma, han demostrado ser una infraestructura fundamental para la generación de valor y bienestar social. Esta diferencia queda clara en un gráfico, inicialmente desarrollado por Dan Atkins en su trabajo con la *Office of Cyberinfrastructure de la National Science Foundation (NSF)*, que enfatiza dos espacios donde el Estado puede accionar: uno -en este caso, las TIC como sector industrial- que profundiza las

¹ El Libro Blanco de la Prospectiva TIC, en <http://bit.ly/1HVf7kE>

estrategias establecidas, y el otro -en este caso, TIC como infraestructura- que además expande el espacio donde situar estrategias de largo alcance y de mayor potencial para desarrollo e innovación:



Nuestro país debe apuntar a incrementar su capacidad de servicios para las industrias y el sector público a través del uso de soluciones digitales. Varios Ministerios (Planificación, Interior, Ciencia y Tecnología, etc.) y otros actores públicos (ANSES, AFIP) están fuertemente comprometidos en esta dirección -que deberían plasmar en un programa de desarrollo. Sin embargo, a diferencia de algunas posiciones más o menos extendidas, el Estado no debe apuntar a "encontrar un ganador", sino a crear prerrequisitos universales para el crecimiento.

Este tipo de disyuntivas son claves al momento de diseñar las políticas públicas que orientarán la difusión de esta nueva revolución de las TIC en la sociedad; libradas a las "fuerzas del mercado" o focalizadas en la búsqueda de "campeones", las TIC -y más aún las tecnologías de la actual etapa- son una poderosa fuerza de concentración de riqueza y profundización de las desigualdades.

Por otra parte, garantizando su accesibilidad y, más aún su utilización productiva, el Estado puede servirse de ellas para democratizar el crecimiento económico y las políticas sociales en beneficio toda la sociedad, y en particular de los más humildes.

A modo de ejemplo, como se comentará más adelante, es posible crear importantes infraestructuras de datos, aplicaciones y servicios (y aún de grandes datos) que pueden cambiar la escala de la agricultura de precisión a niveles hasta hace poco inimaginables, e intervenir en toda la cadena de valor de sus productos, aumentando sensiblemente su productividad y rentabilidad.

Sin embargo, si la introducción de estas nuevas tecnologías se deja librada al mercado, lo más probable es que contribuya a que la renta agraria se concentre más aún que actualmente.

Por otro lado, si el Estado promueve la aplicación de estas tecnologías para los pequeños productores y para la agricultura familiar, podría ayudar a mejorar sensiblemente las condiciones de producción y comercialización de sus productos.

Algunas políticas desarrolladas en la última década, han dado lugar al nacimiento de algunas historias exitosas: algunas empresas de *software* y servicios, desarrollos de videojuegos y servicios públicos. Por otro lado, han atraído en un período relativamente breve algunas actividades de empresas transnacionales, aunque muy poco de investigación y desarrollo.

El tema hacia el futuro es crear un programa para las TIC que ofrezca las mejores condiciones en vistas a la creación de negocios competitivos que generen bienestar. Comprender la naturaleza de la economía de Internet es una de las claves: las nuevas ideas y negocios se crean y difunden más rápido que nunca antes en la historia económica. Una consecuencia de esto es que resulta muy importante eliminar trabas que lo impidan y crear un ambiente abierto para que las ideas se concreten.

Una renovación exitosa requiere un método de operaciones que posibilite la co-existencia de objetivos de largo plazo e implementación rápida. Nuestro país ha tomado algunas medidas de base -desde la Ley de *Software* hasta el desarrollo de una nueva infraestructura de comunicaciones- pero en los próximos años deben decidirse nuevas acciones, tomando los resultados obtenidos y los cambios en el contexto real.

El presente documento se estructura del siguiente modo: en la sección 2 se presenta una revisión crítica y confirmación de los hallazgos del Libro Blanco de la Prospectiva TIC; en la sección 3 se presentan y discuten el diagnóstico y las recomendaciones surgidas del trabajo de prospectiva TIC realizado en el Proyecto MinCyT 2012-2014. En la sección 4 se discute el papel de las TIC en la conformación de una sociedad del conocimiento en Argentina, en relación con los importantes programas públicos de la última década. En la sección 5 se discuten las nuevas tendencias y disrupciones en las TIC, centrando el análisis en las tecnologías asociadas a los grandes datos, en un sentido amplio. Finalmente, en la sección 6 se presentan conclusiones y recomendaciones.

3. EVALUACIÓN CRÍTICA DEL LIBRO BLANCO

El Libro Blanco, elaborado -en 2007 y parte de 2008- colectivamente por un conjunto numeroso y relevante de empresarios, académicos y funcionarios públicos, analizó con profundidad varias áreas tecnológicas importantes, áreas de aplicación de las TIC, y áreas transversales, es decir, las problemáticas claves para lograr el desarrollo de las dos anteriores.

Sobre esa base, se logró identificar los temas críticos que debían ser abordados para promover el crecimiento de las TIC, tanto en sí mismas como también como un *enabler* para el desarrollo económico y social.

Como se indicó en el punto 1, gran parte de los problemas identificados, de las predicciones y de las recomendaciones formuladas se constataron como correctas. Sin embargo, es necesario reconocer que al menos dos fenómenos que se manifestarían poco después de la finalización del estudio prospectivo y que no fueron previstos, cambiaron sustancialmente el escenario, poniendo en cuestión algunas conclusiones y recomendaciones enunciadas en ese momento.

Por una parte, la crisis económica global, iniciada en Estados Unidos y luego expandida principalmente en Europa; por otra parte, las políticas públicas iniciadas por el gobierno nacional y otros, en la región sudamericana.

En efecto, en el Libro Blanco se señala expresamente que:

En las áreas de aplicación se puso foco en el análisis de los mercados. Por ejemplo en el caso de Servicios de IT se presentan las tres variables que hacen atractiva a la Argentina como país destino de la creciente tendencia mundial enfocada hacia el outsourcing-offshoring son: las capacidades de los proveedores, el costo y la comunicación. Se argumenta que para no perder el tren de la oportunidad de esta tendencia, que ya es marcada en el principal proveedor y comprador de servicios IT del mundo, Estados Unidos, las empresas argentinas del sector deben posicionarse en el lugar que están dejando las potencias "I" en tecnologías de la información: Irlanda, India e Israel. Estas ya están encontrando dificultades para satisfacer esta

demanda, sobre todo porque carecen de capital humano, lo que abre el juego a nuevos actores.

Esta misma perspectiva subyace en los fundamentos de la Ley de Promoción de la Industria del *Software* y sus planes asociados. La crisis global puso en cuestión la ventana de oportunidad por la cual debía ingresar la Argentina al escenario global. Si bien en Estados Unidos la industria IT no sufrió la crisis en la misma medida que otras ramas de la producción y los servicios, el mercado europeo se redujo a una expresión mínima y varios países del Este europeo y de Asia se transformaron en competidores directos en el área de los servicios IT.

En resumen, el acceso al mercado estadounidense se mantuvo más o menos estable, se redujo considerablemente el europeo y como se indica más abajo fue reemplazado por el latinoamericano como destino de las exportaciones nacionales de *Software* y Servicios Informáticos (SSI); la promesa del salto cualitativo basado en la exportación de servicios de valor agregado no se cumplió, al menos no en toda la dimensión prevista.

A partir de 2009, el gobierno nacional -en consonancia con gran parte de los sudamericanos- comenzó a desarrollar una serie de programas de infraestructura tecnológica e inclusión digital (más precisamente, social, basada en brindar acceso a la tecnología digital) a gran escala.

Comenzando con la TV Digital e incluyendo los planes Argentina Conectada, Conectar Igualdad, Igualdad Cultural, estas políticas públicas han establecido las bases estructurales para la construcción de una sociedad del conocimiento de alcance universal.

Junto con ellas, se produjo la recuperación de YPF para el patrimonio nacional, la reformulación de AR-SAT como actor central en las comunicaciones y otras acciones relevantes del Estado que cambiaron sustantivamente el escenario. Estas políticas tampoco fueron consideradas en la concepción del escenario futuro subyacente en el Libro Blanco; por cierto, eran muy difíciles de prever en ese momento.

En efecto, en el Libro Blanco, el mercado latinoamericano aparece como una oportunidad menor, y el mercado local fue considerado solamente en términos de algunos sectores productivos y de servicios; la capacidad de demanda del Estado se consideró más bien como un reclamo que como una posibilidad efectiva de desarrollo para el sector TIC.

En resumen, el nuevo escenario global, regional y local muestra los aciertos y las deficiencias del Libro Blanco y las nuevas tendencias tecnológicas que se han comenzado a establecer y dan un nuevo marco para formular planes y proyectos para las empresas nacionales, los centros de investigación, el sistema educativo y el propio sector público para desarrollar un sector TIC amplio y potente.

4. DISCUSIÓN DEL DIAGNÓSTICO Y PROSPECTIVA ELABORADO EN EL PROYECTO MINCYT 2012-2014

4.1. Diagnóstico

El Proyecto MinCyT 2012-2014 elaboró un diagnóstico de la situación actual de las TIC, teniendo en cuenta los cambios efectuados en la industria en el mundo y en nuestro país, a posteriori de la publicación del Libro Blanco de la Prospectiva TIC en 2008. A continuación, se resumen los puntos destacados:

- La crisis global implicó una creciente deslocalización de servicios de los centros de desarrollo y soporte de infraestructura de las empresas transnacionales. Los expertos identifican un impacto positivo por la liberación de recursos humanos, y otro negativo asociado al impacto en las empresas que trabajan para un solo cliente global.
- La falta de asociatividad entre las firmas aparece como un elemento de vacancia. Se señaló la importancia de los consorcios de exportación, aunque la conectividad es necesaria en un sinfín de áreas (sustitución de importaciones, agenda social, etc). La debilidad en la interacción entre el sector empresario y la academia es otra limitante fuerte para abordar proyectos complejos e innovadores.
- Las políticas públicas hacia el sector deben avanzar en la actualización de las leyes y los instrumentos tomando en cuenta las nuevas características de la industria. En materia de fondeo, la principal fuente de financiamiento de las empresas de TIC argentinas es la reinversión de utilidades y la escasez de mano de obra sigue al tope de las preocupaciones. En particular, es importante señalar:
 - Las políticas de promoción del sector (Ley de *Software*, Fondo Fiduciario de Promoción de la Industria del *Software* -FONSOF-, Fondo Tecnológico Argentino -FONTAR-, etc.) han tenido un impacto positivo, aunque se señalan algunas superposiciones entre las distintas dependencias públicas involucradas.
 - Formación de recursos humanos: es el factor que sigue apareciendo como una limitante al crecimiento, junto con la falta de vinculación universidad-empresa. Desde la publicación del Libro Blanco ha habido avances, como la promoción de la incorporación de investigadores y becarios al Consejo Nacional de Investigaciones Científicas y Técnicas -CONICET-, las becas en ciencias e ingenierías, las becas TIC, entre otras. Así también, la reciente iniciativa del Ministerio -adoptada por todo el sistema científico y tecnológico, y las universidades- para introducir cambios en la evaluación del personal científico y técnico.
 - Algunas iniciativas para promover vocaciones hacia la tecnología -y las TIC en particular- en el nivel medio y aún en el inicial (Dale Aceptar, Conectar Igualdad, Program.Ar, y algunas a nivel provincial) resultan positivas pero aún insuficientes, es necesario perseverar en las mismas pero fundamentalmente impulsar un cambio estructural en la educación inicial y media, comenzando por la jerarquización y formación de maestros y profesores.
 - El problema de la disponibilidad de recursos humanos es particularmente una fuente de limitación en las áreas de aplicación importantes como energía, salud e industria, algo menos en agroTIC.
- En relación al incremento de los salarios en moneda extranjera, las consultas muestran dos grupos diferenciados: unos que perciben al tipo de

cambio actual como una amenaza y otros que lo consideran aún competitivo para la actividad que desarrollan. Las políticas deben atender y dar cuenta de esas heterogeneidades.

- El rol del Estado como cliente de las empresas de SSI argentinas es aún poco significativo, apenas representa el 8% de las ventas totales cuando en economías similares tiene un peso del 20%.

- En relación al desarrollo de infraestructura pública aún nos encontramos en la etapa de construcción o despliegue de la infraestructura física. En este punto, es necesario diseñar una estrategia que posibilite la participación más activa de las empresas privadas nacionales de tecnología que se asocien con el Estado y sus empresas públicas que lideran los emprendimientos (ARSAT, INVAP, Educ.ar, RTA). Se trata de una inversión que está haciendo el Estado de unos 10.000 millones de pesos; es pues una oportunidad singular para impulsar el crecimiento de un sector empresario potente e innovador, que podría liderar el desarrollo de las TIC en la región sudamericana. En este punto es necesaria una discusión público-privada para que la inversión pública pueda apalancar un salto el volumen y la complejidad de los proyectos de las firmas nacionales. Consideraciones análogas pueden hacerse respecto de los centros de investigación y universidades. Un caso exitoso en este sentido ha sido el diseño de un *software* de una universidad nacional para la televisión abierta digital. En este sentido es importante señalar:

- El uso intensivo de TIC en agro -siembra, cosecha, control climático, etc- produce incrementos de rentabilidad modestos pero con un importante impacto sobre el volumen de ganancias a partir de la gran escala de producción.

- El uso de las TIC en agroindustria se verifica en un amplio conjunto de áreas. Su penetración, estado evolutivo e impacto es aún una cuestión pendiente de análisis.

- Actualmente el INTA no dispone de un área estratégica en TIC y no posee doctores en informática, electrónica o telecomunicaciones. Argentina no posee una sociedad de TIC para la agricultura, ni edita una revista científica y tan sólo hace 4 años presentó a la comunidad

el Congreso de agro-informática. Estas cuestiones marcan la necesidad de articular los esfuerzos dispersos que actualmente existen en el país. Esto contrasta con el ejemplo de Brasil, que desde hace 25 años tiene en su *Empresa Brasileira de Investigación Agropecuária (EMBRAPA)* un Instituto de Bioinformática, con alrededor de 100 técnicos.

- El cambio climático es el nuevo "tractor" de la I+D y puede aprovecharse para impulsar la inversión en infraestructura de datos (bioinformática). Existen déficits crónicos en información de clima, suelos, recursos naturales, contaminación, uso de la tierra, sistemas hidrológicos, etc.

- En lo referente a salud, la utilización de sistemas de información genera un conocimiento que ayuda a facilitar el control de las enfermedades crónicas y construir un sistema de promoción sanitaria que sea realmente beneficioso para la salud de la comunidad. Por otra parte, el desarrollo de TIC en salud impacta en la industria de las TIC al mejorar el perfil productivo y exportador del país. Si bien existen soluciones tecnológicas para los prestadores de servicio de uso universal, es bien sabido que las soluciones locales y regionales suelen adaptarse más fácilmente a los procesos clínicos locales.

- El proyecto considera prioritaria una política de "compre TIC nacional", e impulsar el desarrollo de mecanismos de fondeo para la adquisición de intangibles. También recomendaron impulsar una mayor integración con la actividad agrícola o de recursos hidro-carburíferos (petróleo y gas), como así también estimular el desarrollo de *software* embebido y la robótica en la producción industrial.

4.2. El futuro de las TIC en las áreas de aplicación seleccionadas: desafíos

El Proyecto MinCyT 2012-2014 que estamos comentando consultó a unos 300 expertos del área, para analizar el posible aporte futuro de las TIC en algunas de las áreas prioritarias del Ministerio, profundizando el análisis exploratorio desarrollado por el Libro Blanco de la Prospectiva TIC².

² La consulta fue organizada por las consultoras del proyecto Verónica Robert y Florencia Barletta, de la Universidad Nacional General Sarmiento.

A continuación se presentan los temas tecnológicos prioritarios que se señalaron por dicho proyecto en cada una de las cuatro áreas seleccionadas, se discuten y añaden algunos otros temas, y se señalan algunas líneas de acción a emprender:

Área de agroTIC

- Aplicación de las TIC a la agricultura de precisión. Se trata de un segmento de muy elevada importancia, por el tamaño del mercado local, posibilidad de sustitución de importaciones, las bajas barreras a la entrada y alto dinamismo de los mercados internacionales. Hay capacidades locales para desarrollarla.
- Trazabilidad, modelaje y sistemas de alertas tempranas: también hay buenas capacidades locales.
- Modelaje para cultivos y pestes más frecuentes.
- Con menor énfasis, sistemas de información agrometeorológica, y sistemas de alertas tempranas: hay oportunidades por el tamaño del mercado local y para el aprovechamiento de la infraestructura desarrollada por los programas de sociedad del conocimiento.

La Comisión Asesora de Expertos (CAE) del proyecto, integrada por referentes importantes a nivel nacional, consideró importantes todos estos temas, aunque dio particular atención a los dos últimos, mayor a la que le dieron los expertos consultados en la encuesta, en especial teniendo las oportunidades que ofrece el nuevo tema de las *Big Data*³. De hecho, todos los hallazgos de la CAE se redefinirán -de hecho ya están siendo reformulados- en el contexto de la nueva revolución de las TIC, en particular, lo que hemos llamado la segunda ola.

Un dato relevante para las agroTIC que marca la tendencia principal a nivel global, y que debería ser tomada en cuenta en función de la competitividad del país, es que los grandes jugadores mundiales en el tema ya están implementando lo que llaman "*data-driven agriculture*" o "*prescriptive agriculture*", y se están disputando el liderazgo en este asunto.

Monsanto y Dupont⁴, que representan más de la mitad del negocio mundial de semillas y granos, ya están implementando soluciones basadas en *Big Data* para, llevado al extremo, apropiarse de "toda" la información relevante del mundo. También los grandes jugadores globales de las TIC están trabajando en la misma dirección (IBM, SAP, CISCO, etc).

Argentina todavía tiene una ventana de oportunidad para entrar en ese juego. En ese aspecto, el trabajo de la Fundación Sadosky junto con AACREA (ver la sección de Conclusiones) resulta de gran importancia. Asimismo, la empresa FRONTEC creada en conjunto por el grupo Los Grobo e INVAP, dedicada a dar servicios a los productores -principalmente basados en imágenes satelitales de alta definición-, es un paso significativo en esa dirección.

A la vez, un conjunto de empresas PyME de TIC, grupos en el INTA y en las universidades desarrollan esfuerzos significativos para explotar datos en diversa escala. En la reciente "*Hackaton*" de *Agro-datos*, organizada por La Fundación Sadosky y la Maestría en Minería de Datos de la Universidad de Buenos Aires, con alrededor de 100 participantes de diferentes disciplinas se pudo observar el potencial de esta disciplina en el país.

En el mismo rumbo se encuentra la Plataforma de Bioinformática Argentina (BIA), una iniciativa en la que participan el CONICET, la empresa Indear y las universidades de San Martín, de Buenos Aires y Católica de Córdoba. La BIA, entre su variedad de investigaciones y servicios, brinda apoyo relevante a soluciones para el agro; entre ellas:

- Análisis de genomas de especies vegetales.
- Análisis de transcriptomas de especies vegetales.
- Metagenómica del suelo.

Es importante valorar el trabajo y la potencialidad de BIA, que está llamada a jugar un papel relevante en la investigación y el desarrollo tecnológico de las TIC en relación con el sector agropecuario.

En otro orden, de algún modo en el otro extremo

³ Por ejemplo, la representante de CONAE en la CAE recordó que el nuevo satélite que se lanzará en 2015 tendrá posibilidades de recoger una gran cantidad de información, por ej., sobre la humedad en los suelos.

⁴ <http://on.wsj.com/1cecdp>, <http://bloom.bg/1AgdcRm>

del sector agropecuario, las TIC pueden jugar un rol relevante en la promoción de los pequeños productores y de la agricultura familiar. La nueva infraestructura de telecomunicaciones (ver sección 4) abre una enorme oportunidad para la inclusión digital y el crecimiento económico de este sector, que representa alrededor del 80% de los productores, y es el principal proveedor de alimentos frescos del país.

El desarrollo servicios sobre la nueva red de fibra óptica y las nuevas comunicaciones satelitales pueden ayudar significativamente para mejorar las condiciones de comercialización, producción, educación y capacitación, salud de los productores familiares. Se trata de un amplísimo abanico de aplicaciones y servicios que también pueden ser una gran oportunidad para empresas nacionales y equipos de I+D, que podrían también ser exportados a los países de la región.

Finalmente, aunque relacionado también con lo señalado en los párrafos anteriores, debe mencionarse la importancia de la utilización de las TIC en la promoción de una producción agropecuaria sustentable, con especial foco en el cuidado del medio ambiente, los métodos y técnicas de irrigación, y la alimentación saludable de la población⁵.

En buena medida, este aspecto se relaciona con el anterior acerca de la promoción de la agricultura en pequeña escala y familiar, e igualmente puede abrir importantes oportunidades para el desarrollo de *software* y servicios para el mercado local y externo.

Acciones a emprender:

- Promover el crecimiento y la articulación de las empresas, grupos de investigación y organismos públicos en el sector agroTIC, particularmente en los temas vinculados a la explotación de datos, apuntando especialmente a conseguir un salto de escala que posibilite explotar datos masivos.
- Promover el desarrollo de la conectividad y de nuevas aplicaciones y servicios, aprovechando la nueva infraestructura de telecomunicaciones provista por el Estado, abarcando las necesidades de todos los actores del agro, con especial énfasis en los pequeños productores y agricultores familiares, y en la sustentabilidad de la producción.

Área de salud

De la consulta hecha a los expertos por el proyecto que venimos discutiendo surgen como temas tecnológicos prioritarios:

- La elaboración de historias clínicas electrónicas.
- La georreferenciación de pacientes y traumas o el desarrollo de sistemas de gestión de pacientes.

Posteriormente consultas y la deliberación con la CAE demostraron que si bien la historia clínica electrónica se presenta en la actualidad como un paso obvio, a pesar de ser un tema de hace treinta años, no es el único relevante. Incluso, puede ser considerado como un primer paso que abra nuevas oportunidades. En ese contexto, se señaló la importancia de la telemedicina, incluyendo trauma, *telenursing*

⁵ En la reciente conferencia HydroGaïa (Montpellier, Francia, 12 y 13 de junio 2013) se presentaron variadas aplicaciones de TIC a distintos temas vinculados con manejo del agua, medioambiente y agricultura. Por ejemplo:

- Irrigación inteligente (*smarter irrigation*): mejorar la eficiencia en el uso del agua para la agricultura. Incluye: usar diversas fuentes de datos disponibles (requiere *data mining*) y aplicar mezclas de agua (*water blending*) a diferentes cultivos; relacionar diferentes modelos agro-hidrológicos y otras técnicas de análisis (usando *machine learning*); planear y utilizar herramientas eficientes para apoyo a la decisión (DSS).
- Temas clave:
 - Asimilación de datos.
 - Integración de modelado con procesos operacionales.
 - *AnalyTIC (Big Data)*.
 - Redes escalables y *pervasive* (ubicuas): conectar diversas tecnologías con diferentes rangos de alcance y fuentes de alimentación (RFID, Bluetooth, ZigBee, WiFi, Celular).
 - Humedad del suelo: imágenes infrarrojas a gran escala.
 - Sensores de telemetría para agricultura.
 - Utilización de pronósticos meteorológicos.
 - Integración de modelado y operaciones: utilización de *high performance computing* (por ejemplo, para modelos de clima, prevención de inundaciones, etc.).
 - Integración de imágenes aéreas para toma de decisiones: vigilancia y visualización para analizar variables y analizar tendencias (fertilidad del suelo, irrigación, productividad, etc.).
 - Sistemas geo-referenciados, incluyendo *streaming* de video on-demand en tiempo real (por ejemplo, CARMEL (Context-Aware Rich-Media Extensible Middleware de IBM).

y prevención de riesgos en el hogar. Por otra parte, la red de televisión digital puede ser utilizada para la implementación de sistemas de alerta temprana.

Algunas cuestiones asociadas a telemedicina, sobre todo las relativas a prevención en caso de epidemias podrían ser tratadas a través de la red de televisión digital dirigiendo instrucciones específicas a los diferentes grupos de riesgo.

Algunas oportunidades y acciones a promover señaladas por el proyecto son:

- Fomentar la compra de *software* desarrollado en el país.
- Impulsar la informatización de redes de salud: meta de “historia clínica electrónica por ciudadano”.
- Contar con infraestructura a nivel nacional para firma digital.

Sin embargo, se podría añadir a lo señalado por el proyecto, que si bien estas oportunidades y acciones resultan adecuadas a la actual situación del país con respecto a las aplicaciones de las TIC en la salud, sería bueno considerar algunas iniciativas a nivel global y local que deberían estar incluidas en un análisis que permita delinear una agenda del futuro en este rubro.

En primer lugar, los programas de *e-health* de la *World Health Organization (WHO)* de las Naciones Unidas, cuyos objetivos y acciones poseen un foco relevante en la utilización de las tecnologías para el cuidado de la salud en las naciones menos desarrolladas.

Los programas y proyectos de *WHO en e-Health* abarcan áreas tales como política y gobernanza, estandarización e interoperabilidad, investigación y reportes globales, *e-learning*, redes y colaboración sur-sur, así como aplicaciones de *e-Health*.

Algunos programas relevantes de *WHO* con respecto a e-salud son:

- *Global Observatory for e-Health*.
- *Health Academy*.
- *WHO e-learning resources for health workforce training*.
- *Governance*.
- *National e-Health Strategies*.

Un programa de particular interés resulta *The Health Internet* cuyos objetivos son de gran alcance, sien-

do relevantes los relacionados con los nuevos desafíos y riesgos que involucra la digitalización de la información de salud de las personas a nivel global.

Esta relativamente nueva situación requiere de investigación tecnológica (algunos de cuyos objetivos han sido mencionados anteriormente) pero particularmente nuevas regulaciones y legislaciones a escala nacional, regional y global, que son y serán motivo de difíciles negociaciones en todos los niveles. Por otra parte, los grandes datos y el análisis de datos de salud están pasando a ocupar un lugar de privilegio en la investigación médica y biomédica.

Algunos proyectos e iniciativas -seleccionados arbitrariamente de entre los cientos actualmente en marcha- son los siguientes:

- La iniciativa *Project Data Sphere* que consiste en una plataforma de datos, disponible para empresas, organismos públicos e investigadores, para la investigación del cáncer.
- La *Plataforma Artemis* que es una aplicación de *Big Data* en unidades de cuidados intensivos para recién nacidos. Artemis posibilita diagnósticos concurrentes de múltiples pacientes a través de análisis en tiempo real y múltiples streams de datos, producidos por diferentes dispositivos de censado.
- El Programa *GENEVA (Gene Environment Association Studies)* es la componente sobre genética de la iniciativa *GEI (Genes, Environment and Health Initiative)* de la *NIH* en los Estados Unidos. El Programa *GENEVA* está dirigido por el *National Human Research Institute* y el *National Center for Biotechnology Information*, el cual gestiona el repositorio de datos para los estudios.

A nivel local, debe destacarse la ya mencionada *Plataforma de Bioinformática Argentina (BIA)*. La BIA brinda apoyo relevante a soluciones para la salud; entre ellas:

- Análisis de relevancia clínica de genomas.
- Análisis de exomas para diagnóstico de enfermedades raras.
- Desarrollo racional y optimización de fármacos.
- Análisis de estructuras de proteínas.
- Análisis de *amplicon deep-sequencing*.
- Análisis de *microarrays*.

Es importante señalar con énfasis, que BIA es un proyecto central que marca una tendencia fundamental para el desarrollo de las aplicaciones de las TIC -en particular la bioinformática- para la salud.

Recientemente, el gobierno nacional [ha anunciado](#) que en los nuevos documentos de identidad (“Nuevo DNI”) se incluirá una historia clínica del ciudadano.

Esto constituye un nuevo paso en la construcción de una infraestructura para la sociedad del conocimiento consistente con otros programas del gobierno nacional que serán discutidos en la sección 4. El alcance que paulatinamente -a medida que se renueve el parque de DNI- irá teniendo este proyecto cubrirá a una vasta porción de la población, y seguramente en unos años a la totalidad.

En estas condiciones, se instala la posibilidad concreta de dar un salto cualitativo hacia la medicina personalizada en la dirección propuesta por *WHO*, antes mencionada, y en particular por *The Health Internet*. Es necesario, por una parte, generar un programa interdisciplinario, interministerial e interprovincial que considere los aspectos sanitarios, tecnológicos, legales y regulatorios que posibiliten la concreción de este objetivo.

Algunas experiencias, en particular las desarrolladas en la Unión Europea -como el [portal salud en línea](#)- pueden aportar al desarrollo de este programa.

En definitiva, es recomendable concebir las recomendaciones realizadas por la CAE (historia clínica electrónica, georreferenciación de pacientes, desarrollo de *software* local, firma digital) en el marco de un único programa que aproveche la infraestructura que brindará el Estado Nacional y avance hacia el cuidado de la salud personalizada por medios electrónicos. Se trata de una gran oportunidad para empresas locales, organismos públicos y grupos de investigación.

Área de energía

De la consulta a los expertos hecha por el proyecto en 2012 surgieron como temas tecnológicos prioritarios:

- Las redes inteligentes de distribución de energía: para ello hay oportunidades por el dinamismo del mercado internacional, las capacidades locales de innovación, el espacio para la sustitución de importaciones y el desarrollo de una sociedad del conocimiento.
- La optimización de procesos tanto en la actividad de *upstream* como *downstream*, y logística en gas y petróleo.
- Logística en gas y petróleo (71%).
- Automatización en viviendas (domótica): hay oportunidades por haber un mercado internacional muy dinámico que favorecía su desarrollo en el país.

Posteriormente las consultas y las deliberaciones de la CAE resaltaron, además de lo anterior, la importancia del aporte de las TIC para el uso más eficiente de la energía. También señalaron como oportunidades, los medidores inteligentes (por ejemplo, para facturación) y sistemas de energía en “ciudades inteligentes” (*smart cities*).

El Proyecto MinCyT 2012-2014 señaló también en su informe final la relevancia en las actuales circunstancias del país de la aplicación de las TIC a la explotación de petróleo y gas. Técnicas sofisticadas de prospección y exploración basadas en la adquisición, procesamiento y representación de grandes volúmenes de datos (*Big Data*) y su visualización con técnicas de *3D imaging* están siendo claves para la explotación exitosa de yacimientos no convencionales en Estados Unidos⁶.

Pero sobre todo, una oportunidad relevante aparece con motivo de la estatización de YPF, que abre perspectivas de utilización de TIC a gran escala tanto en la producción, distribución, logística, gestión, etc. de gas y petróleo.

Se considera relevante analizar e impulsar las oportunidades que abría la estatización de YPF, tanto a

⁶Más allá de las nuevas fuentes de energía que podrán eventualmente cubrir una parte de las necesidades del desarrollo nacional en el futuro, muchas expectativas se depositan actualmente en los llamados yacimientos de petróleo y gas no convencionales, especialmente en el yacimiento de Vaca Muerta. Según noticias ampliamente difundidas, la explotación de estos yacimientos requieren de tecnologías muy sofisticadas y grandes inversiones; las principales técnicas utilizadas se denominan “perforación horizontal” y “fractura hidráulica”. Pero justamente, las *Big Data* son las que pueden permitir una utilización adecuada de estas técnicas: como dice una reciente nota aparecida en Forbes, a partir de las *Big Data* “están llegando grandes cambios a la industria del petróleo y el gas” (“[las Big Data van a acelerar una revolución en la perforación inteligente de petróleo y gas](#)”).

partir de una mayor vinculación con las universidades como con los proveedores. A su vez, también se recomendó apoyar y acompañar las iniciativas referidas a la explotación no convencional de petróleo y gas.

El futuro de YPF como gran empresa de energía, trascendiendo el objetivo del autoabastecimiento energético, con proyección como competidor internacional, depende en buena medida del éxito en la explotación sustentable de los yacimientos de petróleo y gas no convencionales, cuyo caso más relevante es Vaca Muerta.

Alcanzar este gran objetivo requiere de enormes inversiones que exceden la capacidad de los recursos nacionales. Sin embargo, al menos tan importante como el financiamiento, resulta la investigación acerca de las particulares características físicas y geológicas de estos yacimientos, así como el desarrollo de las tecnologías necesarias para lograr la extracción de los hidrocarburos y la sustentabilidad medioambiental de todo el proceso.

En esa dirección, la creación de Y-TEC es un paso relevante hacia dichos objetivos. Tal como se indica en su sitio web *“Y-TEC nace con la vocación de convertirse en un referente internacional en tecnologías aplicadas a la producción de petróleo y gas no convencional”*. En ese marco conceptual, resulta relevante entender y poner en acción la utilización de las TIC como una herramienta imprescindible para lograr dicho objetivo.

Como se indicó más arriba, las TIC están presente a lo largo de toda la cadena de producción y servicios relacionados con la industria del petróleo y gas, sin embargo, en el campo de los hidrocarburos no convencionales las tecnologías asociadas a los grandes datos están comenzando a generar resultados que se traducen en ventajas competitivas fundamentales, a la luz de los enormes costos de explotación de este tipo de explotaciones.

En la sección 3.5.6 se retoma esta cuestión, por el momento es importante señalar que Y-TEC deberá ir construyendo una infraestructura de grandes datos y capacidades humanas y tecnológicas para poder aprovecharlos.

En función de estos objetivos se recomienda enfáticamente:

- Analizar e impulsar las oportunidades que abre la estatización de YPF, a partir de una mayor vinculación con las universidades y con los proveedores. En particular, debe apoyarse fuertemente el crecimiento de Y-TEC y su vinculación con empresas tecnológicas y centros de investigación.
- En el caso particular de las TIC, debe promoverse el desarrollo de un infraestructura de grandes datos y su explotación productiva, especialmente -aunque no solamente- en temas de modelización, simulación y visualización científica. De acuerdo a la experiencia internacional, estas áreas son clave en la explotación no convencional de petróleo y gas.

Área de industria

De la consulta a los expertos, hechas por el proyecto, como para el resto de las áreas de aplicación, en 2012, surgieron como temas tecnológicos prioritarios:

- La informática aplicada a la industria automotriz.
- La automatización.
- TIC en maquinaria agrícola.

Entre las acciones a promover se recomendó:

- Política de “compre TIC nacional”.
- Impulsar el desarrollo de mecanismos de fondeo para la adquisición de intangibles.

En recientes documentos producidos por el Ministerio de Ciencia⁷ se realizó un profundo análisis de la industria automotriz, cuyas conclusiones, en gran medida aplicables a otros sectores de la industria manufacturera, dejan ciertas dudas acerca de las posibilidades reales de crecimiento del sector.

La amplia hegemonía de las multinacionales en cuanto a las decisiones de desarrollo tecnológico, las dificultades en la relación intra- Mercosur y en particular la creciente asimetría en relación con Brasil, dejan un margen relativamente estrecho para el desarrollo local.

El desarrollo de una industria automotriz local y autónoma aparece, al menos, complicado. El desa-

⁷ El futuro del sector automotriz en el mundo 2025, 2014; El futuro del sector automotriz en Argentina y en el MERCOSUR, MinCyT, Argentina 2025, 2014 (en vías de publicación).

rollo tecnológico e innovación local parece confinado a algunas autopartes poco complejas.

De todos modos, también de acuerdo a dichos informes, es posible y necesario desarrollar un programa de investigación y desarrollo, y de formación de recursos humanos de mediano plazo, que coadyuve a conseguir un crecimiento cualitativo de la industria con mayores márgenes de autonomía, algo que Brasil ya ha comenzado a hacer.

En esta dirección, las TIC pueden ser una variable importante para dar respuestas innovadoras a varios de los sistemas que componen un automotor, considerando las tendencias tecnológicas hacia el futuro.

Los sistemas de iluminación inteligente; faros LED; controles de intensidad lumínica; sistemas de seguridad como suspensión predictiva; frenado automático; control de estabilidad; *airbags* en cinturones traseros e inteligentes, de ventanillas; etc.; asistente en intersecciones y atascos; cinturones inteligentes; control de crucero adaptativo (*ACC*); asistente de trayectoria (*line assist*); asistente de cambio de carril (*side assist*).

En todos estos sistemas y subsistemas y en otros, hay lugar para la innovación basada en las TIC. Por otro lado, las TIC son absolutamente relevantes en los vehículos del futuro. Algunas tendencias que ya están siendo recorridas son:

- Conducción autónoma.
- Integración de los controles de audio/climatización, el navegador GPS, la conexión a internet y el equipo de música a través de una pantalla táctil y de la voz.
- Conectividad para vehículos.
- Conexión *wi-fi* con reconocimiento de entorno.
- Comunicación *Car to Car*.
- Monitoreo de personas con ciertas enfermedades.

En todas estas temáticas hay capacidades locales para realizar innovación, desarrollo e investigación (I+D+i) tanto en las empresas como en los equipos de investigación y desarrollo, lo cual muestra que, si se adoptan políticas promoción, se negocia fuertemente con las multinacionales y se buscan acuerdos de mutuo provecho en la región, espe-

cialmente con Brasil, es factible encontrar un sendero de D+I en relación con el sector automotriz.

5. HACIA UNA SOCIEDAD DEL CONOCIMIENTO EN ARGENTINA

El Libro Blanco de la Prospectiva TIC se focalizó principalmente en el análisis del futuro de la industria del *software*. En el proyecto desarrollado por el Ministerio en 2012-2014, teniendo en cuenta el grado de desarrollo de esta industria y su consolidación, se trató de analizar el apoyo que la misma y todo el área TIC (incluyendo electrónica y comunicaciones) podría ofrecer a la nueva infraestructura que está creando el Estado Nacional para promover una sociedad del conocimiento (Argentina Conectada, Conectar Igualdad, Igualdad Cultural, Televisión Digital Abierta).

Como se dijo en el diagnóstico, *“en relación al desarrollo de infraestructura aún nos encontramos en la etapa de construcción o despliegue de la infraestructura física”*, sobre todo teniendo en cuenta el abandono de la investigación y desarrollo tecnológico en comunicaciones, como consecuencia de la privatización y extranjerización de los servicios telefónicos. Resulta clave entonces entender cuáles serán los futuros pasos en cuanto a: i) gestión, mantenimiento, expansión y actualización de esa infraestructura y su, ii) aprovechamiento y explotación. También resulta importante entender los pasos para ampliar la participación de empresas nacionales en algunos rubros y de algunos equipos de I+D universitarios, entre otros.

En este tema, cabe hacer una mención a la gobernanza de las infraestructuras TIC, tanto en materia de eficiencia energética y *“Green IT”* como en materia de los servicios capaces de proveer dicha infraestructura y la infraestructura de datos entre otras cuestiones. El desarrollo de ARSAT como una gran empresa pública de comunicaciones, así como el crecimiento en cantidad y calidad de las infraestructuras y los servicios provistos por el Estado y el sector privado en *IT*, requieren de esfuerzos público-privados para abordar este tema.

Para el aprovechamiento y utilización de la infraestructura que se está promoviendo para una sociedad del conocimiento, dos temas son fundamentales, el gobierno digital y los contenidos digitales:

- *e-Government*: eficiencia y efectividad de gobiernos y administraciones será una demanda pública creciente. La transparencia, así como la agilización de trámites públicos es crucial. Al respecto, hay que llamar la atención sobre la necesidad de especificar los requerimientos técnicos y legales para el desarrollo de *e-government*.
- Contenidos digitales y culturales: la sociedad del conocimiento también requerirá de una gestión eficiente de sus contenidos culturales. En tal caso, será necesario el desarrollo de repositorios y bibliotecas digitales, así como una agregación distribuida de contenidos, incluyendo la digitalización de trabajos científicos, educativos y culturales; además de la previsión de accesos abiertos y datos abiertos; administración inteligente de información, y aprendizaje mediado/potenciado por la tecnología.

Estos temas requieren de investigación, desarrollo e innovación en Argentina en los próximos años a fin de poder explotar el despliegue de la infraestructura pública en curso (Argentina Conectada, Conectar Igualdad, los sistemas de información científica del Ministerio, Igualdad Cultural, BACUA, CDA, etc.).

- Componentes y sistemas: para el desarrollo de las TIC en general y, en modo particular, para el apoyo a la infraestructura resulta también importante impulsar el desarrollo de la micro y nano-electrónica. Dentro de estas áreas, se destacan por su importancia los sistemas y componentes embebidos *-smart systems-*, la fotónica y las llamadas *OLAE (organic and large-area electronics)*.

6. NUEVAS TENDENCIAS Y DISRUPCIONES EN LAS TIC

Además de las nuevas tendencias y desafíos considerados hasta aquí, en cuanto a la contribución de las TIC a una sociedad del conocimiento y a las áreas prioritarias del Ministerio, el Proyecto Min-CyT 2012-2014 empezó a analizar (pgs. 74 y ss.) también algunas nuevas tendencias y disrupciones que se abren desde el campo de las TIC, y que se espera dominarán el mundo hacia el 2020: *Big Data*, *cloud computing* e Internet de las cosas⁸.

Estas tendencias han sido analizadas con mayor profundidad después del proyecto, a partir del análisis de las tendencias mundiales y de la contribución de expertos argentinos en el país y en el exterior. Asimismo, es importante tener en cuenta que en estos campos los avances son continuos, por lo que habría que actualizar y monitorear continuamente estos cambios, como ya ha empezado a hacerlo el Ministerio a través de la Fundación Sadosky. Las tendencias en cuestión son las siguientes:

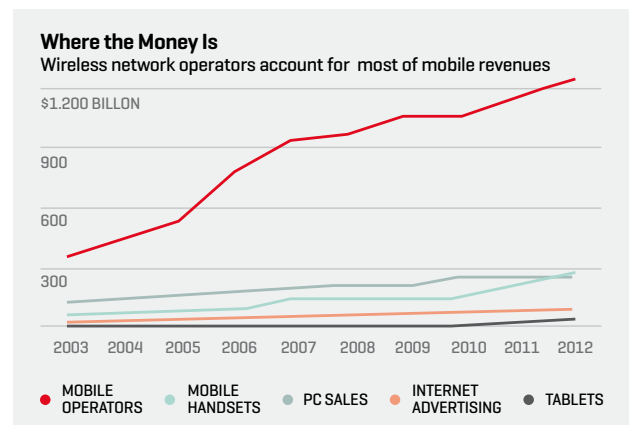
6.1. Computación móvil

La expansión mundial de la conectividad (a Internet) móvil, alcanza ya a más de 1.100 millones de personas en el planeta, un mercado inmenso y casi virgen (en términos de lo que podría ocurrir en relación a lo que realmente ocurre), en términos de aplicaciones y servicios útiles para los usuarios. Algunos datos obtenidos de *MIT Technology Review* y otras fuentes, indican que 5.800 millones de personas aún no están conectadas, de las cuales 4.500 millones nunca han accedido a Internet.

Es decir, el gran negocio de la computación móvil recién comienza (o tal vez aún no comenzó), la gran tajada de la torta se la llevan las compañías de comunicaciones móviles (*wireless carriers*): globalmente 900 de ellos ganan US\$1.3 trillones anuales, alrededor de 4 veces la ganancia sumada de Google, Apple, Microsoft e Intel, los famosos grandes ganadores de las TIC, aun cuando

los constructores de dispositivos -especialmente Apple- consiguen más ganancias individualmente.

De todas formas, está claro que a escala global estamos aún en la etapa del despliegue de la infraestructura, falta todavía mucho -casi todo- en cuanto al agregado de valor sobre la infraestructura. El siguiente gráfico, muestra la realidad del mercado mundial:



Source: Benedict Evans Enders Analysis.

En 2007, el precio promedio de un teléfono móvil era de US\$120 y en declive; los analistas hablaban de "saturación del mercado" puesto que casi cualquiera que pudiera comprar uno ya lo tenía. Pero desde entonces, los precios han caído en un 50%, y las ganancias por las ventas de dispositivos móviles se han duplicado.

Las aplicaciones y servicios aún representan el rubro de menor facturación en computación móvil. Los avisos comerciales para móviles se estiman en solamente US\$9 billones, sin embargo es en este punto donde aparecen las mayores oportunidades.

Facebook tiene una audiencia mensual nunca antes alcanzada (alrededor de 1.000 millones de personas), y en enero de 2013 se estima que, por primera vez, la audiencia de dispositivos móviles fue mayor que la de PC.

⁸ En especial, cabe mencionar los aportes de Mario Nemirovsky, experto de la Universidad de California y que trabaja en el Centro de la Supercomputación de Barcelona, a Gabriel Taubin, de la Universidad de Brown, y Nora Sabelli, de la NSF y Stanford.

6.2. Internet de las cosas

Junto a esta explosión de los dispositivos móviles, redes sociales, etc. aparece la llamada "Internet de las cosas" (IoT). Según el *Cluster of European Research Projects on Internet of Things (CERP-IoT)*:

*Internet de las cosas es una parte integral de la Internet del futuro y podría definirse como una infraestructura de red global dinámica con capacidades de autoconfiguración, basada en protocolos de comunicación interoperables donde las 'cosas' físicas y virtuales tienen identidades, atributos físicos y personalidades virtuales, usan interfaces inteligentes, y están completamente integradas en la red de información. En la IoT, se espera que las 'cosas' sean participantes activas en los procesos de distinto tipo para los cuales estén capacitadas y habilitadas para actuar, y que se comuniquen entre ellas y con el medio, intercambiando datos e información censada acerca del ambiente, en tanto reaccionan autónomamente ante los eventos del "mundo real/físico" y lo influyen ejecutando procesos que disparan acciones y crean servicios con o sin intervención humana directa"*⁹.

Si todos los objetos y personas en su vida cotidiana estuvieran equipadas con tales identificadores (también llamados etiquetas o *tags*), podrían ser gestionados e inventariados a través de computadoras; equipando todos los objetos en el mundo con dispositivos de identificación minúsculos que podrían transformar la vida diaria, la producción, el comercio, etc.

Por ejemplo, las industrias no se quedarían sin stock, ni tampoco producirían en exceso, ya que todos los involucrados conocerían qué productos son los que efectivamente se demandan y se consumen.

6.3. Big Data

"...es, en el sector de tecnologías de la información y la comunicación, una referencia a los sistemas que manipulan grandes conjuntos de datos (o data sets). Las dificultades más habituales en estos casos se centran en la captura, el almacenado, búsqueda,

*compartición, análisis, y visualización. La tendencia a manipular ingentes cantidades de datos se debe a la necesidad en muchos casos de incluir los datos relacionados del análisis en un gran conjunto de datos relacionado, tal es el ejemplo de los análisis de negocio, los datos de enfermedades infecciosas, o la lucha contra el crimen organizado"*¹⁰.

El límite superior de procesamiento se ha ido desplazando a lo largo de los años, de esta forma los límites que estaban fijados en 2008 rondaban los órdenes de *petabytes* a *zetabytes* de datos.

Los científicos con cierta regularidad encuentran limitaciones debido a la gran cantidad de datos en ciertas áreas, tales como la meteorología, la genómica, la conectómica, las complejas simulaciones de procesos físicos, y las investigaciones relacionadas con los procesos biológicos y ambientales, las limitaciones también afectan a los motores de búsqueda en internet, a los sistemas de finanzas y a la informática de negocios.

Los *data sets* crecen en volumen debido en parte a la introducción de información ubicua procedente de los sensores inalámbricos y los dispositivos móviles (por ejemplo las VANETs), del constante crecimiento de los históricos de aplicaciones (por ejemplo de los logs), cámaras (sistemas de teledetección), micrófonos, lectores de *radio-frequency identification*.

La capacidad para almacenar datos de la humanidad se ha doblado a un ritmo de cuarenta meses desde los años ochenta. En 2012, cada día cerca de 2.5 quintillón (2.5×10^{18}) *bytes* de datos fueron creados.

6.4. Cloud Computing

La elevada penetración de la computación en nube (o *cloud computing*) es una tendencia clara de la Sociedad de la Información y el Conocimiento. Esa masiva generalización se vio favorecida por el desarrollo de las redes de telecomunicaciones, la extensión de los servicios de banda ancha, la amplia capacidad de los sistemas informáticos, la creciente "virtualización" del entorno personal y el mundo de las aplicaciones.

⁹ El término Internet de las Cosas fue introducido por Kevin Ashton en 1999, en tanto el concepto se popularizó principalmente a través de las publicaciones del Auto-ID Center de MIT. En general, se toma como prerrequisito para definir la IoT la Identificación por radiofrecuencia (RFID) de las 'cosas' a ser integradas a la red, aunque una amplia variedad de tecnologías son útiles para hacerlo, por ejemplo NFC, códigos de barras, QR codes y marcas de agua digitales.

¹⁰ De Wikipedia, artículo "*Big Data*".

Según la definición del *IEEE Computer Society* es un paradigma en el que la información se almacena de manera permanente en servidores de Internet y se envía a las caches temporales de los clientes, lo que incluye equipos de escritorio, dispositivos móviles, etc.

Esto genera un nuevo escenario profesional y de consumo que afecta capacidades, usos y costumbres, así como métodos de trabajo y cadenas de valor. Prueba de la importancia que reviste, algunos países ya trabajan en desarrollar plataformas para fomentar el desarrollo de *software* en la web como refleja el proyecto Morfeo de la Comunidad Europea.

Por otra parte, los avances en la investigación sobre *Big Data* permiten predecir que en los próximos años se verán progresos en diferentes disciplinas científicas y de negocios basadas en la generación automática de múltiples modelos sobre la base de encontrar patrones y/o tendencias soportadas por la semántica de dichos modelos y en la comprensión de los contextos de múltiples patrones en conjuntos de datos complejos.

Esto significará un paso adelante decisivo respecto del análisis o toma de decisiones basados en un único modelo preciso o en un único patrón o tendencia en un conjunto de datos complejo. Se trata de un salto cualitativo en las formas de generar nuevo conocimiento, que se apoya en un salto tecnológico que involucra las capacidades de obtener información a gran escala de múltiples fuentes, procesarlas eficientemente y obtener nuevas y múltiples relaciones semánticas entre tendencias y patrones.

Por otro lado, la computación en nube también se ha convertido en un motor de innovaciones como muestran algunos desarrollos en la industria del *software*, en particular los asociados al *Software* como Servicio (*SaaS*) o Plataformas como Servicio (*PaaS*)¹¹.

6.5. Hacia una integración de estas tendencias y disrupciones

Estas tendencias son, obviamente, importantes en sí mismas. Cualquier país u organización que pretenda avanzar en su desarrollo tecnológico debe dedicar importantes esfuerzos a la investigación, desarrollo, innovación y negocios en ellas.

Sin embargo, parece claro que las próximas disrupciones tecnológicas relacionadas con las TIC -y por cierto, con sus aplicaciones más relevantes en otras disciplinas, "duras" y "blandas"- tendrán mucho que ver con la intersección de las mismas.

Estamos así en un nuevo escenario de convergencia que está generando las condiciones para profundos cambios en términos económicos, científicos, culturales, etc. Para muestra, bastan algunos ejemplos interesantes, como se muestran en el siguiente recuadro:

- *Deep Learning*: una nueva área de *Machine Learning*, creada con el objetivo de acercar la disciplina a una de sus utopías fundantes, la inteligencia artificial. *Deep Learning* se ha convertido recientemente en un foco de investigación y negocios de Google que eventualmente le permitirá obtener ventajas decisivas para sus negocios (viejos y nuevos). El nuevo paradigma aparece entonces como inteligencia artificial repotenciada utilizando redes neurales sobre *Big Data* para generar servicios sobre dispositivos móviles a escala global.
- *Quantifying the Impact of Human Mobility on Malaria* [Amy Wesolowski, Nathan Eagle, Andrew J. Tatem, David L. Smith, Abdisalan M. Noor, Robert W. Snow, Caroline O. Buckee. *Science* 12 October 2012: Vol. 338 no. 6104 pp. 267-270]: utilización de millones de datos generados por poblaciones que se mueven –dejando rastros a través de sus comunicaciones móviles con teléfonos baratos– para investigaciones epidemiológicas.

¹¹ El *SaaS* alude a la disponibilidad de aplicaciones a las que puede accederse mediante un navegador, algo que antes sólo era posible si dicha aplicación estaba en el servidor. Algunas iniciativas son verdaderamente innovadoras (Facebook, Dropbox, etc.) y otras son simplemente la versión web de paquetes existentes.

PaaS refiere a plataformas de *software* que estuvieron tradicionalmente asociadas a los sistemas operativos para PC, pero el ágil desarrollo de la nube y las tecnologías de red propiciaron su irrupción y generalización. Las plataformas de *software* permiten el desarrollo y distribución de aplicaciones sin los costos de los elementos e infraestructuras tradicionales. Autores como Michael Cusumano del MIT hablan de la "configuración de un ecosistema de innovación abierto, en el que el proveedor de la plataforma trabaja con los proveedores de complementos para ofrecer aplicaciones que ayuden a elevar el valor de la plataforma"; pero en la práctica el bajo costo de desarrollo y de puesta en el mercado gracias a la web, hace que subsistan productos con poca diferenciación y pequeña base de usuarios.

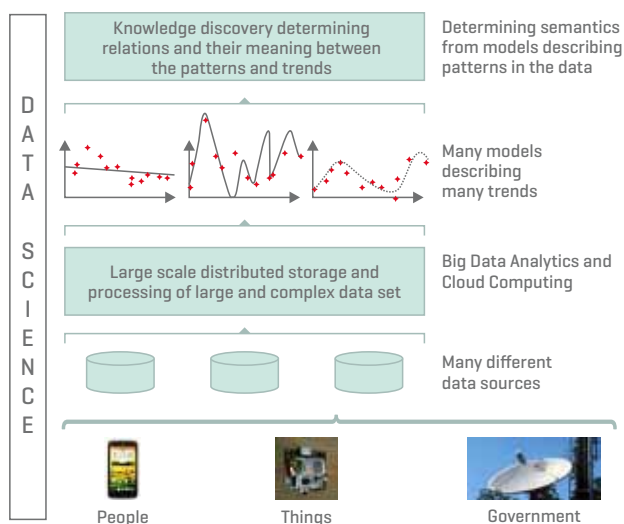
- Proyecto *Voices* de la UE (mvoices.eu): parte de *Mobile Web for Social Development Roadmap* que apunta a promover las capacidades locales y de emprendedores a través de la utilización de las TIC, en este caso a través del procesamiento de la voz como forma de superar barreras de acceso al conocimiento y la tecnología.
- **D4d: Data for Development** ha sido un desafío del Grupo Orange de telecomunicaciones orientado a equipos de investigación a nivel mundial, cuyo objetivo ha sido la utilización de registros de comunicaciones móviles (CDRs) para conseguir resultados innovadores en áreas de interés

como salud, transportes, emergencias, turismo en África. Los resultados fueron presentados recientemente en la conferencia **NetMob 2013** organizada por MIT.

Esta convergencia es disruptiva porque más y más acceso a la computación (*pervasive computing*) y más y más datos (*big data*), no sólo es mucho más, sino que puede ser diferente¹².

En la figura siguiente, se presenta un esquema de la situación:

ESCENARIO DE CONVERGENCIA DE LAS TECNOLOGÍAS¹³
 [CLOUD COMPUTING+BIG DATA] + [CHEAP PHONES + IOT]



¹² Como señala Nora Sabelli en una comunicación al proyecto : "...en cierta manera, porque la combinación cambia el balance de poder entre expertos y laicos, además de cambiar la ciencia en sí. Los físicos se dieron cuenta [de ello] hace mucho, y ahora los sociólogos están viendo cambiar sus teorías en base a mayores datos, muchos de ellos resultado del uso de "social media" por la gente común". Y continúa: "A scientific revolution is just beginning. It has the potential to create an era of science-based innovation that could completely eclipse the last half century of technology-based innovation; and with it, a new wave of global social, technological and economic growth".

¹³ Versión modificada de 2013: *Convergence of Big Data and Cloud Computing*, posted by Anton Heijs, dec 31, 2012.

El concepto de *Big Data* está todavía muy relacionado a lo que se llama *Big Science*. La cuestión, como ha ocurrido tantas veces en la historia, es ¿cómo se puede hacer para que *Big Data* pase de *Big Science* a otros aspectos de la vida, como “*Big Government*” o “*Big Business*”?

El corazón de la respuesta es la Ciencia de los Datos (*Data Science*), la disciplina que trata acerca de transformar los datos disponibles en nuevas ideas por medio de los cuales una organización puede tomar acciones o decisiones. El “microscopio” digital¹⁴, que está siendo inventado en este momento para obtener y manipular *Big Data*, permitiría rápidamente analizar y predecir eventos de todo tipo, mucho más exactamente y científicamente; descubrir estructuras y conexiones que posibilitarán obtener impensadas tendencias, patrones, *insights* y soluciones, y tomar mejores decisiones. Aparecen y se potencian viejas y nuevas técnicas: “*Data Driven Decisions*” y “*Predictive Analysis*”.

La Ciencia de los Datos incluye un conjunto de disciplinas que, en conjunto, proveen los conocimientos, técnicas y métodos para utilizar provechosamente el nuevo microscopio. Tal como se comentó anteriormente, *machine learning* es una de las claves para hacerlo; junto con ella, algoritmos y técnicas de estadística avanzada y de computación distribuida y de alto desempeño, forman el “núcleo duro” de la nueva ciencia.

6.6. La era de los Grandes Datos: esperanzas, problemas y recomendaciones

Es necesario comprender plenamente la “revolución” que va a significar *Big Data* y su convergencia con “*pervasive & trusted computing*” (computación ubicua y segura). *Big Data* refiere al sector de tecnologías de la información y la comunicación que manipulan grandes conjuntos de datos.

Las dificultades más habituales en estos casos se centran en la captura, el almacenado, búsqueda, análisis, y visualización de los datos. La tendencia a manipular ingentes cantidades de datos se debe a la necesidad en muchos casos de incluir datos relacionados, tal es el ejemplo de los análisis de negocio, los datos de enfermedades infecciosas, o la lucha contra el crimen organizado.

Las redes que surgen de las interacciones entre personas, sensores de todo tipo y organizaciones van dejando una trama de información que es un recurso clave para la ciencia, la producción y la sociedad. Es fundamental en esta dirección entender el futuro de esta nueva tendencia.

Por otra parte, el crecimiento de esta rama o disciplina basada en la informática abre nuevas preguntas en materia de regulación y política, por ejemplo, aspectos legales, privacidad, etc. Al mismo tiempo, pone nuevos desafíos a la ciencia y a la tecnología relacionados con la infraestructura física y humana, y la explotación de esta convergencia (*cloud computing*, computación de alto rendimiento -*HPC*-, redes fijas, móviles e híbridas, *machine learning*, estadística avanzada, visualización, sistemas complejos).

Pero fundamentalmente, ofrece nuevas e impensadas oportunidades en temas tales como salud pública, medicina personalizada, educación, ciencias “duras” y “blandas”, medios de comunicación, *marketing* y finanzas, entre muchos otros. Algunos de estos tópicos y aplicaciones -agro, salud, energía, industria- han sido discutidos en secciones anteriores, en lo que sigue se intentará aclarar un aspecto específico que resulta ser, a juicio del autor, el motivo principal de la disrupción que los grandes datos traen consigo.

Más allá de los requerimientos de equipamientos, gestión de una impensada hasta hace poco cantidad y variedad de datos, nuevos y viejos problemas acerca de la procedencia, “limpieza” y reducción de los datos, la cuestión cualitativamente diferente que es necesario investigar es el proceso interactivo, que involucra como nunca antes a humanos y máquinas, y posibilita transformar datos en conocimiento.

Este proceso se denomina inferencia, un término utilizado desde hace siglos pero que en la era de los grandes datos adquiere un significado muy particular. Será esta capacidad la que distinga la nueva era científico-tecnológica, en la cual toda clase de entidades, desde pequeñas y en algún sentido irrelevantes “cosas”, hasta seres humanos aportarán datos a una red que tendrá inteligencia para aprender y generar conocimiento genuinamente nuevo.

¹⁴ Se ha comparado el advenimiento de las *Big Data* a lo que fue la aparición del microscopio para la ciencia.

Un tema que será central en relación con estas redes podría denominarse sistemas complejos para la toma de decisiones en redes de hombres y máquinas. Los efectos de estos sistemas complejos son difíciles de sobreestimar, de hecho sistemas de este tipo están siendo aplicados en escenarios de guerra por las Fuerzas Armadas de Estados Unidos con resultados nefastos aunque impactantes.

Más allá de esto, su potencial en otro tipo de escenarios, como catástrofes naturales, industrias de procesos y manufacturas, redes eléctricas y de distribución de energía en general parece realmente inmenso. Los sistemas de este tipo son esencialmente el resultado de la irrupción de los Grandes Datos en la disciplina de la toma de decisiones, una rama importante de la matemática aplicada llamada investigación de operaciones. Esta temática, si bien tiene áreas importantes que dependen de los avances de la ciencia de los datos, presenta especificidades muy importantes cuyo estudio trasciende el presente informe y que deberán ser estudiadas por separado.

Volviendo a nuestro tema de interés, es claro que resulta vital identificar y abordar, en términos de programas de investigación básica, aplicada y desarrollo tecnológico, las principales líneas y problemas que permitirán avanzar la frontera del conocimiento, a la vez que dominar y desarrollar los aspectos tecnológicos que posibilitarán la apropiación de los resultados prácticos de esta nueva disrupción.

Enfrentar con éxito el desafío de los grandes datos requiere tener en mente, tal como se ha dicho y reiterado a lo largo de este documento, que nos encontramos frente a la “segunda ola” del desarrollo de una nueva revolución tecnológica, por lo cual no es esperable obtener soluciones simples o “llave en mano”, ni tampoco que las mismas provengan de una única disciplina; más bien, se pueden ofrecer caminos para avanzar en la comprensión del fenómeno y soluciones parciales para algunos de los desafíos.

En las próximas secciones, se mencionarán brevemente dichos temas y finalmente se expondrá una lista de desafíos de carácter práctico, así como una lista de problemas de mediano y largo plazo que requieren investigación de base y aplicada. La intención de estas enumeraciones es orientar los

programas que promuevan la investigación y el desarrollo en el campo de los grandes datos.

La investigación y el desarrollo necesarios para el análisis de grandes datos trasciende las fronteras de una única disciplina, y una de las principales conclusiones de este estudio es la necesidad de interdisciplinariedad, o tal vez convergencia¹⁵, para abordar los problemas de los datos masivos.

Los expertos en computación, encargados de la construcción de sistemas de *Big Data*, deben desarrollar una profunda comprensión de los aspectos relacionados con la inferencia, en tanto los estadísticos deben tomar en cuenta la escalabilidad, los aspectos algorítmicos y los problemas relativos a la necesidad de toma de decisiones en tiempo real.

Los matemáticos también tienen un rol importante que cumplir, puesto que áreas como álgebra lineal y optimización (que ya están contribuyendo al análisis de grandes datos) seguirán creciendo en importancia. Por otra parte, como se ha mencionado el rol del juicio humano es esencial, de modo tal que la contribución de científicos sociales y psicólogos así como la de expertos en usabilidad y visualización serán de gran valor.

Finalmente, los expertos en los diferentes dominios y los usuarios de la tecnología tienen un rol esencial que jugar en cualquier sistema para el análisis de datos. Esto es particularmente cierto en el terreno de los grandes datos a causa de la explosión de decisiones de diseño y las posibles direcciones que el análisis puede seguir.

Grandes esperanzas y grandes desafíos

En la medida en que los datos masivos puedan ser aprovechados de manera eficaz, crece la esperanza de que la ciencia ampliará sus fronteras y de que la tecnología será más adaptable, personalizada, y robusta.

Esto permite imaginar, por ejemplo, un sistema de atención a la salud en el que se registran datos cada vez más detallados de cada individuo -incluyendo datos genómicos, celulares y medioambientales-, y en el que dichos datos pueden combinarse con los de otras personas y con resultados de la investigación biológica y médica fundamental, para que pue-

¹⁵ *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond* (2014). http://www.nap.edu/catalog.php?record_id=18722

dan diseñarse tratamientos optimizados para cada individuo. Es posible también imaginar numerosas y originales oportunidades de innovación y nuevos negocios que combinan conocimientos de preferencias y necesidades, en el nivel de cada persona, con descripciones muy refinadas de bienes, servicios y capacidades para crear nuevos mercados.

Es natural ser optimista acerca de estos prospectos. Varias décadas de investigación y desarrollo en bases de datos y plataformas de búsqueda han dado lugar a la construcción de una valiosa experiencia en el diseño de tecnología escalable centrada en los datos. En particular, estos campos han posibilitado el advenimiento de la computación en la nube (*cloud computing*) y otras plataformas distribuidas adecuadas para el análisis de datos masivos. Más aún, la innovaciones en los campos de *machine learning*, *data mining*, estadísticas y teoría de algoritmos han dado lugar a la aparición de métodos de análisis que pueden aplicarse a conjuntos de datos cada vez mayores.

Sin embargo, tal optimismo debe atemperarse sobre la base de la comprensión de las grandes dificultades que aparecen al intentar obtener los objetivos previstos. En las próximas secciones se describen algunas de ellas.

Inferencia estadística

En parte, estas dificultades son las familiares en las implementaciones de bases de datos de gran porte -encontrar y mitigar cuellos de botella, obtener generalidad y simplicidad en la interfaz de programación, propagar correctamente los metadatos, robustez frente a las fallas de *hardware* y aprovechar las plataformas de *hardware* paralelas y distribuidas-, todo ello a una escala sin precedentes.

Pero los nuevos desafíos de los datos masivos van más allá de los problemas de almacenamiento-indexación-consulta que han sido ámbito de los sistemas de bases de datos clásicos (y de los sistemas de búsqueda clásicos). Tienen que ver con el nuevo y más ambicioso objetivo de la inferencia.

Inferencia es el problema de transformar datos en conocimiento, donde el conocimiento muchas veces se expresa en términos de entidades que no

están presentes en los datos *per se*, pero sí lo están en los modelos que se utilizan para interpretarlos. Es imprescindible el rigor estadístico para justificar el salto inferencial de los datos al conocimiento y allí aparecen muchas dificultades relacionadas con la aplicación de principios estadísticos a los grandes datos. Pasar por alto estos fundamentos puede llevar a resultados que, en el mejor de los casos, pueden no ser útiles, y aún ser perjudiciales, en el peor.

En cualquier debate sobre datos masivos e inferencia, es esencial ser consciente de que es muy posible transformar datos en algo que parece conocimiento pero que realmente no lo es. Más aún, es muy difícil darse cuenta de qué es lo que ha ocurrido. En efecto, muchas cuestiones son las que inciden en la calidad de la inferencia. Una de las más importantes es el llamado "sesgo muestral"; los datos pueden haber sido recolectados de acuerdo a un cierto criterio (por ejemplo, en una forma que favorece a los ítems "grandes" frente a los "pequeños"), pero las inferencias y decisiones tomadas pueden estar referidas a un criterio de muestreo diferente. Esta cuestión resulta ser particularmente severa en muchos conjuntos de datos masivos, los cuales muchas veces consisten de muchas sub-colecciones de datos, cada una recolectada de acuerdo a un criterio de muestreo particular y con poco control de la composición global.

Existen muchas fuentes de potenciales errores en el análisis de grandes datos, muchas de las cuales se deben al interés en las "colas largas" ("*long tails*"¹⁶) que frecuentemente acompañan a las colecciones de datos masivos. Los eventos en la "cola larga" pueden ser extraordinariamente raros, aún de un conjunto muy grande. Por ejemplo, en aplicaciones para *marketing* personalizado -o en general, para aplicaciones informáticas centradas en el consumidor-, donde el objetivo es, crecientemente, ofrecer servicios muy refinados y personalizados puede haber pocos datos disponibles de muchos individuos, aún en conjuntos de datos muy grandes.

En las ciencias es frecuente que el objetivo sea encontrar fenómenos raros o inusuales y que la evidencia para tales fenómenos sea débil, particularmente cuando se consideran las tasas de error asociadas con búsquedas sobre espacios caracterizados por grandes clases de hipótesis.

¹⁶ De Wikipedia: *long tails*.

Otras fuentes de error prevalentes en los grandes datos incluyen la naturaleza multidimensional de muchos conjuntos de datos, su heterogeneidad, sesgos debidos a patrones de muestreo con poco control y procedencia desconocida de ítems en una base de datos. En general, el análisis de datos está basado en suposiciones, y las suposiciones subyacentes a los métodos clásicos de análisis de datos suelen no ser válidas en el campo de los grandes datos.

Otro tema importante es la “procedencia”: muchos sistemas incluyen niveles de inferencia en los cuales los “datos” no son las observaciones originales sin el producto de un procedimiento inferencial de algún tipo. Esta situación ocurre, por ejemplo, cuando faltan algunas entradas en los datos originales. En un gran sistema que incluya inferencias interconectadas puede resultar difícil evitar circularidades, lo cual puede introducir sesgos adicionales y puede entonces amplificar el ruido.

Finalmente, cuando se consideran muchas hipótesis, aparece con fuerza el tema del control de las tasas de errores. En efecto, los conjuntos de grandes datos, generalmente involucran el crecimiento, no meramente en el número de los individuos representados (las “filas” de la base de datos) sino también en el número de los descriptores de dichos individuos (las “columnas” de las base de datos).

Más aún, muchas veces estamos interesados en la capacidad predictiva asociada con combinaciones de descriptores y esto puede llevar a una explosión exponencial en el número de hipótesis consideradas, con consecuencias severas sobre las tasas de error. Una apelación ingenua a una “ley de los grandes números” está lejos de poder justificarse; en todo caso, los riesgos asociados a fluctuaciones estadísticas en realidad pueden aumentar a medida que crece el tamaño de los conjuntos de datos.

En resumen, los grandes datos traen nuevos desafíos en el campo de la estadística, en especial en lo referido a las inferencias que pueden realizarse sobre los mismos. En efecto, si bien los estadísticos han desarrollado herramientas para abordar, en principio, los temas antes mencionados, en el contexto de los grandes datos es necesario tener cuidado por dos razones principales:

1. Todas las herramientas estadísticas están basadas en supuestos -explícitos o implícitos- acerca de las características de los conjuntos de datos y

la forma en la que fueron tomadas las muestras; sin embargo, dichas suposiciones pueden ser violadas en el proceso de construir los conjuntos de datos masivos.

2. Las herramientas para evaluar errores de procedimiento, así como para diagnósticos, son procedimientos computacionales que pueden resultar computacionalmente inviables en la medida que los conjuntos de datos cambian de escala (es decir, se transforman efectivamente en “*Big Data*”).

Convergencia multidisciplinar

El análisis de grandes datos no es el ámbito de una única disciplina, se trata de una empresa profundamente interdisciplinar. Las soluciones de los problemas relacionados con grandes datos requieren de la fusión de ideas provenientes de las ciencias de la computación y estadística, con contribuciones esenciales de la matemática pura y aplicada y de varias ramas de la ingeniería, notablemente procesamiento de señales y teoría de la información.

Científicos de dominio y usuarios de la tecnología también deben participar en todo el proceso de diseño de los sistemas para el análisis de grandes datos. Una amplia variedad de temas, que resultan críticos para el análisis de grandes datos (especialmente los relacionados con privacidad), requieren contribuciones de juristas, economistas y otros expertos en ciencias sociales, aunque estos aspectos quedan fuera del alcance de este estudio. En general, cuando se reúnen perspectivas interdisciplinarias para abordar el análisis de datos masivos, es posible discutir provechosamente ventajas y desventajas considerando conjuntamente las problemáticas computacionales, estadísticas, científicas y humanas que conforman el marco de un problema.

Cuando se considera un problema aisladamente, es frecuente intentar resolver un problema que es más general que lo que efectivamente se requiere y en tal caso puede no haber una solución factible para dicho problema general. En cambio, una mirada interdisciplinaria puede orientar a los investigadores hacia un cambio fundamental del foco de análisis. Por ejemplo, la ausencia de una intuición apropiada, puede llevar a analizar el peor caso del comportamiento algorítmico de una solución, lo cual puede ser muy costoso o aun sin sentido, cuando una mirada totalizadora podría revelar que

el comportamiento del caso promedio es totalmente apropiado desde una perspectiva estadística. Análogamente, el conocimiento del comportamiento típico en la generación de consultas puede posibilitar confinar un análisis a un subconjunto relativamente simple de todas las posibles consultas que deberían haber sido consideradas en un caso más general.

Finalmente, las conocidas dificultades de la programación paralela en situaciones genéricas puede ser eludida enfocando la atención en clases de algoritmos estadísticos que pueden ser eficientemente implementados con un conjunto de algoritmos paralelos sencillos, los cuales pueden sugerir modelos naturales de almacenamiento y acceso a los datos sobre plataformas de *hardware* distribuidas.

Variedad

Una gran diversidad de fuentes de datos potencian el rápido crecimiento del volumen de los grandes datos. Sin embargo, unas pocas formas de los mismos son las que crean desafíos particularmente interesantes. En primer lugar, gran cantidad de datos incluyen habla y lenguaje humano y crecientemente el interés consiste en extraer aspectos de la semántica subyacente a los mismos. En este sentido, algunos ejemplos son análisis de sentimientos, modelado relacional y el análisis semántico requerido para los sistemas automáticos de respuestas a preguntas, traducciones, etc.

Segundo, el exponencial crecimiento de datos de imagen y video que conforman la mayor parte de los datos está creando un vasto rango de desafíos en temas a gran escala como comprensión, procesamiento de imágenes, visión computacional y análisis semántico.

Tercero, cada vez más, los datos están siendo etiquetados con "*tags*" geo-espaciales y temporales, creando así una amplia variedad de desafíos relacionados con el mantenimiento de la coherencia en relación con las escalas de tiempo y espacio.

Cuarto, gran cantidad de conjuntos de datos masivos incluyen redes y grafos que posibilitan realizar inferencias acerca de la articulación de nociones semánticamente poderosas como "centralidad" e "influencia".

La profundidad del análisis requerido por las fuentes de datos antes mencionadas implica abordar

problemas muy complejos y/o no resueltos en inteligencia artificial y matemáticas que van más allá de las tareas -de corto plazo- relacionadas con aumentar la escala de los algoritmos existentes. Parece evidente que los grandes datos pueden dar lugar a un nuevo apalancamiento de tales problemas, entre ellos, la traducción automática del lenguaje natural es un ejemplo frecuentemente citado en la literatura.

Máquinas + personas

El análisis de los grandes datos crea nuevos desafíos en la interfaz entre personas y computadoras. Como se ha señalado, muchos conjuntos de datos requieren comprensión semántica que trasciende el alcance de las aproximaciones algorítmicas actualmente conocidas, por tanto es necesaria la contribución del juicio de los humanos.

Este *input* puede ser proporcionado por el analista de datos, cuyo juicio es necesario a lo largo de todo el proceso de análisis, desde la formulación de las hipótesis a la gestión de "*trade-offs*" (por ejemplo, error *versus* tiempo) y a la selección de las preguntas que podrán formularse al sistema. Junto con este aporte, también puede considerarse el de los expertos y usuarios de la tecnología a través de *crowdsourcing*, una herramienta potencialmente poderosa que debe utilizarse con cuidado, dado los muchos tipos de errores y sesgos que puede introducir. En cualquier caso, existen muchos desafíos que deben encararse en relación con el diseño de visualizaciones e interfaces y, más generalmente, en la vinculación del juicio humano con los algoritmos de análisis.

Tiempo real

Gran parte de las fuentes de grandes datos operan en tiempo real produciendo *streams* de datos que pueden desbordar largamente los circuitos de análisis. Más aun, es habitual la necesidad o el deseo de tomar decisiones rápidamente, tal vez en tiempo real.

Estos requisitos temporales ilustran un ejemplo particularmente claro de la necesidad de un intenso diálogo entre investigadores del campo de la estadística con los de las ciencias de la computación. En efecto, la investigación en estadística raramen-

te considera las restricciones debidas a la toma de decisión en tiempo real en el desarrollo de los algoritmos de análisis de datos. Recíprocamente, la investigación en ciencias de la computación raramente considera la complejidad computacional de los algoritmos para gestionar riesgo estadístico.

Facilitar el uso

Uno de los desafíos prácticos relevantes es el desarrollo de *middleware* -componentes de *software* que vinculan las especificaciones de análisis de datos de alto nivel con las arquitecturas de los sistemas distribuidos en el bajo nivel. Buena parte del trabajo sobre estas componentes de *software* puede facilitarse tomando herramientas ya desarrolladas en varios programas del campo de la computación científica, sin embargo, es necesario cambiar el foco, considerando las restricciones que imponen las necesidades estadísticas sobre las soluciones algorítmicas existentes.

Existe también una importante demanda de adaptación del *software* a los usuarios finales, de modo tal que un usuario no experto pueda realizar análisis de grandes datos sin tener una comprensión completa de los sistemas subyacentes ni de los temas relativamente complejos de la estadística. Sin embargo, no es esperable que el objetivo final en relación con el *software* para el análisis de datos masivos sea desarrollar soluciones "llave en mano". El ejercicio efectivo del juicio humano siempre será necesario en el análisis de datos y este juicio no puede sino estar basado en una comprensión al menos mínima y conceptual de los mecanismos computacionales y los principios estadísticos.

El desarrollo del análisis de grandes datos debe proceder en paralelo con un gran esfuerzo en la educación de los estudiantes y la fuerza de trabajo en el pensamiento estadístico y el pensamiento computacional. Entrenar estudiantes para trabajar con grandes datos requerirá desarrollar experiencia con datos masivos reales y con infraestructuras computacionales que revelen los verdaderos problemas asociados con los datos masivos. La disponibilidad de *benchmarks*, repositorios (de datos y programas) e infraestructuras de computación serán requisitos indispensables para la formación de los futuros "científicos de datos".

La misma observación puede hacerse en relación con la investigación académica: las nuevas ideas

relevantes surgirán solamente si los académicos son expuestos a problemas de datos masivos del mundo real.

No hay soluciones "llave en mano"

Finalmente, es importante señalar que el análisis de grandes datos no es un problema específico o una metodología particular. Frecuentemente, los datos son heterogéneos y la mejor manera de atacar un problema puede consistir en encontrar una descomposición en subproblemas a partir de la cual la mejor solución pueda elegirse sobre la base de razones de carácter computacional, inferencial o interpretacional. El descubrimiento de tales subproblemas puede resultar, en sí mismo, un problema inferencial.

Por otro lado, en muchas ocasiones, los datos posibilitan obtener visiones parciales de un problema. En tales casos, la solución puede consistir en fusionar múltiples fuentes de datos. Estas perspectivas de segmentación *versus* fusión no son contrapuestas o conflictivas, simplemente muestran la necesidad de recurrir al conocimiento del dominio y a mecanismos de abstracción y diseño adecuados para obtener la combinación adecuada.

Naturalmente, emerge la esperanza de obtener procedimientos estandarizados que puedan ser usados eficazmente sobre cualquier conjunto de datos masivos, de manera similar al uso de la Transformada Rápida de Fourier en el procesamiento de señales clásico. Sin embargo, al menos por el momento, es natural ser pesimista con respecto a la viabilidad de tales procedimientos. Esto no significa que no vayan a emerger procedimientos generales y rutinas útiles. Lo que es importante enfatizar es que el análisis de grandes datos no puede, en general, reducirse a procedimientos del tipo "caja negra" que los usuarios puedan utilizar sin pensar.

Por el contrario, el diseño de un sistema de análisis de grandes datos requiere habilidades ingenieriles y experiencia, en tanto el despliegue de tal sistema necesita de decisiones de modelado específicas, capacidades para evaluar aproximaciones, tratar con diagnósticos y robustez. En todo caso, las mejores expectativas deben situarse en la aparición de nuevas plataformas de *hardware* y *software* orientadas específicamente al análisis de grandes datos y en la emergencia de una nueva clase de in-

genieros capacitados para la creación y utilización de tales plataformas en el contexto de la solución de problemas del mundo real.

Algunas conclusiones sobre las Big Data

Las temáticas (y las precauciones) enunciadas anteriormente permiten capturar los verdaderos desafíos que traen consigo los grandes datos masivos. Es decir, posibilitan “separar la paja del trigo” en relación con los discursos del *marketing* de las empresas que focalizan en temas de grandes infraestructuras y soluciones “llave en mano”, algo que aún está bastante lejos de poder ser realidad.

Los verdaderos desafíos de los datos masivos están en el área del análisis de los datos. En primer lugar, la realización efectiva y eficiente de la transformación de los datos en conocimiento, esto es, en el campo de la ejecución efectiva de inferencias a través de mecanismos computacionales. Estos desafíos deben ser abordados a través de un importante y sostenido esfuerzo de investigación, sólidamente basado en principios computacionales e inferenciales (lógicos y estadísticos). Estos programas de investigación deben apuntar a desarrollar infraestructuras computacionales que materialicen, efectiva y eficientemente, los principios de inferencia los cuales, a su vez, están basados en consideraciones de escala.

La investigación debe tomar en cuenta también los ciclos de toma de decisión en tiempo real y la gestión de *trade-offs* entre velocidad y exactitud. En otro orden, la investigación debe orientarse a producir herramientas novedosas para incluir a los humanos en los ciclos de análisis de datos en todas sus etapas, reconociendo que el conocimiento es muchísimas veces subjetivo y dependiente del contexto y que muchos aspectos de la inteligencia humana no serán reemplazados -al menos en el futuro previsible- por máquinas.

Abordar la era *Big Data* presenta una cantidad de desafíos tanto en la gestión como en el análisis de los datos, que requieren nuevas formas de pensar y nuevas aproximaciones. Estos desafíos se extienden desde la generación de los datos y su preparación para el análisis, llegando a los relacionados con las políticas para compartirlos y utilizarlos, incluyendo los siguientes:

- Tratar con fuentes de datos altamente distribuidas.
- Trazar la procedencia de los datos, desde la generación hasta la preparación.
- Validar los datos.
- Tratar con los sesgos de muestreo y heterogeneidad.
- Trabajar con diferentes formatos y estructuras de datos (incluyendo datos no estructurados y semi estructurados).
- Desarrollar algoritmos que exploten eficazmente las arquitecturas paralelas y distribuidas.
- Asegurar la integridad de los datos.
- Asegurar la seguridad de los datos.
- Posibilitar descubrimiento e integración de los datos.
- Posibilitar el *data sharing*.
- Desarrollar métodos eficientes para visualizar datos masivos.
- Desarrollar algoritmos escalables e incrementales.
- Desarrollar métodos y técnicas para análisis de datos y toma de decisiones en tiempo real.

Encarar los desafíos y pensar en el mediano/largo plazo posibilita definir algunos de los principales temas de investigación que deberían abordarse para avanzar la frontera del análisis de los grandes datos:

- Representación de datos, considerando las caracterizaciones de los datos en bruto (*raw data*) y las transformaciones que habitualmente se aplican a los datos. Particularmente importantes son las transformaciones que buscan reducir la complejidad representacional de los datos.
- Complejidad computacional, en particular temas que ayuden a caracterizar los recursos computacionales necesarios y el balanceo en la utilización de los mismos.
- Construcción de modelos estadísticos en el contexto de *Big Data*, incluyendo particularmente limpieza y validación de datos.

- Muestreo, como parte del proceso de recolección de datos y como una metodología clave para reducción de datos.
- Métodos para incluir a las personas en el ciclo de análisis de datos, particularmente a través de técnicas como “*crowdsourcing*” -en la cual los humanos participan como fuente de datos en el entrenamiento de los algoritmos de aprendizaje-, y visualización, que no solamente ayuda a las personas a comprender los resultados de un análisis, sino también provee de *input* para la revisión de modelos.

Abordar exitosamente los desafíos y líneas de investigación descriptos requiere -como es obvio- de la participación de investigadores de diferentes disciplinas, incluyendo a los grandes usuarios de los grandes datos. En esa dirección, es recomendable la creación de una institución que lidere y articule las múltiples actividades necesarias alrededor de la temática. Por tal motivo, se recomienda la creación de un Centro de Ciencias de los Datos, que podría tomar diversas formas (podría ser un centro radicado en alguna universidad, también una red de laboratorios o una combinación de ambas). Pero, claramente, el Centro debería tener una misión directriz a nivel nacional y debería articular su actividad con el Centro de Modelado de Aplicaciones Tecnológicas radicado en el Polo Científico Tecnológico del Ministerio, la Plataforma de Bioinformática Argentina (BIA) antes mencionada y el Programa de Ciencias de los Datos de la Fundación Sadosky.

Asimismo, resulta de vital importancia que, integrado al Centro de Ciencia de los Datos o articulado con él, se genere un Centro de Visualización. Entre ambas iniciativas se podrían sentar las bases para un salto cualitativo para el desarrollo científico y tecnológico del país.

Ya en la discusión de la Comisión Asesora del Proyecto MinCyT 2012-2014, se destacó la importancia de este tema y hubo acuerdo en recomendar la creación de dichos centros. Se insistió, como lo muestra la experiencia internacional, en que deberían ser abiertos a la academia y a la industria y que deberían servir a diferentes dominios o áreas de aplicación.

Un gran tema para empezar sería el de la extracción no convencional de petróleo y gas por YPF; también habría una gran oportunidad para CONAE y ARSAT, por la gran cantidad de datos que posee

CONAE y especialmente los que obtendrá con su nuevo satélite para medir la humedad del suelo a partir de 2015.

En esta dirección, se ubica el Programa de Ciencias de los Datos (PCD) de la Fundación Sadosky. La creación del PCD fue una de las conclusiones de las *Jornadas de Definición Estratégica en Big Data* llevadas a cabo en agosto de 2013 por recomendación de la Comisión Asesora del Proyecto mencionado del Ministerio, en la cual participaron dos expertos internacionales y alrededor de 20 empresas, grupos investigación y representantes de organismos públicos. El PCD se ha propuesto impulsar desde todos los aspectos relevantes la creación y el desarrollo de áreas de investigación, formación de recursos humanos, aplicaciones relevantes, fomento de empresas, etc. En efecto, su programa de trabajo para el año 2014 incluyó los siguientes objetivos y acciones:

Desarrollo Tecnológico

- Plataforma de Datos Geolocalizados de Agro de la Argentina – *AgroBigData*.

Financiamiento a la investigación y desarrollo

- Generación de líneas de financiamiento específicas del Ministerio de Ciencia, Tecnología e Innovación Productiva, direccionadas a iniciativas de Grandes Datos.
- Creación de un cluster de empresas de *Big Data* y laboratorios de Ciencia de los Datos.
- Creación de un Centro Nacional de Ciencia de los Datos (CNCD).

Infraestructura

- Fomento al desarrollo de infraestructura y creación de un “*Cloud Nacional*”.
- Impulso a la Plataforma Nacional de Bioinformática.

Educación y formación de RRHH

- Co-organización del *Track de Big Data* en la Escuela de Ciencias Informáticas del Departamento de Computación – FCEyN – UBA.
- Generación de posgrados en Ciencia de Datos.
- Realización de una Escuela Anual de Ciencia de los Datos a partir de 2015.
- Realización de un Congreso o Jornadas Argentinas de *Big Data* a partir de 2015.

Hasta el momento, se ha avanzado en la Plataforma de datos geocalizados de agro de la Argentina – *agroBigData* en conjunto con AACREA (Asociación de empresarios agropecuarios), previéndose una primera versión de la misma para fines de 2014. Se han otorgado 10 becas de maestría en universidades de Estados Unidos a través del Programa BEC. AR y se otorgarán 10 becas para estancias breves a través del mismo. Se ha establecido que *Big Data* sea un tema prioritario en las convocatorias de la Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) y del CONICET. Se ha concretado el *trak de Big Data* en la Escuela de Ciencias Informáticas 2014 en la FCEyN-UBA y se han realizado numerosas acciones de difusión y divulgación de la temática tanto en ámbitos académicos como empresarios.

7. CONCLUSIONES Y RECOMENDACIONES

A lo largo de este documento, se ha pretendido continuar con el desarrollo de una línea de pensamiento prospectivo sobre las TIC en Argentina que comenzó en el año 2007 a través de distintos foros de debate presenciales y virtuales y tuvo una primera síntesis en el Libro Blanco de la Prospectiva TIC, publicado en 2009. A lo largo de 2012-2014, se desarrolló un proyecto, también promovido por el Ministerio que revisó, completó y actualizó el contenido del Libro Blanco, obteniendo nuevas conclusiones a la luz de los avances a escala global y nacional. Sin embargo, el trabajo finalizado en 2013, no alcanzó a tratar en profundidad algunas disrupciones y nuevas tendencias que aparecieron con mucha entidad y fuerza últimamente en todo el mundo.

La irrupción de los Grandes Datos promete ser el motor de una transformación de muchos aspectos importantes de la vida económica, social y cultural en el mundo, incluyendo temas estratégicos, militares, científicos, tecnológicos, productivos, sociales, comerciales e institucionales. Mirada en retrospectiva, la “revolución” de los grandes datos aparece como “la segunda ola” de una transformación que se inicia alrededor de una década atrás con la irrupción y generalización del uso de una enorme variedad de dispositivos –notablemente los dispositivos móviles- que llegaron a las manos de miles de millones de personas en todo el mundo.

Estos dispositivos son cada vez más “inteligentes”, en el sentido de que permiten a las personas no solamente comunicarse sino realizar computaciones cada vez más complejas. Por otra parte, más allá de muchas y muy válidas objeciones de carácter jurídico y más aún ético, estos aparatos generan fabulosas cantidades de datos que son almacenados, clasificados, “limpiados” y procesados utilizando las inmensas redes de comunicaciones y los servidores que las sustentan, y cuya gobernanza está hegemonizada por un relativamente pequeño conjunto de estados poderosos y grandes compañías.

Un fenómeno similar aparece a nivel de las cosas que nos rodean, a través del uso de sensores e identificadores implantados en los más diversos objetos de diferentes tamaños –incluyendo por cierto los cuerpos de los animales y personas-,

conectados también a las redes de comunicaciones; esto se conoce como la Internet de las Cosas. En resumen, estamos frente a una red global de personas, cosas y máquinas que interactúan, intercambiando datos, generando información y produciendo nuevos conocimientos.

Toda revolución tecnológica tiene aspectos extraordinariamente positivos y aspectos extraordinariamente peligrosos; la que estamos discutiendo tiene de ambos en altas dosis. Desde el punto de vista de la prospectiva, entendida como una herramienta de política científica y tecnológica, el primer paso para enfrentar efectivamente los nuevos desafíos es realizar los esfuerzos necesarios para comprender el nuevo fenómeno y sus teorías, métodos y técnicas, que han comenzado a dar sus primerísimos pasos. En ese sentido, es que se ha considerado necesario el estudio presentado en la Sección 6, y muy particularmente, las consideraciones sobre nuevos desafíos tecnológicos y líneas de investigación del apartado 5.5. Por cierto, dichas recomendaciones solamente refieren a actividades de investigación, desarrollo tecnológico y educación. En efecto, se considera fundamental que nuestros científicos y profesionales se apropien y desarrollen los conocimientos y técnicas propios de esta nueva etapa de la revolución de las TIC. No hacerlo implicaría un riesgo estratégico para el futuro; avanzar en estas líneas, en cambio, puede significar una oportunidad única para potenciar el desarrollo de nuestro país como una economía basada en el conocimiento y un salto en la competitividad internacional.

La promoción sin precedentes de la ciencia y la tecnología llevada a cabo desde el gobierno nacional, junto con el enorme salto en el desarrollo de la infraestructura de comunicaciones y la promoción de las TIC de manera general, crean un marco dentro del cual no solamente es factible la apropiación de las tecnologías de la nueva oleada, sino su utilización práctica para el desarrollo nacional. Estos objetivos requieren de una masa crítica de medidas innovadoras, alineadas consistentemente en la misma dirección. Por esta razón, es necesaria la creación de un programa que incluya varias líneas de acción interdependientes sobre las cuales

nuestro país debe avanzar rápida y decisivamente. La definición completa de estas líneas excede el alcance de este documento, sin embargo, cuatro de ellas de carácter estructural deben conformar su fundamento:

1. Construcción de una arquitectura nacional de servicios. Esta línea facilita la creación de servicios que atraviesen las fronteras organizacionales, reduciendo el trabajo redundante y mejorando el servicio a los usuarios.

La transformación en la infraestructura tecnológica inevitablemente crea un vacío donde las estructuras sociales continúan siendo funcionales a la etapa previa, aun cuando la nueva ya está naciendo. Argentina debe aprovechar la renovación de su infraestructura tecnológica para generar espacio para nuevos servicios y nuevos negocios. Una sociedad basada en el conocimiento requiere apertura, acceso a la información y servicios sencillos y eficientes para todos los ciudadanos.

De acuerdo al reciente reporte sobre gobierno electrónico de Naciones Unidas, muchos Estados Miembros están cambiando su rumbo, de un modelo organizacional disperso a una estructura centralizada, donde el gobierno funciona como un "todo". Los ciudadanos y empresas pueden así obtener servicios desde un único punto (muchas veces un portal gubernamental), independientemente de la organización que produce el servicio. Al mismo tiempo, la organización se vuelve más eficiente. Esta estrategia no implica que la gobernanza del sistema de servicios públicos digitales sea autoritaria, se trata de una estrategia organizacional.

En Argentina, la autonomía provincial y aún municipal, además de una trayectoria cambiante y hasta caótica a lo largo de décadas, ha llevado a una arquitectura de servicios fragmentada. Sin embargo, aún con la actual legislación, es posible conseguir una arquitectura nacional de servicios eficiente. Las propuestas de la [Agenda Digital](#) cuya actividad comenzó en 2009 y aún continúa, son importantes en temáticas centrales como la interoperabilidad de todos los sistemas de información del sector público.

La infraestructura generada a través del programa Argentina Conectada y la eventual aprobación de la Ley Argentina Digital, abren una gran oportunidad para transformar la actual fragmentación en una arquitectura nacional de servicios como se ha logrado implementar en [Estonia](#). Las mejoras en

los nuevos DNI que se comenzarán a distribuir en 2015 son importantes, aunque aún insuficientes, en ese sentido.

Construir una arquitectura nacional de servicios uniforme significa, básicamente, que los datos en los sistemas creados para diferentes funciones estarán disponibles a través de interfaces abiertas para todos los sistemas que los necesiten. Cada sistema gestiona sus propios datos y asegura que los datos requeridos por otros sistemas están disponibles a través de una plataforma de despacho (*delivery*) en tiempo real y en el formato correcto.

Otros aspectos fundamentales para la arquitectura de servicios son:

- Implementación de intercambios de datos en tiempo real entre el sector público y las empresas, lo que reduciría mucho trabajo rutinario e incrementaría el volumen de información actualizada. Esta tarea posibilitaría además un control más eficiente de la evasión impositiva, el trabajo informal, etc.; pero además posibilitaría la modernización de la gestión de las empresas, en especial de las PyME, así como su acceso eficaz a programas de promoción públicos, fondos de inversión y también la interacción entre las firmas.
- Mejorar la cantidad, calidad y confiabilidad de las conexiones de banda ancha fija y móvil es imprescindible para el desarrollo de una economía de servicios digitales. El Estado nacional y algunas provincias han dado pasos fundamentales, sin embargo, es necesario un esfuerzo mayor y el establecimiento de objetivos claros en cuanto a los alcances, la seguridad y los niveles de servicios requeridos. Junto con esto, es necesario un esfuerzo continuo para mejorar las conexiones internacionales, en especial con la región sudamericana, perseverando en la construcción del anillo de comunicaciones sudamericano.
- Promover una iniciativa nacional de datos abiertos. En la dirección propuesta por la Agenda Digital Argentina, es de gran relevancia sostener un amplio y comprehensivo programa de datos abiertos. En la era de los Grandes Datos, es necesario tener claro que los datos son una materia prima similar a los recursos naturales y su procesamiento y utilización tienen un enorme potencial para la generación de nuevos servicios y nuevos negocios.

- Abrir los repositorios de datos del sector público produciría enormes beneficios. Ejemplos de este tipo de repositorios incluyen información geográfica, estadísticas y datos producidos por proyectos de investigación financiados por el sector público. El [Sistema Nacional de Repositorios Digitales](#) es un muy buen ejemplo. La promoción de la investigación y desarrollo de datos abiertos y datos masivos es una de las claves para generar nuevos servicios y negocios digitales. Como se discutió extensamente en la Sección 5, crear experticia en el sector requiere inversiones significativas en análisis de datos, algoritmos, etc., en todos los niveles de la educación y de las ciencias.
2. Un programa de I+D+i a diez años. Este programa debe reunir a los actores clave tales como universidades, centros de investigación, empresas y proveedores de financiamiento.

Argentina debe responder a sus necesidades de desarrollo productivo y a sus oportunidades en el campo de las TIC con un programa que posibilite la creación de clusters de competencias relevantes para alentar el crecimiento y la inversión en innovaciones. En este punto, es de gran relevancia considerar que estamos en medio de lo que hemos llamado la segunda ola de la nueva revolución de las TIC, caracterizada por la centralidad de las tecnologías relacionadas con procesamiento de datos masivos, incluyendo su generación a través de múltiples tipos de dispositivos -fijos, móviles-, análisis, seguridad, etc., tal como se ha discutido extensamente en las secciones previas.

En el diagnóstico presentado en la Sección 2, así como en los informes internacionales, se muestra que las empresas del sector TIC tienen dificultades para conseguir trabajadores competentes. A escala global, una de cada tres empresas tiene dificultades para reclutar empleados. En consecuencia, desarrollar clusters de competencias críticas es una precondition para el desarrollo del país y de la economía digital en particular.

A partir del análisis desarrollado en el Libro Blanco, el Proyecto MinCyT 2012-2014 y su actualización en el presente informe queda claro que el éxito de la Argentina requiere:

- Desarrollar capacidades, en profundidad, en procesamiento de datos, en especial alrededor de las temáticas relacionadas con los ejes planteados en el Plan Argentina Innovadora 2020 y de acuerdo a las tendencias tecnológicas predominantes.
 - Asegurar la creación de clusters de competencias críticas en tecnologías clave, incluyendo servicios digitales y contenidos, juegos, seguridad de datos, movilidad y *Big Data*.
 - Creación de redes que integren investigación, aplicación, productización y comercialización vinculando universidades, centros científicos y tecnológicos, organismos públicos y empresas.
 - Integrar a las TIC como parte importante en las políticas de educación general.
3. Un programa de financiamiento abierto y diverso que cubra comprehensivamente las necesidades de *startups* y el crecimiento de las empresas.

Más allá de las dificultades circunstanciales para obtener inversión y financiamiento externo, es necesario establecer un plan de estímulos al desarrollo de productos y servicios con foco en la innovación. El sector SSI en particular ha demostrado capacidad de crecer y exportar a pesar de la crisis global que cerró los mercados europeos.

El desarrollo de la infraestructura de comunicaciones y la nueva ley Argentina Digital establecen un marco para desarrollar productos y servicios innovadores para el mercado local con posibilidades de exportación, al menos a escala regional. Por otro lado, es muy importante analizar las posibilidades que abren las alianzas establecidas con potencias emergentes como China, Rusia, India, por tratarse de enormes mercados en expansión y, en especial en el caso chino, con capacidades complementarias con nuestro país. En este punto, es interesante comprender que nuestros desarrolladores de *software* superan en términos generales la media de sus colegas chinos¹⁷. Es necesario y factible financiar el desarrollo de innovaciones tecnológicas utilizando el poder de demanda del Estado, de manera amplia y comprehensiva, focalizando en la construcción de una amplia plataforma de servicios públicos, apuntando a conseguir el liderazgo regio-

¹⁷ El autor ha corroborado esto en varias ocasiones con motivo de colaboraciones con empresas chinas en el contexto del desarrollo de *software* para la TV Digital.

nal en ese aspecto y estableciendo colaboraciones con las nuevas potencias, en especial con China.

Por otra parte, debe intensificarse el estímulo con capital semilla para la creación de nuevas empresas. La experiencia del programa para emprendedores del FONSOFT muestra que existen potencialidades importantes en el sector SSI, así como en la electrónica y las comunicaciones, sectores en los que es crucial promover el surgimiento de nuevas empresas. A la vez, es necesario estimular el desarrollo de capital de riesgo privado, algo que ha comenzado a crecer aunque de manera aún insuficiente.

Por otra parte, es necesaria una política que ayude a consolidar a las empresas locales de SSI que han crecido notablemente durante la última década, impulsadas por la Ley de *Software* y otras políticas públicas. Algunas de ellas ya han logrado internacionalizarse (por ejemplo, Globant y ASSA) pero hay una camada que está avanzando en esa dirección y requiere de estímulos específicos para estabilizarse como jugadores internacionales.

Por otro lado, es importante que dicho crecimiento se realice focalizando en la innovación y no solamente en el desarrollo de proyectos de mediana o baja complejidad, generalmente concebidos y diseñados en el exterior. En ese sentido, la participación en los grandes proyectos públicos constituye una oportunidad para lograrlo. A la vez, es necesaria la inversión a riesgo tanto pública como privada y muy especialmente la protección de la propiedad intelectual de las innovaciones. En este punto, debe considerarse muy seriamente alguna alternativa del tipo de la "*Innovation Box*"¹⁸ implementada en los Países Bajos desde 2010.

4. Un método operativo que posibilite establecer objetivos de largo plazo e implementación rápida. Para definir esta línea de manera más concreta, se podría establecer un grupo experto en el nivel de la Presidencia o la Jefatura de Gabinete con acuerdo parlamentario, con la intención de que pueda operar trascendiendo los cambios de administraciones. La tarea del grupo experto es la de monitorear y facilitar la implementación de las medidas, crear una base de acuerdos que permita la dirección sin interferencia de las operaciones y mantener un método de operación en red.

La perspectiva del programa es de una década, sin embargo, nadie puede prever qué medidas podrían requerirse aún para los próximos pocos años. En consecuencia, la recomendación metodológica inicial consiste en formular medidas cada año o dos años, que sean revisadas y eventualmente reformuladas. El grupo experto y la red que se formará a su alrededor propondrán las nuevas medidas en los años subsiguientes.

¹⁸ <http://bit.ly/1GFoKSX>



**Presidencia
de la Nación**

Ministerio de
Ciencia, Tecnología
e Innovación Productiva



Secretaría de
Planeamiento y Políticas