SMA Superintendencia del Medio Ambiente Gobierno de Chile

Gobierno de Chile gob.cl

# Inteligencia **Ambiental**

## Una Estrategia de Ciencia de Datos al Servicio del Medio Ambiente y las Personas

Superintendencia del Medio Ambiente

Septiembre 2021

**HOJA DE RUTA 2020-2023**

# Créditos

Superintendencia del Medio Ambiente

**División de Seguimiento e Información Ambiental**

**Redacción**

Sebastián Elgueta Alarcón

Cristóbal De La Maza Guzmán

Francisco Donoso Galdames

**Edición**

Jimena Quilodrán Herrera

**Diseño**

Patricio De La Rosa Muñoz

**Agradecimientos**

A las personas, equipos y áreas de la SMA que han colaborado en el diseño e implementación de esta estrategia, incluyendo a la División de Seguimiento e Información Ambiental, la División de Fiscalización y Conformidad Ambiental, la Fiscalía, las Oficinas Regionales, el Departamento de Administración y Finanzas, la Oficina de Comunicaciones y la Oficina de Transparencia, Participación y Atención Ciudadana.

A las personas, organismos e instituciones externas a la SMA que han apoyado la implementación de esta estrategia de #InteligenciaAmbiental.

Inteligencia Ambiental. Estrategia 2020 - 2023 por **Superintendencia del Medio Ambiente de Chile** se distribuye bajo una **Licencia Creative Commons Atribución-NoComercial-SinDerivadas 4.0 Internacional.**
Basada en una obra en **https://portal.sma.gob.cl/publicaciones/IA/Inteligencia_Ambiental_SMA.pdf.**

# Contenidos


**1 Introducción** 4

**2 Contexto** 5

2.1 ¿Qué es la SMA? 5

2.2 La SMA tiene un desafío Big Data 6

2.3 La SMA con un foco digital y orientada a datos 8

2.4 La SMA como un organismo vivo 13

**3 Inteligencia Ambiental** 18

3.1 Objetivo general 18

3.2 Objetivos específicos 19

3.3 Principios 20

3.4 Pilares de la estrategia 21

3.5 Alianzas 24

**4 La Plataforma de Inteligencia de Datos (PID)** 28

4.1 Modelo conceptual 29

4.2 Captura de datos 30

4.3 Analítica avanzada 39

4.4 Acciones 44

4.5 Aprendizaje 52

**5 Hoja de ruta** 54

5.1 Captura de datos 55

5.2 Analítica avanzada 57

5.3 Acciones 59

**6 Referencias** 61

# 1. Introducción

El objetivo de la Superintendencia del Medio Ambiente (SMA), como principal órgano del estado encargado de asegurar el cumplimiento de la normativa ambiental, es la protección de la salud de las personas, y de la diversidad de ecosistemas que conforman el invaluable patrimonio natural nacional. Para ejercer estas funciones la SMA debe lidiar con múltiples normativas ambientales y sus especificidades, tener una cobertura en todo el territorio nacional, así como dar solución a permanentes contingencias que requiere de una reacción rápida y efectiva para mitigar efectos ambientales indeseados. En este sentido, la presente **Estrategia de Inteligencia Ambiental** tiene como objetivo mejorar la gestión institucional de la Superintendencia del Medio Ambiente, a través del **uso intensivo de tecnologías, la ciencia de datos y la inteligencia artificial,** de modo de contribuir a robustecer la fiscalización, fortalecer la potestad sancionatoria y el incentivo al cumplimiento; potenciar el seguimiento y el acceso a información pública en materia ambiental.

En particular, la Estrategia busca fortalecer la detección temprana de desviaciones o irregularidades de la normativa ambiental, y la consecuente adopción oportuna de las medidas o acciones que correspondan. Adicionalmente, en el contexto de la crisis ambiental en que vivimos, amplificada por el cambio climático, se busca generar y consolidar información y conocimiento sobre el estado del medio ambiente y su evolución en el tiempo, para aportar a la toma de decisiones y la mejora continua de políticas públicas.

De este modo, se propone el desarrollo de productos focalizados tanto para promover el cumplimiento de la normativa ambiental, como para realizar un seguimiento intensivo de los ecosistemas del país. Esto incluye aplicaciones de reporte electrónico de compromisos ambientales, verificaciones automática del cumplimiento normativo, conexión en línea de sensores remotos, uso y análisis de imágenes satelitales, desarrollo de modelos predictivos, fiscalizaciones remotas y masivas, entre otras.

El presente documento se estructura en cuatro secciones. En primer lugar, se expone un capítulo de contexto donde se explica las funciones de la Superintendencia y el gran desafío que tiene, como principal institución de fiscalización y cumplimiento ambiental del país. En segundo lugar, se presenta el concepto general de la Estrategia de Inteligencia Ambiental, sus objetivos, definiciones generales, principios y pilares fundamentales. En tercer lugar, se detalla el desarrollo de la Plataforma de Inteligencia de Datos (PID), como el eje central de la Estrategia, desglosando su marco conceptual y componentes. Finalmente, se ilustra una hoja de ruta 2020-2023 con proyectos o iniciativas para lograr los objetivos planteados.

En síntesis, **la estrategia plantea el desarrollo de una inteligencia colectiva híbrida, conformada por personas y tecnologías, que permitan optimizar la gestión de la SMA a fin de los desafíos relacionados con la protección del medio ambiente y la salud de la población.**

# 2. Contexto

## 2.1 ¿Qué es la SMA?

La Superintendencia del Medio Ambiente (SMA) es un servicio público descentralizado, creado el año 2010 en virtud de la Ley N° 20.417[1] y que comenzó a operar con plenas competencias en diciembre del año 2012. Le corresponde a la SMA de forma exclusiva ejecutar, organizar y coordinar el seguimiento y fiscalización de las Resoluciones de Calificación Ambiental, de las medidas de los Planes de Prevención y/o de Descontaminación Ambiental, del contenido de las Normas de Calidad Ambiental y Normas de Emisión, y de los Planes de Manejo, cuando corresponda, y de todos aquellos otros instrumentos de carácter ambiental que establezca la ley. Las funciones de fiscalización de la Superintendencia pueden desarrollarse a través de sus propios funcionarios, mediante la encomendación a Organismos Sectoriales, o por medio de terceros debidamente acreditados y autorizados por la institución.

Adicionalmente, la institución debe promover e incentivar el cumplimiento de la normativa ambiental, orientando a sus regulados en la comprensión de sus obligaciones, y en caso de detectarse infracciones, imponer las sanciones que correspondan. También es responsable de gestionar las denuncias ambientales presentadas por la ciudadanía y autoridades, aplicando sus facultades.

VISIÓN

*"Un país consciente y responsable, que protege el medio ambiente y se desarrolla de manera sostenible".*

MISIÓN

*"Proteger el medio ambiente y la salud de la población, asegurando el cumplimiento de la normativa ambiental".*

[1] https://www.bcn.cl/leychile/navegar?idNorma=1010459

## 2.2 La SMA tiene un desafío Big Data

La verificación del cumplimiento de la normativa ambiental y el monitoreo del estado del medio ambiente es un desafío que requiere integrar y procesar una gran cantidad de datos proveniente de distintas fuentes con objeto de activar respuestas de manera oportuna, eficaz y eficiente. De esta manera la SMA debe contar con un alto estándar tecnológico y capacidades de detección temprana de desviaciones e incumplimientos en distintas unidades fiscalizables, de acuerdo con lo establecido en el artículo 33 de su Ley Orgánica.

**Artículo 33 de la Ley N° 20.417:**

*La Superintendencia contará con una plataforma electrónica que le permita la adecuada administración del Sistema Nacional de Información de Fiscalización Ambiental y su aplicación útil para la detección temprana de desviaciones o irregularidades y la consecuente adopción oportuna de las medidas o acciones que correspondan.*

*Dicha plataforma deberá permitir, además, mantener un historial sistematizado y actualizado de las acciones de fiscalización realizadas respecto de cada sujeto fiscalizado y sus resultados, así como de los procesos sancionatorios de que sean objeto y los resultados de ambos, que permita obtener información útil para la priorización de futuras fiscalizaciones y para la resolución de futuros procedimientos sancionatorios.*

De acuerdo con lo establecido en la Resolución Exenta SMA N° 1184/2015, se denomina Unidad Fiscalizable a la “unidad física en la que se desarrollan obras, acciones o procesos, relacionados entre sí y que se encuentran regulados por uno o más instrumentos de carácter ambiental de competencia de la Superintendencia”. De esta forma, **existen más de 16.000 Unidades Fiscalizables en el territorio que deben cumplir con la normativa ambiental** (Figura 1) y que incluyen desde pequeñas instalaciones, tales como panaderías o locales nocturnos, hasta grandes establecimientos asociados a actividades económicas como faenas mineras o procesos industriales. Muchas de ellas deben cumplir con más de un instrumento normativo a la vez y reportar una gran cantidad de datos, análisis y mediciones a la SMA de forma periódica, según obligaciones y compromisos.

Por otro lado, se reciben **miles de denuncias ciudadanas al año,** relacionadas con diversos motivos que incluyen por ejemplo ruidos y olores molestos, afectación de ecosistemas, y desarrollo de actividades sin permisos ambientales. En diciembre de 2020 se habilitó el Portal Ciudadano de Denuncias, que permitió digitalizar el ingreso de éstas, ya que antes sólo se recibían en formato físico. Adicionalmente, se recopilan **antecedentes de distintos servicios sectoriales con competencia ambiental,** que incluyen denuncias, actividades de fiscalización e información consolidada en sus plataformas electrónicas.

***Figura 1:*** *Hay más de 16.000 Unidades Fiscalizables en el país que deben cumplir la normativa ambiental.*

En este contexto, la SMA debe gestionar las *4V's del Big Data:* un conjunto masivo de datos **(Volumen),** generados de forma continua **(Velocidad),** con características diversas en cuanto a tipología y fuentes de obtención **(Variabilidad)** y que deben ser correctos y precisos para ser utilizados **(Veracidad).** Así, con un trabajo mancomunado que integre desarrollo tecnológico basado en datos (aplicaciones, modelos matemáticos y estadísticos, procesos automatizados) con la experiencia basal de los profesionales especialistas en materias de fiscalización y sanción, se podría disminuir los tiempos de gestión, identificar masiva y certeramente desviaciones a la normativa ambiental, y más aún, fortalecer el carácter preventivo de la institución con acciones que eviten impactos ambientales y sociales relevantes.

Por lo tanto, responder de forma oportuna, eficaz y eficiente requiere de una gestión óptima de la información. La cantidad masiva de datos disponibles, la capacidad de procesamiento de información a costos accesibles y las posibilidades de utilización y desarrollo de algoritmos de última generación, hacen indispensable priorizar la incorporación de la inteligencia artificial y la ciencia de datos como herramientas esenciales para cumplir los objetivos institucionales (MinCiencia, 2021). Estas herramientas permiten cumplir los objetivos, implementar mejores políticas públicas y tomar mejores decisiones (OECD, 2019b).

La adopción de este tipo de tecnologías en la Superintendencia del Medio Ambiente irá en directo beneficio del medio ambiente y de las personas. Su implementación busca mejorar el nivel de servicio a la comunidad y proponer soluciones a los grandes desafíos ambientales del país como el cambio climático y sus efectos nocivos, mediante el desarrollo de herramientas que van a permitir liberar al personal de tareas automatizables y expandir su productividad mediante sistemas inteligentes (MinCiencia, 2021).

## 2.3 La SMA con un foco digital y orientada a datos

El crecimiento de los datos y las tecnologías digitales están transformando a la sociedad y a sus organizaciones, con enormes implicancias para las operaciones cotidianas de organismos públicos, quienes deben adaptarse para entregar una mejor calidad de servicios, de forma simple y efectiva. En este contexto, un **"Gobierno Digital"** se caracteriza por potenciar las siguientes dimensiones enfocadas en generar valor público, a través del uso de tecnologías (OECD, 2018):

- **Digital por diseño:** que transforme sus procesos y cultura para entregar servicios digitales.
- **Orientado a datos *(Data driven)*:** que considere los datos como un activo estratégico.
- **Interconectado:** que colabore con el sector público, privado, la academia y la sociedad civil.
- **Abierto:** que potencie la transparencia y responsabilidad.
- **Centrado en las personas:** que considere las reales necesidades de los ciudadanos.
- **Proactivo:** que se anticipe y responda de forma rápida.

### Transformación digital en Chile

En el ámbito nacional, las 6 dimensiones planteadas por la OECD fueron incorporadas como principios de la Estrategia de Transformación Digital del Estado 2018-2022 (Gobierno Digital, 2019), la cual cuenta con 6 líneas de acción:

- ✔ **Identidad Digital:** que busca posicionar a la "Clave Única" como el exclusivo medio de identificación digital en el Estado (firma electrónica, gestión de datos personal y buzón de notificación).
- ✔ **Cero Filas:** que está orientada a digitalizar los trámites públicos y la interoperabilidad de sus instituciones.
- ✔ **Cero Papel:** que consiste en digitalizar los procesos internos de las instituciones del Estado, para hacerlos más eficientes y eliminar de forma gradual el uso de papel (expediente y oficina de partes virtual).
- ✔ **Estado Basado en Datos:** que se refiere a fomentar el uso de datos y potenciar la optimización de políticas públicas y la automatización de procesos, mediante el uso de tecnologías emergentes como la inteligencia artificial.
- ✔ **Ciberseguridad:** que se enfoca en la generación de una política nacional para proteger al Estado de las amenazas a la seguridad de la información.
- ✔ **Un Estado que mira al futuro:** que busca instalar como práctica permanente en las instituciones la mirada de futuro en su gestión y entrega de servicios, es decir, instalando una cultura institucional que esté a la vanguardia, utilizando tecnologías emergentes en sus soluciones.

Un organismo público orientado a datos se caracteriza no sólo por reconocerlos como activos, sino también por realizar esfuerzos para facilitar la gestión, intercambio y utilización de ellos, junto con su aplicación para transformar el diseño, la entrega y el monitoreo de sus servicios, además de promover la publicación y el uso de ellos por parte de otros organismos (OECD, 2019a). Existen muchas oportunidades para utilizar los datos en la generación de información y conocimiento, mediante tecnologías emergentes asociadas a la ciencia de datos e inteligencia artificial y la aplicación de otras tecnologías disruptivas como internet de las cosas (Internet of Things - IOT), computación en la nube (cloud computing), cadenas de bloques (blockchain) o realidad aumentada (augmented reality - AR).

**¿Qué entendemos por Ciencia de Datos?**

Existen múltiples definiciones para este concepto. Según Provost & Fawcett (2013), "la ciencia de datos es un campo interdisciplinario que incluye una serie de principios, procesos y técnicas para la extracción de conocimiento desde el análisis de los datos". Por su parte, Van der Aalst (2011) propone que "es un campo interdisciplinario que busca generar valor a partir de los datos, que incluye procesos de extracción, preparación, exploración, transformación, almacenamiento y visualización de datos".

La ciencia de datos requiere de distintos conocimientos y habilidades. Según el diagrama de Venn propuesto por Drew Conway en una conferencia del año 2010, un científico de datos debe contar con habilidades informáticas para extraer y tratar los datos, conocimiento matemático para tratar y analizar los datos y experiencia relevante en el dominio de interés, para interpretarlos, como en el caso de conocimiento sobre ciencias y normativas ambientales (Figura 2).

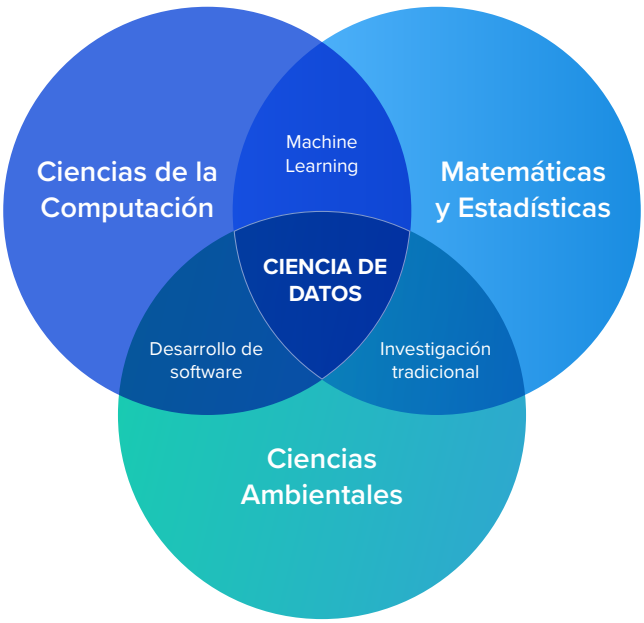



***Figura 2:*** *Ciencia de Datos Ambientales (adaptado de Palmer, 2015).*

## ¿Qué entendemos por Inteligencia Artificial (IA)?

El grupo de expertos de la Organización para la Cooperación y el Desarrollo Económico, define IA como "un sistema computacional que puede, para un determinado conjunto de objetivos definidos por humanos, hacer predicciones y recomendaciones o tomar decisiones que influyen en entornos reales o virtuales. Los sistemas de IA están diseñados para operar con distintos niveles de autonomía e intentan realizar funciones cognitivas similares a las de los humanos".

¿Cuál es la diferencia entre Inteligencia Artificial (IA), Machine Learning (ML) y Deep Learning (DL)? En muchas organizaciones se utilizan estos conceptos como sinónimos, pero, si bien están relacionados, hay diferencias relevantes (Figura 3). Existe un amplio consenso que IA es un concepto que se refiere a la capacidad de las computadoras de emular el aprendizaje humano, mientras que ML es una forma de IA que permite a un sistema aprender de los datos en lugar de aprender mediante la programación explícita. En tanto, DL se refiere a técnicas específicas de ML que utilizan redes neuronales, en múltiples niveles, generalmente utilizados para resolver problemas de alta complejidad (OECD, 2019c).

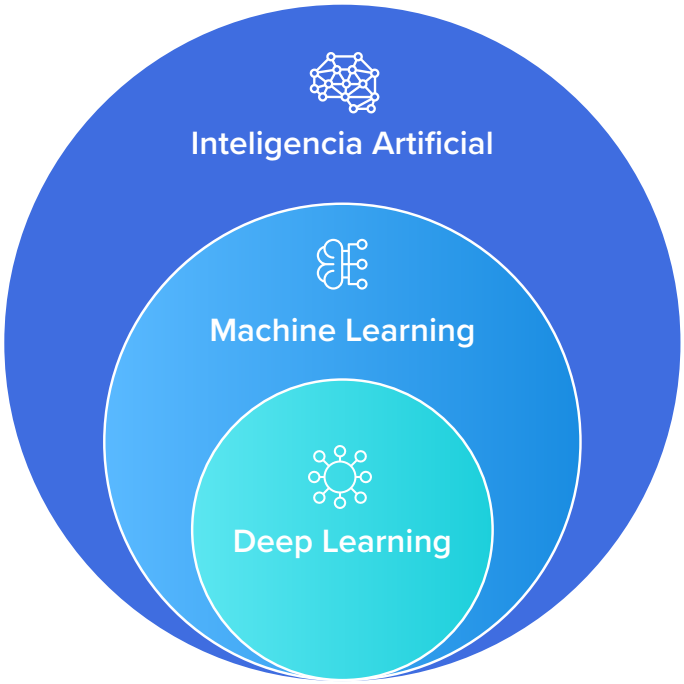



***Figura 3:*** *Relación entre IA, ML y DL (adaptado de OECD, 2019c)*

## La inteligencia artificial en el sector público

Durante los últimos años se han publicado diversos documentos que contienen experiencias y recomendaciones para desarrollar estrategias de ciencia de datos e inteligencia artificial en organismos de Gobierno. A nivel internacional destacan los trabajos realizados por UNESCO (2019), Deloitte (Eggers et al, 2019) y OECD (2019b), mientras que, a nivel local, un hito relevante es la publicación de la Política Nacional de Inteligencia Artificial de Chile (MinCiencia, 2021), durante el segundo semestre de 2021.

Sin duda, la inteligencia artificial puede potenciar el trabajo que realiza la SMA. Si bien existe una brecha importante entre las capacidades institucionales actuales y las necesarias para innovar en esta área, la presente estrategia permitirá generar las herramientas fundamentales para la incorporación efectiva de estas tecnologías y ponerse a la vanguardia en el desarrollo de herramientas aplicadas en la temática ambiental en un futuro próximo, tal como lo han realizado agencias estatales de otros países (NITI Aayog, 2018).

En particular, existen diversas aplicaciones de inteligencia artificial que se pueden implementar en organismos públicos como la SMA, incluyendo el procesamiento de lenguaje natural para extraer conocimiento desde grandes colecciones de textos, el desarrollo de modelos predictivos a través de técnicas de *machine learning*, y el reconocimiento de voz en llamados telefónicos para mejorar la atención ciudadana, entre otros (Figura 4).

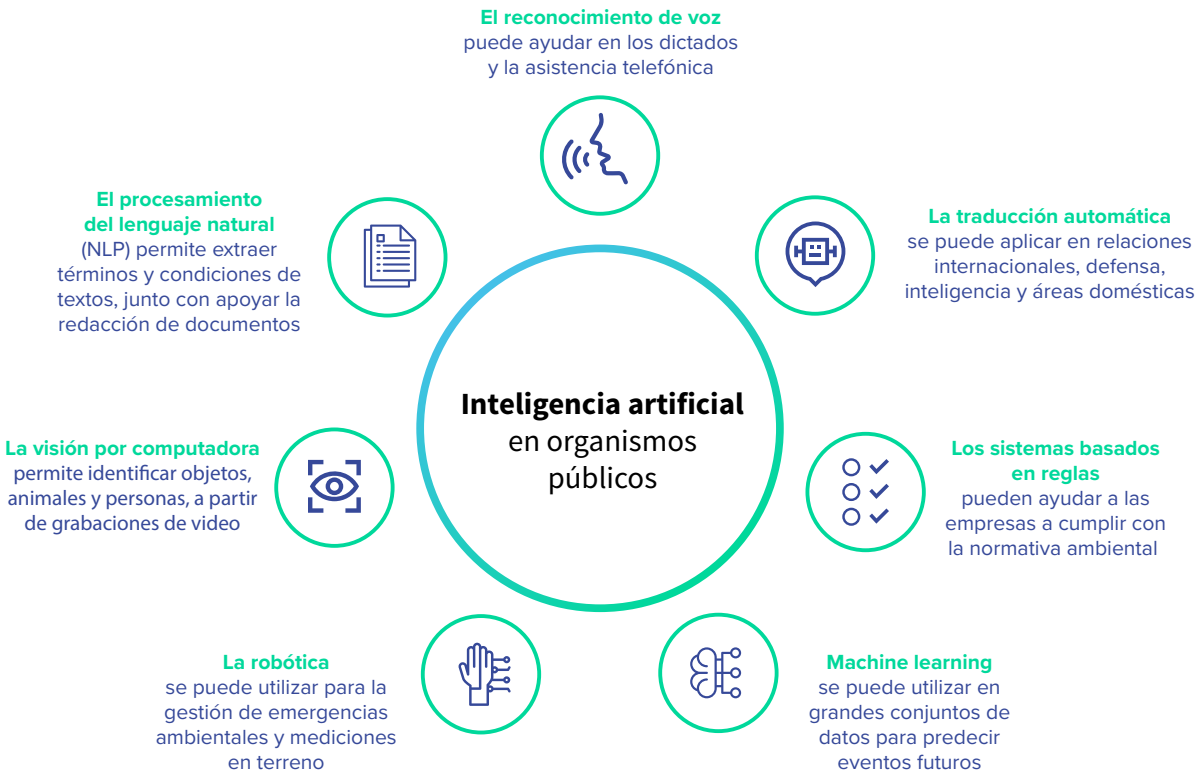



***Figura 4:*** *Aplicaciones posibles de Inteligencia Artificial en Organismos Públicos (adaptado de Eggers et al, 2019).*

Precisamente, durante los años 2015 y 2019, para convertir a la SMA en un organismo orientado a datos, se implementó la estrategia de transformación digital denominada "**Cumplimiento Ambiental 2.0**", que se publicó el 2017[2] (Figura 5). Esta estrategia se basó en el enfoque propuesto por la Agencia Ambiental de Estados Unidos (EPA, 2014), denominado "Next Generation Compliance", a través del desarrollo de 5 ejes estratégicos: monitoreo avanzado a través de sensores remotos e imágenes satelitales; reporte electrónico de compromisos y obligaciones; transparencia de la información para generar un efecto disuasivo; innovación aplicada a casos de fiscalización y sanción; además de promover el uso de tecnologías en la regulación ambiental. De esta forma, la presente Estrategia de Inteligencia Ambiental se construye a partir de las capacidades y avances logrados anteriormente.

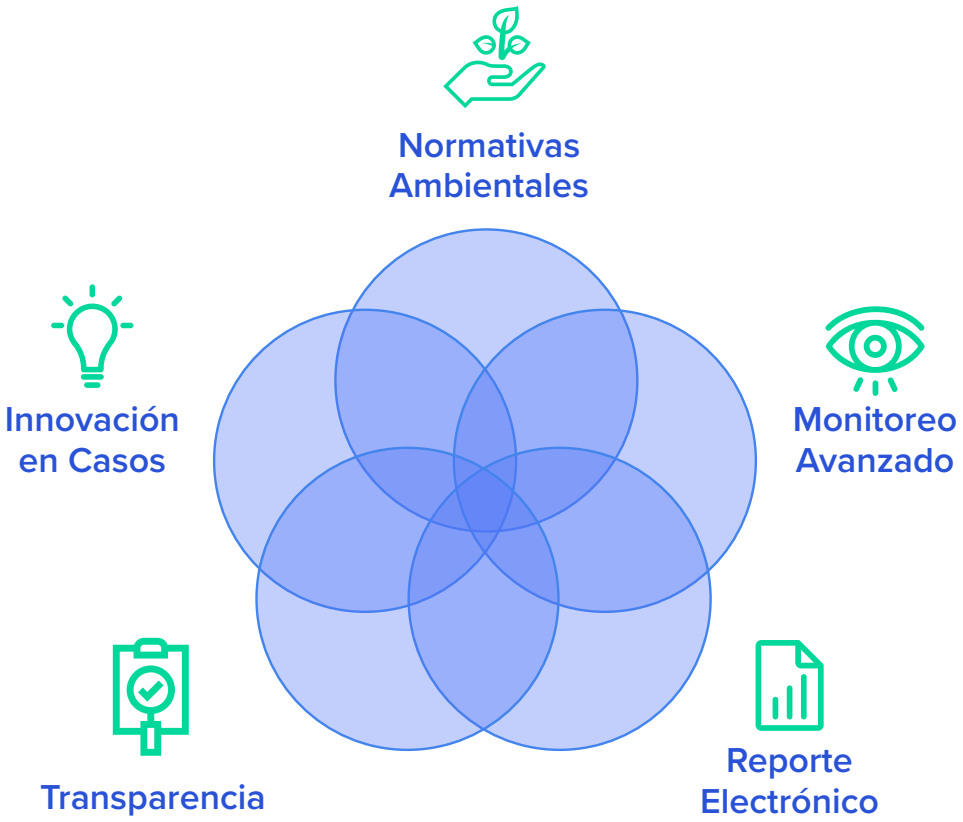



***Figura 5:*** *Ejes de la estrategia 2015-2019 de la SMA, "Cumplimiento Ambiental 2.0".*

[2] https://portal.sma.gob.cl/index.php/publicaciones/

## 2.4 La SMA como organismo vivo

Si bien los avances en la ciencia de datos e inteligencia artificial ofrecen grandes beneficios, optimizando tiempos y mejorando la eficiencia de diversos procesos, las máquinas no son capaces aún de pensar y razonar como los seres humanos. Por lo tanto, las personas seguirán ocupando roles centrales en las organizaciones y sus funciones serán complementadas con tecnologías que permitirán automatizar tareas y facilitar la toma de decisiones.

Ésta simbiosis se conoce como **"Inteligencia Híbrida",** definida por Dellermann et al. (2019a) como "*sistemas que tienen la capacidad de lograr objetivos complejos, mediante la combinación de inteligencia humana y artificial, obteniendo resultados superiores a los que cada uno podría lograr por separado y mejorando de manera continua por medio del aprendizaje mutuo*".

Por un lado, la inteligencia humana se caracteriza por su flexibilidad y, por otro lado, las máquinas son particularmente efectivas para resolver tareas repetitivas, que requieren procesamiento rápido de grandes cantidades de datos para reconocer patrones complejos, por lo cual se trata de capacidades complementarias y sinérgicas que permiten lograr mejores resultados (Dellermann et al., 2019b).

La inteligencia artificial fue concebida en conexión con las ciencias de la administración, para luego derivar a otras ciencias como las cognitivas o de la computación (Simon, 1991b). Durante los últimos años, en el desarrollo de la inteligencia artificial se ha enfatizado su foco de aplicación original, incorporando elementos asociados a potenciar el desempeño de organizaciones, en el marco de lo que se conoce como "**Inteligencia Colectiva"** (Wolley et al, 2010), que propone que grupos de individuos sin mayor experiencia, actuando de manera coordinada, mediante la tecnología apropiada, pueden superar el desempeño de expertos actuando por separado (concepto referido usualmente como sabiduría en grupo) (Dellermann et al., 2019b).

De este modo, se genera un circuito de aprendizaje, en el cual la inteligencia artificial contribuye a mejorar decisiones y acciones, a través de la automatización de tareas y la provisión de predicciones, mientras que la inteligencia humana permite generar nuevos datos, entrenar modelos analíticos y construir herramientas tecnológicas. Recientemente a partir de lo anterior, se ha propuesto el uso de un concepto más amplio denominado **"Inteligencia Colectiva Híbrida",** que se refiere a la integración de individuos, colectivos de personas y tecnologías de inteligencia artificial, que corresponde al enfoque que se propone adoptar en esta estrategia (Peeters et al., 2021).

De esta manera, se plantea potenciar el desarrollo de sistemas inteligentes, controlados y retroalimentados por personas, que cuentan con experiencia y conocimiento de los procesos de la institución (Figura 6). **Coordinar el conocimiento y las capacidades humanas colectivas, con la precisión y las capacidades analíticas de sistemas inteligentes, permitirá que la SMA realice mejor sus funciones, aumentando significativamente su cobertura de acción para prevenir impactos ambientales.**

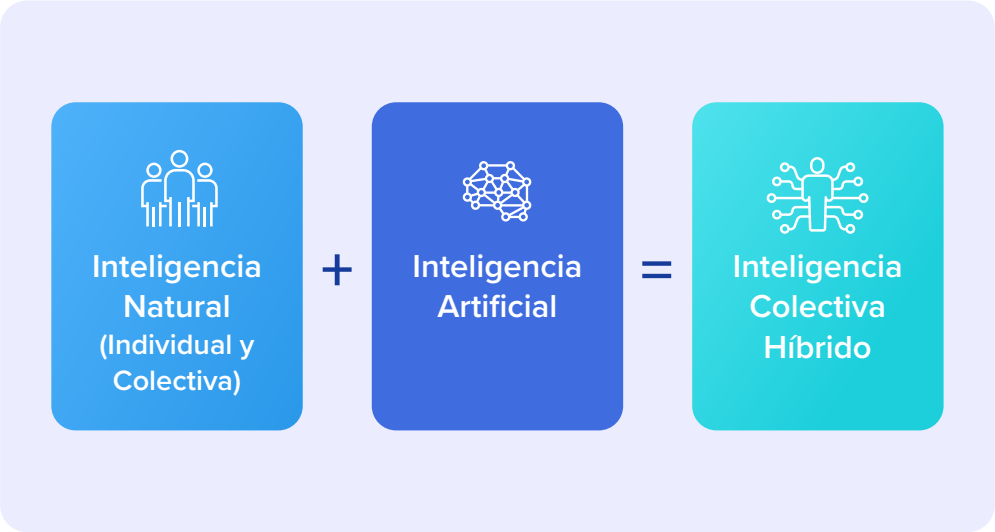



***Figura 6:*** *Inteligencia colectiva híbrida.*

El desarrollo de una **inteligencia colectiva híbrida** en la SMA para optimizar sus funciones está en total sintonía con el paradigma moderno de una organización ágil (Figura 7). Este tipo de organizaciones son estables y dinámicas, conformadas por una red de equipos que interactúan en una cultura centrada en las personas, que operan mediante ciclos rápidos de aprendizajes y decisiones facilitadas por las tecnologías, para responder de manera ágil, tal como actúa un organismo vivo, en constante modificación al activar los procesos productivos que le son propios (Maturana & Varela, 1987).

***Figura 7:*** *Nuevo paradigma de organización ágil (adaptado de Aghina et al, 2018).*

Según lo planteado, mientras las personas destacan en escenarios que requieren intuición, creatividad, emoción, juicio o empatía (OECD, 2019b), las máquinas han mostrado un desempeño superior en tareas que requieren de decisiones lógicas (Kahneman, 2002; Dellermann et al., 2019b). Tanto en situaciones en que la inteligencia artificial provee de conocimiento que luego es evaluado por un tomador de decisión, o en el caso inverso cuando es el humano el que genera, entrena y valida un algoritmo, poniendo a disposición de la máquina conocimiento para resolver un problema, ambos funcionan como colaboradores, como profesores que se enseñan mutuamente (Figura 8) (Dellermann et al., 2019b).

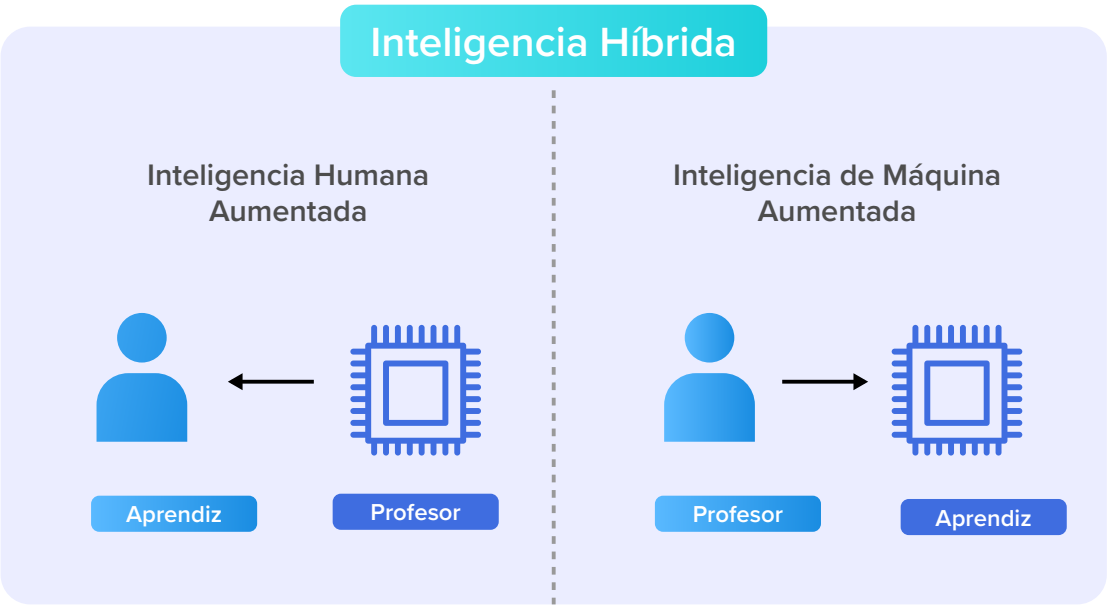
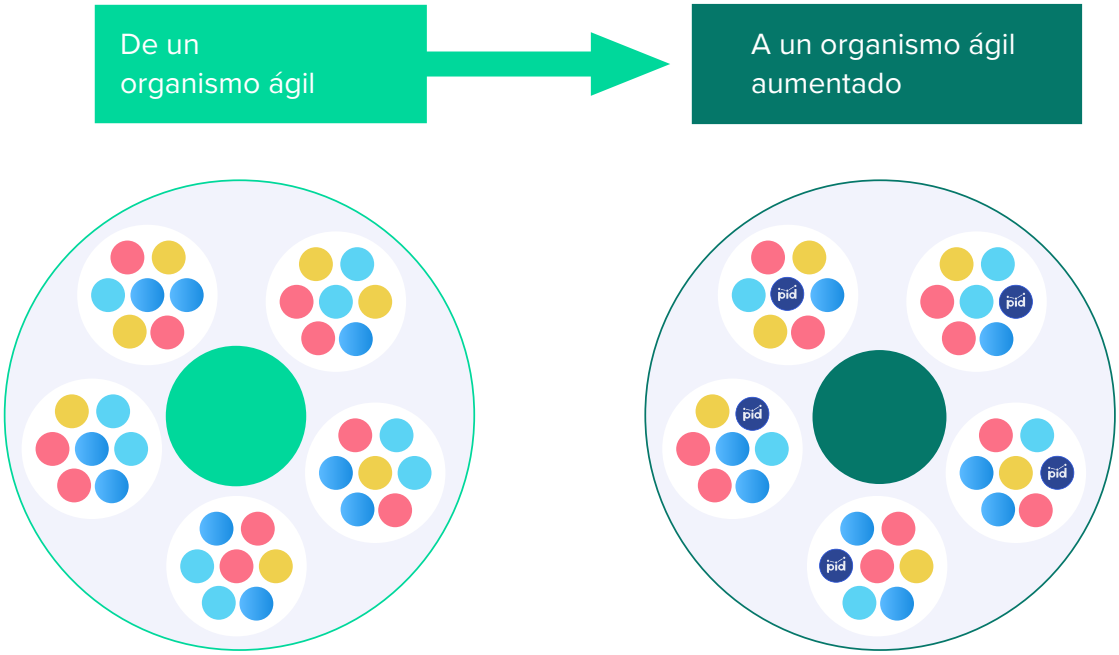



***Figura 8:*** *Inteligencia híbrida y los roles de las personas y máquinas (en base a Dellermann et al., 2019b).*

## Sinergias

Es mediante esta colaboración, que humanos y máquinas pueden evolucionar y alcanzar niveles de desempeño superiores. En aplicaciones reales, en contextos dinámicos de alta incertidumbre, las capacidades combinadas de ambos son indispensables (Dellermann et al., 2019b). En cambio, en algunos escenarios donde las personas experimentan sesgos cognitivos o fatiga en el proceso de toma de decisiones (Simon, 1997; Kahneman et al, 1982), incluso podría ser conveniente anular la decisión preferida por la inteligencia humana y favorecer la recomendada por la máquina.

El trabajo mancomunado de equipos de máquinas y humanos, coordinados de manera adecuada utilizando el estado del arte en ciencias del comportamiento permite un salto en productividad, creando nuevas soluciones a problemas ambientales complejos y facilitando una cultura flexible y dinámica que actúe de manera cohesionada (NSTC, 2019; OECD, 2019b). Equipos híbridos pueden darse en distintas configuraciones, desde pares de individuos, pequeños grupos, hasta grupos numerosos. Al potenciar la coordinación de estos equipos se propone evolucionar de un organismo ágil a un organismo ágil aumentado, mediante el desarrollo de una Plataforma de Inteligencia de Datos (PID) (Figura 9).




PID: Plataforma de Inteligencia de Datos

***Figura 9:*** *Nuevo paradigma de organización ágil (adaptado de Aghina et al, 2018).*

Cuando nos referimos a un organismo aumentado, algunos aspectos claves a considerar corresponden al fomento de relaciones basadas en la colaboración, el intercambio de ideas e información, el desarrollo de talentos y el uso de tecnologías que faciliten todas sus funciones (Aghina et al, 2018). En este contexto, la tecnología perfectamente integrada es fundamental para todos los aspectos de una organización, como un medio para generar valor y adaptarse a los cambios, en contraste con las organizaciones tradicionales, donde las tecnologías se consideran capacidades de soporte, que ofrecen servicios, plataformas o herramientas específicas al resto de la organización, según lo definido por las prioridades, los recursos y el presupuesto.

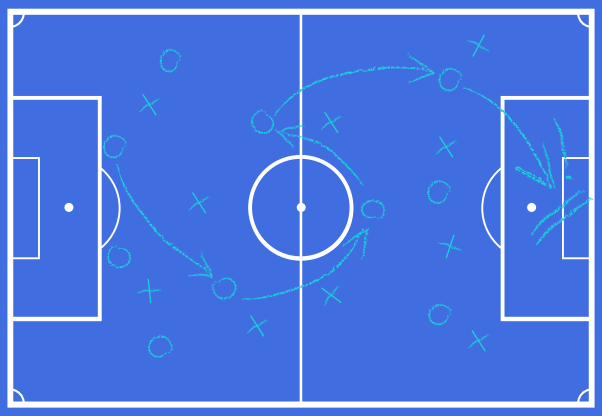
Para alcanzar estos resultados se debe contar con una red de equipos empoderados y autónomos que operen con altos estándares de desempeño profesional, responsabilidad, conocimiento, transparencia y colaboración, para lo cual es fundamental crear un ecosistema de innovación que permita el intercambio de experiencias y conocimientos (Aghina et al, 2018). El siguiente paso será integrar a la SMA en el ecosistema digital, potenciando la transferencia tecnológica entre la academia, el sector público y privado, motivando el desarrollo de capacidades, tanto dentro, como fuera del aparato estatal.

Para el éxito de este proceso, es indispensable disponer de sistemas amigables, que posibiliten la interacción humano-máquina con el menor esfuerzo posible, para su uso en diferentes contextos y en diferentes aplicaciones, favoreciendo sistemas con interfaces y paneles de control idóneos, y herramientas para su uso en diversos escenarios, desde sistemas estacionarios como un computador o sala de control, a otros móviles y usables en terreno como tabletas, anteojos o relojes; que puedan asistir al usuario en el ejercicio de sus tareas potenciando su atención, percepción y memoria, además de aportar recomendaciones útiles cuando se requiera, incluso sin necesidad de realizar consultas (NSTC, 2019).

El desarrollo de esta **inteligencia colectiva híbrida** en la SMA permitirá procesar grandes cantidades de datos de manera eficiente, evaluar un mayor número de soluciones alternativas, identificar incumplimientos en base a patrones no perceptibles al ojo humano, estructurar información en base a conocimiento experto, reconocer fácilmente situaciones similares experimentadas en el pasado, diseñar estrategias de manera automática, anticipando acciones a problemas que se presenten, y en general prestar servicios con mayor oportunidad a la comunidad (OECD, 2019b). Cabe destacar que todo lo anterior debe estar completamente alineado con la Estrategia Institucional, mediante una dirección clara y coherente (Aghina et al, 2018).

**“La estrategia, no la tecnología, es el motor de la transformación digital de una organización”**
(Kane et al, 2015).

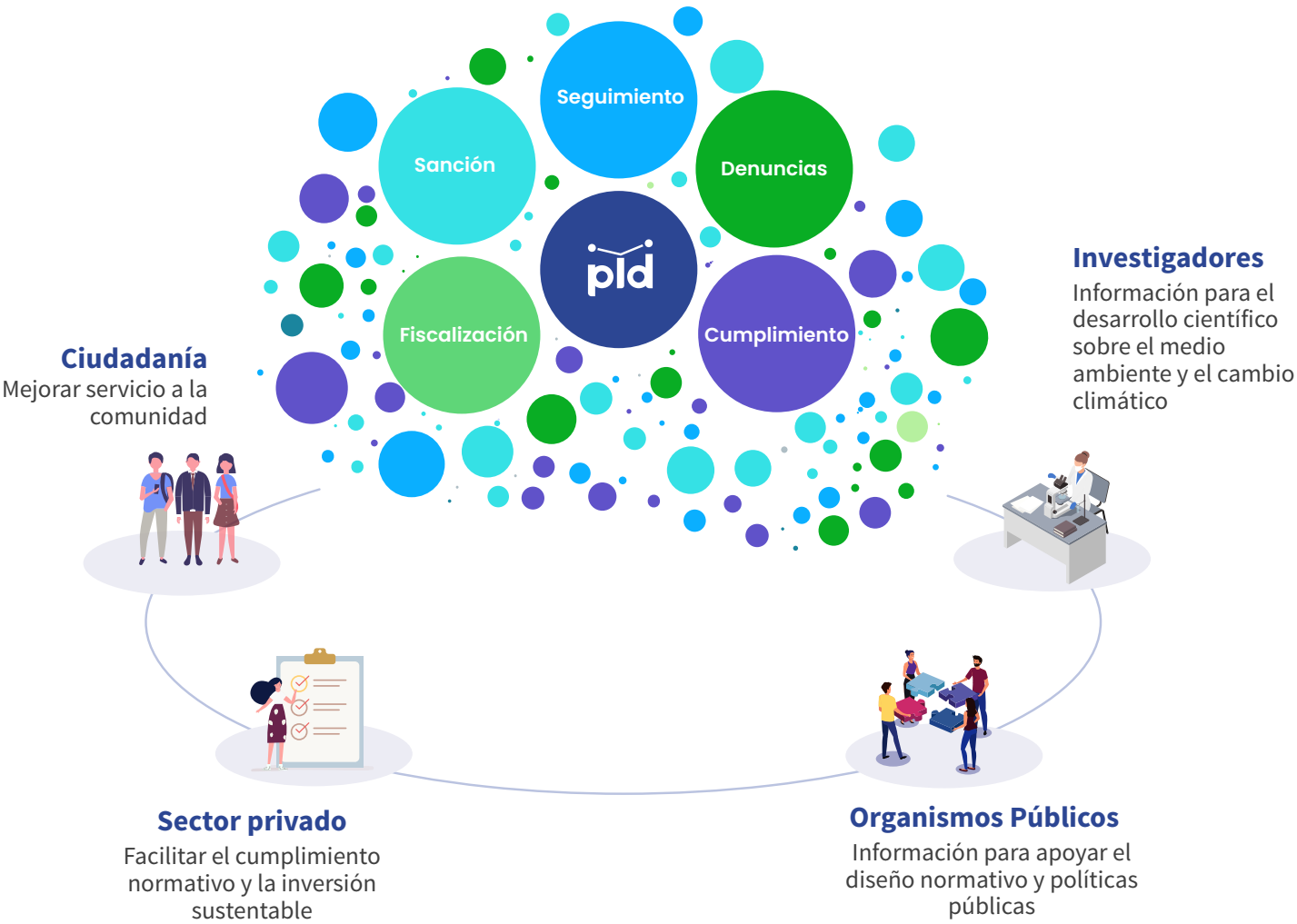
Según este estudio, basado en las respuestas de 4.800 ejecutivos de 27 industrias y 129 países, una organización no se puede reinventar digitalmente, a menos que tenga una estrategia clara, que aborde la integración de personas, procesos y tecnologías, para lo cual es clave fomentar una cultura que se adapte a los cambios y donde la gestión del riesgo juegue un rol clave.



# 3. Inteligencia Ambiental

## 3.1 Objetivo general

La SMA busca optimizar la gestión institucional para mejorar significativamente su nivel de servicio a la comunidad. Esto incluye tanto la verificación del cumplimiento de la normativa ambiental, como la generación de información y conocimiento sobre el estado del medio ambiente, de forma remota, masiva y automatizada, a través del uso de tecnologías de información, la ciencia de datos y la inteligencia artificial, combinadas con la experiencia y conocimientos de los equipos de trabajo de ésta y de otros organismos (Figura 10).




PID: Plataforma de Inteligencia de Datos

***Figura 10:*** *Diagrama conceptual Inteligencia Ambiental*

## 3.2 Objetivos específicos

- **Optimizar la gestión interna:** Automatizando procesos y desarrollando herramientas analíticas que generen respuestas oportunas, eficaces y eficientes.
- **Mejorar la calidad del servicio a la ciudadanía:** Optimizando la gestión de denuncias y facilitando el acceso a la información pública, por medio del Sistema Nacional de Información de Fiscalización Ambiental (SNIFA).
- **Promover el cumplimiento normativo y el comportamiento responsable:** Entregando una retroalimentación temprana y permanente al sector privado sobre sus obligaciones ambientales, previniendo la generación de impactos negativos sobre ecosistemas y poblaciones humanas.
- **Apoyar el diseño normativo y sobre políticas públicas:** Generando información y conocimiento sobre la eficacia de la regulación ambiental y el estado del medio ambiente, retroalimentando a los responsables del diseño de políticas y regulaciones.
- **Fomentar el desarrollo científico:** Facilitando información que genere conocimiento sobre el estado del medio ambiente y los impactos del cambio climático.

## 3.3 Principios

- **Eficacia**: desarrollar servicios y productos tecnológicos que permitan cumplir con la misión institucional.
- **Eficiencia**: impulsar una cultura de innovación y agilidad, orientada a la mejora continua de procesos de la institución, integrando personas y tecnologías.
- **Imparcialidad**: fomentar el cumplimiento de las regulaciones y buenas prácticas para evitar sesgos discriminatorios en el uso de herramientas de inteligencia artificial, respetando los derechos de las personas y generando mecanismos transparentes y trazables.
- **Seguridad**: velar por la integridad, disponibilidad y confidencialidad de la información, además de la protección de datos personales, con énfasis en el diseño, la implementación, el monitoreo y la ejecución de medidas de ciberseguridad de manera integrada con otros agentes del Estado.
- **Sostenibilidad**: impulsar una cultura digital, que minimice el uso del papel y de recursos energéticos, siendo consecuente con la implementación de la visión institucional.
- **Coordinación:** fomentar la interoperabilidad digital con otros organismos, para compartir información de forma eficiente y optimizar procesos conjuntos.
- **Cooperación:** generar alianzas estratégicas y una apertura que fomente la innovación colaborativa, en el ámbito interno y con organismos nacionales e internacionales.
- **Transparencia:** facilitar el acceso a información consolidada, a través de los canales digitales de la institución, de forma activa y fomentando estándares abiertos de datos, software y hardware.

***Figura 11:*** *Planificación colaborativa de proyectos de innovación en la SMA.*

## 3.4 Pilares de la estrategia

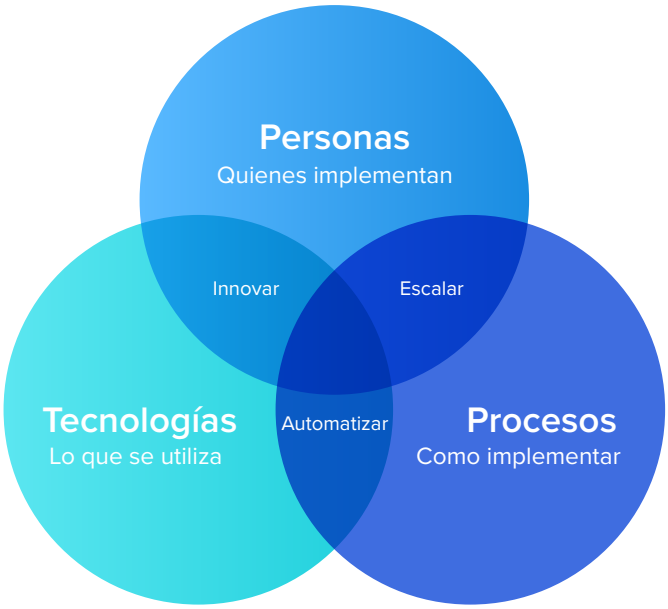



***Figura 12:*** *Pilares de la estrategia (adaptado de Bell, 2005).*

### Personas:

Con toda seguridad se necesitan personas capacitadas y motivadas para el uso de tecnologías de información, ciencia de datos e inteligencia artificial. Considerando que se trata de una transformación cultural disruptiva, es muy importante la gestión del cambio para facilitar el proceso, así como también la identificación, el desarrollo y la retención de talentos. Para esto, se deben fomentar capacitaciones en prácticas de mejora continua y agilidad, empero de una gestión colaborativa, adaptativa y eficiente de los proyectos tecnológicos, a través de iteraciones, incrementos y entregas continuas que agreguen valor permanentemente, mediante productos mínimos viables (Schwaber & Sutherland, 2020). Además, es necesario capacitar a las personas para generar conocimientos técnicos con respecto al uso de herramientas, el desarrollo de modelos analíticos y la adopción de prácticas DevOps para el despliegue y entrega continua de servicios (Kim et al, 2016), entre otros.

Se promoverá la capacitación del personal en herramientas analíticas avanzadas, eliminando sistemáticamente tareas administrativas rutinarias y repetitivas, focalizando el trabajo en tareas de mayor complejidad y valor, alineando los incentivos de desarrollo profesional del funcionariado con las demandas ciudadanas, y facilitando la formación de capacidades para beneficiarse de estas tecnologías (OECD, 2019b). Es imprescindible que el personal adquiera los conocimientos mínimos para comprender las fortalezas y debilidades de la ciencia de datos y la inteligencia artificial, permitiendo un monitoreo efectivo del comportamiento de sistemas que cuenten con mayores grados de automatización y autonomía.

Se debe mantener siempre un control efectivo de máquinas inteligentes autónomas en la toma de decisiones públicas, en caso de fallas o eventos inesperados (OECD, 2019b). Los sistemas autónomos pueden no reaccionar de acuerdo con las normas éticas, morales y sociales que la comunidad ha consensuado formal o informalmente, y que considera indispensables para una mejor convivencia (NSTC, 2019).

Sin lugar a dudas, es fundamental realizar un uso responsable de la ciencia de datos y de la inteligencia artificial. Se deben considerar los temas éticos y legales que correspondan, por parte de todos los equipos involucrados, incluyendo tomadores de decisión (directivos), especialistas jurídicos profesionales encargados de la ejecución técnica de los proyectos (Denis et al, 2021).

A su vez, será necesario promover la formación y atracción de capital humano avanzado de diferentes disciplinas, ya sea de manera permanente, como temporal, con una propuesta de trabajo atractiva. Así, se desarrollará investigación aplicada en temáticas de valor público, en espacios de desarrollo profesional estables y motivantes. Se facilitará el trabajo a distancia y la equidad de género en el desarrollo de sistemas de inteligencia artificial (Eggers et al, 2019; MinCiencia, 2021), en función de los recursos disponibles.

## Tecnologías:

Para la adopción de éstas tecnologías es necesario contar con herramientas informáticas avanzadas, infraestructura adecuada y servicios para la captura, almacenamiento, procesamiento, intercambio y publicación de datos de alta demanda. En particular, es necesario considerar una infraestructura escalable que permita el crecimiento y la seguridad de la información, a través de servicios en la nube y la computación de alto desempeño. También es importante impulsar el desarrollo de una plataforma electrónica modular, que permita la integración de distintos componentes, privilegiando el uso de herramientas de código abierto y promoviendo software y hardware libre.

Todo esto con el objetivo de integrar tecnologías que permitan optimizar el desempeño institucional, con diferentes grados de autonomía, desde asistir en tareas básicas y funcionales, completando las capacidades humanas en tareas que demandan una alta carga cognitiva, como el procesamiento de grandes cantidades de datos, hasta la automatización completa de servicios, liberando completamente a las personas en algunas tareas específicas, que de todos modos deben ser monitoreadas (NSTC, 2019; Eggers et al, 2019).

El uso intensivo de ciencia de datos y tecnologías permitirá aumentar las capacidades individuales, incrementando la capacidad de desarrollar un mayor número de tareas, con mayor calidad y a menor costo (OECD, 2019b). Si bien aún se ve lejana la posibilidad de desarrollar herramientas que requieren de inteligencia general, que permitan el desarrollo de múltiples tareas y en diferentes contextos (NITI Aayog, 2018), no se descarta aportar al desarrollo de sistemas con habilidades propias del pensamiento humano, como la abstracción analítica, la representación de problemas, la planificación estratégica, la generalización de conceptos, la identificación activa de información relevante (Settles, 2009), la utilización de conocimiento experto o su transferencia desde otras disciplinas cuando existe escaza o nula información (NSTC, 2019); con desempeño indistinguible al de la inteligencia humana (Turing, 1950).

Lo anterior permitirá un mayor uso de tecnologías en apoyo a tareas relacionadas con áreas como las ciencias jurídicas, donde la acción humana podría ser irremplazable, realizando tareas como la revisión de documentos legales de manera automática, generando escenarios posibles para la resolución de un litigio, o apoyando la redacción de instrumentos jurídicos en tiempo real. Sin perjuicio de lo anterior, siempre se debe tener en cuenta que el desarrollo de sistemas debe corresponder a necesidades institucionales prioritarias, identificando de manera permanente la disponibilidad de nuevas tecnologías para integrarlas en tiempo y forma al sistema de trabajo (NITI Aayog, 2018).

Por otra parte, las tecnologías utilizadas deben estar diseñadas al servicio de las personas, tanto externas como internas a la institución, con respeto a su integridad física y síquica. La Superintendencia, como desarrollador y comprador de herramientas tecnológicas, puede transformarse en catalizador de nuevos emprendimientos y promotor de estándares éticos de desarrollo en la industria, resguardando en el caso de contratos con externos la posibilidad de establecer normas estrictas que aseguren la adquisición de sistemas con un nivel de transparencia, privacidad y responsabilidad adecuada para la función pública (OECD, 2019b).

La aplicación de estas herramientas comprende el diseño e implementación de políticas públicas, que pueden afectar de manera significativa la calidad de vida de la población. Por consiguiente, se debe priorizar la cabal comprensión de la operación de los diferentes algoritmos que asistirán o realizarán una acción en particular, evitando el uso de sistemas de tipo "*caja negra*", y fomentando la capacidad de interpretar, explicar y transmitir efectivamente las razones de una recomendación a las personas usuarias (NSTC, 2019). Esto con el objetivo de rendir cuentas de manera efectiva ante la comunidad, con respecto a las decisiones adoptadas, disponiendo de repositorios públicos que faciliten la reproducibilidad, verificación externa y evaluación de las herramientas desarrolladas (OECD, 2019b).

## **Procesos**:

Alcanzar una interacción efectiva entre humanos y máquinas requiere de experiencias que faciliten la confianza en la utilización de estos nuevos sistemas, fomentando su uso intuitivo y una supervisión activa (NSTC, 2019). Se trata de un cambio de paradigma que va a ampliar el concepto de fuerza laboral, redefiniendo lo que entendemos por trabajo en equipo, pues combina el esfuerzo de personas en conjunto con máquinas autónomas, aprovechando al máximo sus fortalezas y minimizando errores en la ejecución de tareas (OECD, 2019b). Más aún, el mismo proceso de desarrollo de herramientas de inteligencia artificial puede fomentar la creación y consolidación de grupos transversales de trabajo y equipos multidisciplinarios (Eggers et al, 2019).

Para enfrentar este nuevo paradigma, es necesario actualizar y estandarizar procedimientos de forma dinámica, por medio de un enfoque de mejora continua, considerando los avances tecnológicos disponibles para automatizar y optimizar procesos. Esto implica internalizar que, en los proyectos basados en el uso de tecnologías y la ciencia de datos, las personas tienen un rol central para identificar oportunidades de mejora de procesos institucionales. También es necesario desarrollar procedimientos y políticas específicas para la gestión de datos y de documentos, incluyendo procesos de captura, almacenamiento, procesamiento, intercambio y publicación, considerando las implicancias legales y éticas, especialmente con respecto al uso de inteligencia artificial (González et al, 2020 y Denis et al, 2021).

Es fundamental instalar procedimientos y procesos para impulsar la innovación y la transformación de los servicios públicos con foco en las personas, incorporando la lógica experimental para generar experiencias que sean escalables, de acuerdo con lo promovido por el Índice de Innovación Pública, desarrollado por el Laboratorio de Gobierno y el Banco Interamericano[3]. Cabe destacar que la primera entrega de este Índice, durante el año 2020, reconoció a la SMA como uno de los organismos públicos nuevos (creados después del año 1980) con mayores capacidades para innovar, debido a sus prácticas y procesos.

[3] https://lab.gob.cl/indice-de-innovacion

**Estructura orgánica de la SMA:**

Para impulsar el desarrollo de esta estrategia y la integración de sus 3 pilares (Personas, Procesos y Tecnologías), en diciembre de 2020, mediante la Resolución Exenta SMA N ° 2.516, se creó la **División de Seguimiento e Información Ambiental**, la cual tiene la rectoría técnica en ciencia de datos y está orientada al desarrollo de herramientas que permitan optimizar los procesos de la institución. En conjunto con la **División de Fiscalización y Conformidad Ambiental (rectoría en ciencias ambientales)**, la **Fiscalía (rectoría en ciencias jurídicas)**, el **Departamento de Participación y Relacionamiento Comunitario (rectoría técnica en ciencia ciudadana)**, y el **Departamento de Administración y Finanzas (rectoría técnica en ciencias de la administración)**, la estructura de la SMA refuerza el enfoque multidisciplinario y científico requerido para llevar a cabo sus funciones.

En particular, la nueva División de Seguimiento e Información Ambiental está conformada por una **Sección de Gestión de la Información**, que incluye una Oficina de Infraestructura y Redes (hardware), una de Desarrollo y Mantención de Sistemas (software) y una Oficina especializada en ciencia de datos (denominada "Inteligencia Ambiental"), además de una **Sección de Mejoramiento Continuo**, orientada a la mejora de procesos bajo un enfoque Lean (Kenworthy, 2017).

## 3.5 Alianzas

Un aspecto fundamental para consolidar la estrategia es la generación de alianzas con Organismos Gubernamentales, Universidades y Centros de Investigación, considerando el interés mutuo en el desarrollo de proyectos que integren temas ligados al medio ambiente y cambio climático y de aplicaciones de ciencia de datos e inteligencia artificial. De este modo, se busca fomentar un ecosistema de innovación abierto que permita el intercambio de experiencias y conocimientos. La innovación colaborativa entre distintas organizaciones mediante alianzas estratégicas, contribuyen al desarrollo de un ecosistema de ciencia, tecnología, conocimiento e innovación (OECD, 2019b).

- **Comité asesor de expertos**

Se conformará un comité de expertos en ciencia de datos e inteligencia artificial que asesorará el desarrollo y actualización de esta estrategia, invitando a representantes de distintos sectores (academia, industria, Estado y Sociedad Civil) que puedan aportar visiones distintas y complementarias.

• **Convenios con servicios públicos**

Se fomentará la firma de convenios y alianzas con organismos públicos para la interoperabilidad de servicios y el intercambio de información, a través de medios digitales (Application Programming Interface: API). En este contexto, se considerarán actores claves en materia ambiental e información atingente al desempeño de la SMA tales como el Ministerio del Medio Ambiente, el Servicio de Evaluación Ambiental, el Servicio de Impuestos Internos, la Tesorería General de la República, la Contraloría General de la República, el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, y los organismos públicos miembros de la Red Nacional de Fiscalización Ambiental (RENFA)[4].

• **Convenios con Universidades y Centros de Investigación**

Se fomentará la materialización de convenios, la realización de investigaciones aplicadas, y la postulación a fondos y programas concursables, tanto nacionales, como internacionales. De esta forma, interesa acercar a docentes, investigadores, profesionales y futuros profesionales para el desarrollo de proyectos conjuntos con la SMA que permitan avanzar en el desarrollo científico y, al mismo tiempo, en la generación de soluciones ambientales.

• **Colaboración con organizaciones internacionales**

Se promoverá instancias para intercambiar conocimientos y experiencias con otros organismos y redes internacionales, tales como INECE[5] (Red Internacional de Fiscalización y Cumplimiento Ambiental), REDLAFICA[6] (Red Latinoamericana de Fiscalización Ambiental) y agencias de fiscalización y cumplimiento ambiental de otros países. También se incentivará la difusión de experiencias a través de publicaciones en revistas científicas. De este modo, la SMA compartirá los avances logrados y aprenderá también de las experiencias de otros organismos.

[4] https://renfa.sma.gob.cl/

[5] https://inece.org/

[6] http://www.redlafica.org/

**A continuación, se presentan algunas alianzas y avances logrados:**

• **Fondo Empatía**: Consiste en una cooperación de diversos organismos para promover la aplicación de la Inteligencia Artificial enfocada a resolver problemas públicos en Latinoamérica. Durante el año 2020, el fondo seleccionó un proyecto conjunto entre la SMA y la Universidad Adolfo Ibáñez, para implementar un modelo que permita predecir la calidad del aire y la ocurrencia de episodios críticos, en base a los datos de emisiones de industrias, estaciones de calidad del aire y meteorología[7], como herramienta para priorizar acciones de fiscalización oportunas.

• **Data Science For Social Good - DSSG**: Se trata de un programa intensivo de 3 meses que coordina la Universidad de Warwick en Inglaterra, en conjunto con la Universidad de Múnich de Alemania, el Instituto Alan Turing del Reino Unido y la Fundación DSSG. La iniciativa ayuda a organizaciones sin fines de lucro y organismos de Gobierno para mejorar sus servicios, a través del desarrollo de herramientas de ciencia de datos por parte de estudiantes universitarios[8]. Durante el año 2021 la SMA fue seleccionada para desarrollar un modelo que permita clasificar y priorizar denuncias al momento de su ingreso, en función de la información histórica de casos terminados, con objeto de optimizar los tiempos de gestión.

• **Colaboración con Universidades**: Se firmó un convenio con la Universidad Adolfo Ibáñez, que permitió la realización de una tesis colectiva (*capstone*) del magíster de ciencia de datos para el desarrollo de un modelo que permita clasificar y priorizar denuncias al momento de su ingreso, en función de la información histórica de casos terminados, de forma paralela a la iniciativa con DSSG. Durante el año 2021 también se están impulsando trabajos colaborativos con la Escuela de Ingeniería de la Pontificia Universidad Católica de Chile y el Centro de Investigación en Inteligencia de Negocios, del Departamento de Ingeniería Industrial de la Universidad de Chile, entre otros.

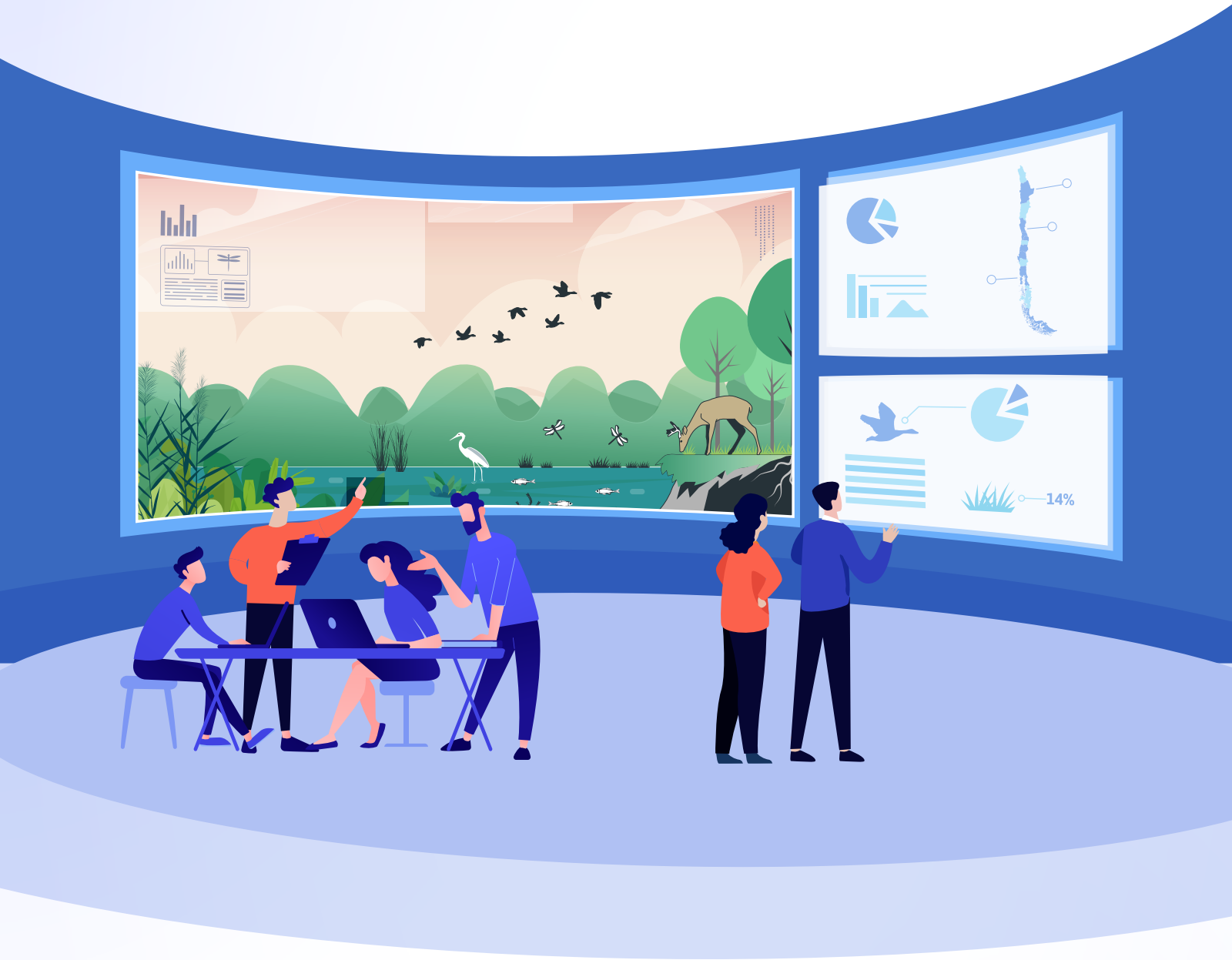
• **Fondos concursables:** La SMA se adjudicó tres proyectos de Investigación científica y tecnológica, mediante el concurso IdeA I+D 2021 del Fondo de Fomento Ciencia y Tecnología (FONDEF), que consideran el uso de Big Data y de *machine learning* para fortalecer el seguimiento ambiental. El primer proyecto se denomina "Desarrollo de una Plataforma de Inteligencia de Datos (PID) para un nuevo Sistema Nacional de Seguimiento Ambiental", se desarrollará en conjunto con la Fundación Ciencia y Vida, y su objetivo es procesar los datos de análisis y mediciones reportadas por empresas que cuentan con Resoluciones de Calificación Ambiental, con énfasis en agua y biodiversidad. El segundo proyecto se denomina "Sistema de Alerta y Monitoreo Satelital de Áreas de Relevancia Ambiental (SAMSARA): casos de estudio en humedales urbanos, turberas y bosque esclerófilo", se desarrollará en conjunto con la Universidad Adolfo Ibáñez y su objetivo es desarrollar un sistema de monitoreo automatizado de ecosistemas de alto interés, mediante la utilización de imágenes satelitales. El tercer proyecto se titula "Desarrollo de herramienta computacional para la certeza jurídica en materia ambiental", presentado con la Universidad de los Andes, y busca apoyar la elaboración de textos con contenido jurídico relevante.

---

[7] https://www.empatia.la/proyecto/inteligencia-artificial-para-el-cuidado-del-medio-ambiente-y-la-salud-de-la-poblacion/

[8] https://warwick.ac.uk/research/data-science/warwick-data/dssgx/5 https://inece.org/

# 4. Plataforma de Inteligencia de Datos (PID)




**“El propósito de la computación es la comprensión, no los números”.**

Richard Hamming (1962).

## 4.1 Modelo conceptual

La Plataforma de Inteligencia de Datos (PID) es el componente fundamental para el desarrollo de esta estrategia y corresponde a la consolidación de la plataforma electrónica establecida en el artículo 33 de la Ley Orgánica de la SMA. Mediante la PID se espera emular la capacidad humana de sentir, comprender, actuar y aprender (Figura 13) (Bataller & Harris, 2016). Si bien en un inicio, y como prueba de concepto, el foco ha estado puesto en desarrollar proyectos que simulen inteligencia de manera acotada y automaticen tareas rutinarias mediante algoritmos de cálculos o sistemas expertos con reglas de decisión específicas, se está avanzado en la incorporación de algoritmos más sofisticados que perciban el contexto en que se encuentran, aprender al observar nueva evidencia y adaptarse a nuevas circunstancias (Dellermann et al., 2019b).

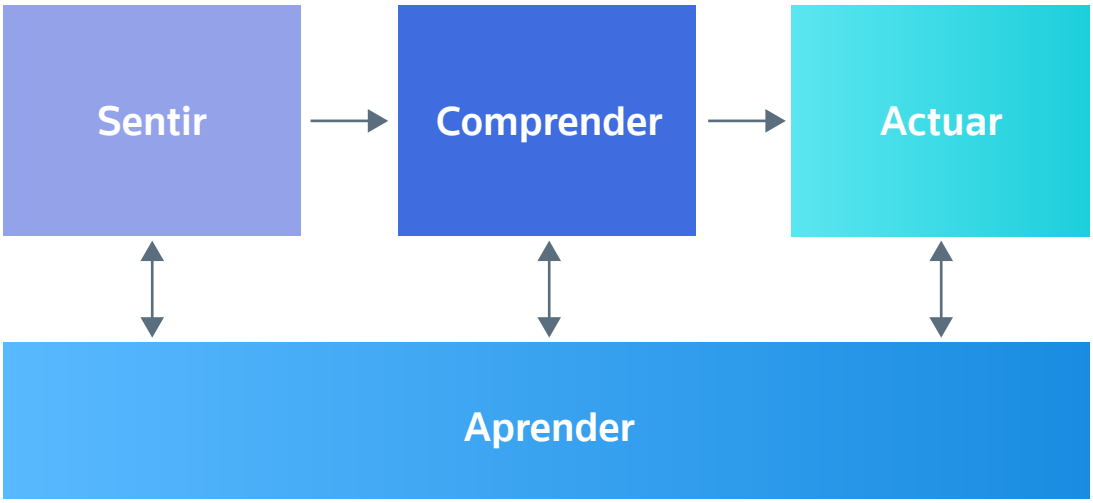



***Figura 13:*** *Principales habilidades humanas a considerar en el desarrollo de herramientas de inteligencia artificial (basado en Bataller & Harris, 2016)*

En términos conceptuales, la PID está conformada por diversos módulos, que incluyen sistemas de información, bases de datos y herramientas analíticas para generar la información y el conocimiento necesario que permita activar respuestas oportunas, eficaces y eficientes por parte de la SMA y de los distintos organismos y actores involucrados en la temática ambiental. En particular, la PID cuenta con 3 capas principales: (i) captura de datos, incluyendo procesos de consistencia y validación; (iii) analítica avanzada para generar información; y (iii) acciones para facilitar gestiones y decisiones. Más aún, se espera que el equipo interno de la SMA, organizaciones y personas aporten datos y se retroalimenten de la misma PID, generándose **un círculo virtuoso de inteligencia colectiva híbrida para la protección del medio ambiente y la salud de la población** (Figura 14).

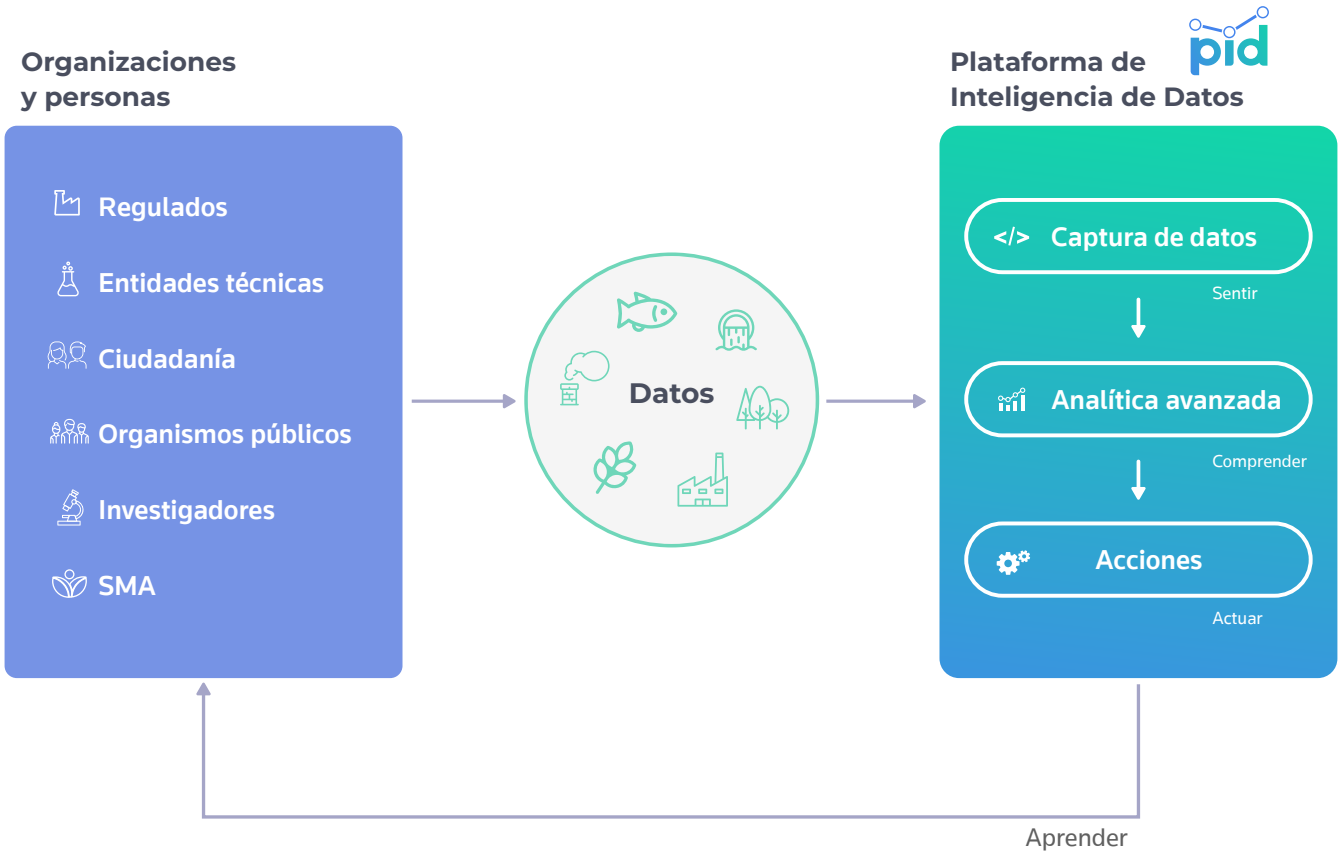



***Figura 14:*** *Plataforma de Inteligencia de Datos (PID) y su integración con organizaciones y personas.*

## 4.2 Captura de datos

Los humanos percibimos nuestro entorno mediante los sentidos. En el caso de una máquina inteligente, este proceso se realiza mediante sensores que capturan y procesan datos de distinta índole como imágenes, sonidos, o texto, logrando la capacidad de sintetizar datos visuales, interpretar el lenguaje humano o analizar archivos de audio (OECD, 2019b). El desarrollo de iniciativas de inteligencia artificial requiere de una cantidad suficiente de datos de calidad, actualizados de manera continua que permitan entrenar algoritmos de última generación (NSTC, 2019; OECD, 2019b).

La PID facilitará el acceso a grandes volúmenes de datos, correctamente etiquetados y de alta calidad, depurados para facilitar el desarrollo de analítica avanzada, recopilados de manera estratégica para potenciar el desempeño de modelos estadísticos, y en formatos estructurados para facilitar el entrenamiento permanente y continuo de modelos de aprendizaje automático (Eggers et al, 2019). Esta información estará públicamente disponible, facilitando no solo el desarrollo de soluciones en materia ambiental, sino también de diversas áreas científicas, incluyendo la ciencia de datos, teniendo el carácter de bien público (NSTC, 2019).

## Plataforma electrónica

La SMA es una de las instituciones referentes de datos ambientales a nivel nacional. En efecto, miles de sujetos regulados por la normativa ambiental deben reportar periódicamente el cumplimiento de compromisos, análisis y mediciones de variables ambientales, a lo cual se suma la información que se levanta internamente en los procesos de denuncia, fiscalización, seguimiento, cumplimiento y sanción ambiental. Para esto, durante años se ha desarrollado y consolidado una plataforma electrónica conformada por distintos sistemas de reporte electrónico que son utilizados por la Comunidad Regulada[9], Entidades Técnicas, Ciudadanía, Organismos Públicos, más sistemas de gestión interna, los cuales se encuentran integrados por medio de un Catastro de Unidades Fiscalizables (CUF), que corresponde a un repositorio que alimenta a los sistemas con información única y transversal sobre empresas reguladas (Figura 15).

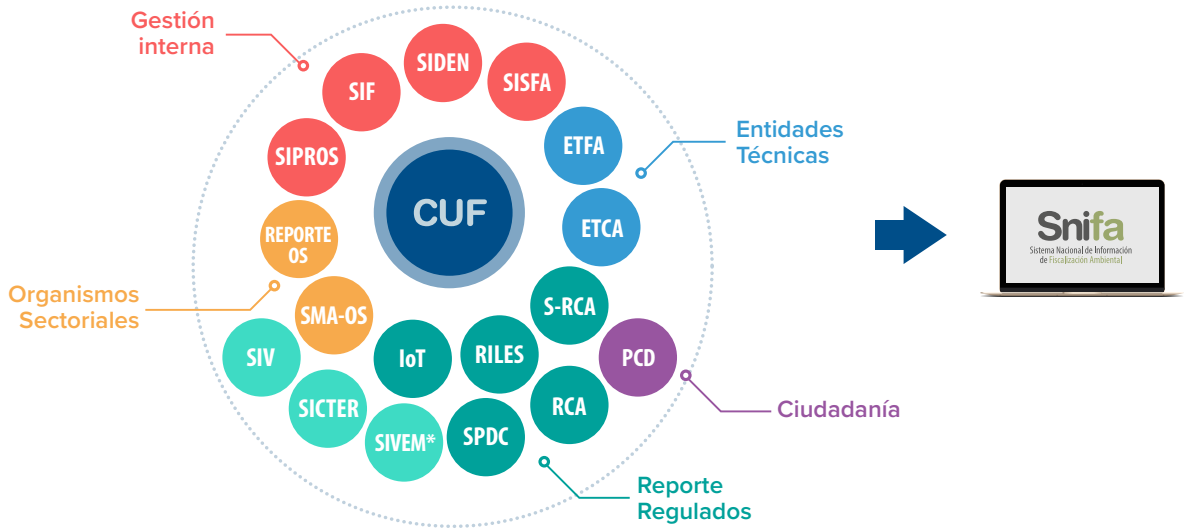



***Figura 15:*** *Plataforma electrónica de la SMA, conformada por distintos sistemas que capturan datos. La información generada se encuentra disponible públicamente a través de SNIFA.*

**Digitalización de trámites:**

Una de las metas planteadas para aumentar la captura de datos y fortalecer el nivel de servicio de la SMA es lograr el 100% de trámites digitalizados al cierre del año 2021, de acuerdo con el enfoque de "Cero filas", entendiendo un trámite como toda acción necesaria para que una persona natural y/o jurídica acceda a un producto entregado por una institución pública, generando un procedimiento que finaliza con una respuesta para el solicitante. Al inicio del año 2021, más del 90% de los trámites de la SMA eran digitales, de un total cercano a 50, incluyendo el ingreso de denuncias, la autorización de Entidades Técnicas, y, principalmente, el reporte de compromisos y obligaciones establecidas en instrumentos de carácter ambiental. Esto ha permitido que la SMA sea uno de los organismos públicos que más ha avanzado en la digitalización de trámites durante los últimos años, según el Registro Nacional de Trámites (RNT)[10]. Sin perjuicio de lo anterior, es necesario mejorar continuamente la captura de datos, para lo cual se deben considerar proyectos de actualización de los sistemas, en función de la experiencia de los usuarios y los requerimientos de la institución.

[9] https://portal.sma.gob.cl/index.php/portal-regulados/sistemas-de-reporte/

[10] https://tramites.gob.cl/

Pese a estos avances, las funciones de la SMA requieren de esfuerzos adicionales de captura de datos para complementar la información disponible, abarcando nuevos campos y formatos. La sistematización de datos desde el origen facilita enormemente las etapas posteriores de validación, por lo cual, para el desarrollo de la PID, se debe privilegiar el reporte de datos en formatos estructurados, mediante formularios web, planillas y APIs, de forma adicional a la entrega de documentos consolidados que suelen incluir análisis y conclusiones a partir de los datos levantados. Cabe destacar que la SMA está facultada para establecer forma y modo de reporte, privilegiando los medios electrónicos y, adicionalmente, puede solicitar información para la conformación y mantención del Sistema Nacional de Información de Fiscalización Ambiental, según lo indicado en el artículo 32 de su Ley Orgánica.

En este sentido, es necesario considerar que existen diferentes tipologías de datos por capturar. En primer lugar, aquellos que se pueden clasificar con relación a su estructura, lo cual tiene un impacto no solo en la captura, sino que también en su almacenamiento y procesamiento, identificándose 3 grandes grupos: Estructurados, Semiestructurados, y No estructurados. Por otro lado, los datos que la SMA necesita procesar pueden clasificarse como indicadores en términos ambientales, siguiendo el modelo FPEIR: Fuerza motriz - Presión - Estado - Impacto - Respuesta (Smeets & Weterings, 1999), ampliamente utilizado a nivel nacional e internacional (MMA, 2020), a través del cual se busca mostrar las relaciones causales entre el medio ambiente y la actividad humana (Figura 16).

***Figura 16:*** *Tipos de datos por capturar (modelo FPEIR adaptado de Smeets & Weterings, 1999).*

**Tipos de datos a capturar: según estructura y naturaleza**

*Clasificación de datos según estructura:*

- **Estructurados**: que cuentan con formatos fijos según parámetros específicos para organización en esquemas relacionales (Excel, CSV).
- **Semiestructurados**: que no tienen formatos fijos, pero contienen etiquetas y otros marcadores (HTML, XML, JSON).
- **No estructurados**: que no tienen una organización o jerarquía interna clara (Word, PDF, imagen, audio, video).

*Clasificación de datos, siguiendo el modelo "FPEIR":*

- **Fuerzas motrices:** Se refieren a factores o variables indirectas que están detrás de las presiones más específicas que afectan al medio ambiente. Ejemplos: demografía, actividades económicas, marco institucional, etc.
- **Presiones:** Se refieren a factores o variables directas que afectan el estado de los componentes del medio ambiente de manera individual o colectiva. Estas presiones pueden ser de orden antrópico o procesos naturales. Ejemplos: emisiones atmosféricas, descargas de residuos líquidos, ruidos, entre otros.
- **Estado del medio ambiente:** Se refiere a la situación en que se encuentran los componentes del medio ambiente, producto de fuerzas motrices y de presiones. Ejemplos: indicadores de biodiversidad (flora, fauna y fungi), calidad del aire, calidad del agua, meteorología, entre otros.
- **Impactos**: Se refiere a las consecuencias de cambios en el estado del medio ambiente, tanto en la calidad de vida o salud de la población, como en servicios ecosistémicos. Ejemplos: pérdida de biodiversidad, daño ambiental, contaminación ambiental, entre otros.
- **Respuesta:** Se refiere a las acciones que realizan tanto las autoridades, como la sociedad en general, ya sea en orden a disminuir los impactos ambientales o para adaptarse a éstos. Estas acciones afectan el estado de los componentes del medio ambiente, así como las presiones y las fuerzas motrices. Justamente, este ámbito incluye la gestión realizada para proteger el medio ambiente, a través de políticas y regulaciones, fiscalizaciones, procedimientos sancionatorios, medidas cautelares, cumplimiento de programas, entre otros.

Particularmente, los esfuerzos de captura en esta estrategia se concentran en tres líneas de trabajo: (i) generación de planillas estandarizadas y formularios web para datos discretos; (ii) reforzar el envío de datos en tiempo real para variables continuas vía APIs; y (iii) ampliar el uso de imágenes satelitales de acceso libre para extraer datos de variables ambientales. Para el primero de ellos, se destacan los avances logrados con respecto al reporte estandarizado de datos de la componente agua, gracias a lo cual se han capturado millones de registros en diversos puntos de monitoreo a nivel nacional, lo que se espera replicar para otros componentes, como por ejemplo datos de biodiversidad o el seguimiento de rubros específicos como refinerías o rellenos sanitarios, entre otros.

## Reporte continuo de datos

En tanto, el reporte de datos en tiempo real ha experimentado un crecimiento sostenido durante el año 2021, mediante la conexión de sensores remotos y continuos, lo que se conoce como Internet de las Cosas (IoT en sus siglas en inglés), tecnologías que funcionan como complemento al reporte periódico de datos discretos, típicamente incluidos en la normativa ambiental (por ejemplo: mediciones realizadas por un laboratorio). Para esto, se ha desarrollado una plataforma tecnológica que permitió la conexión masiva de cientos de sensores, que llegarán a miles en un corto plazo, por medio de un protocolo de integración máquina a máquina, basado en tecnología API REST (Figura 17), que permite el envío automatizado de datos desde sistemas informáticos de empresas reguladas. De este modo, las instrucciones establecidas por la SMA para el reporte de datos en formatos estandarizado, ya sea de manera periódica o en tiempo real[11], serán claves para ampliar la cobertura del monitoreo ambiental.

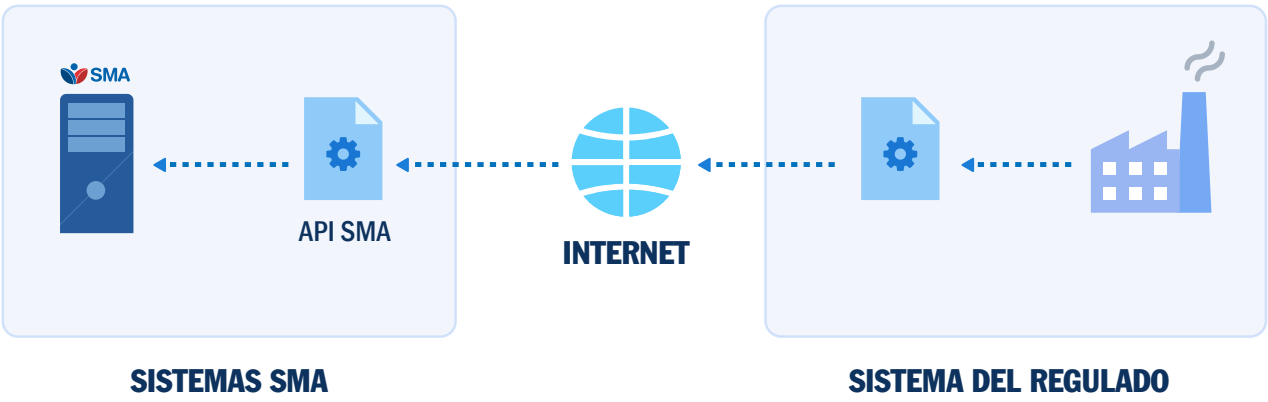



***Figura 17:*** *Reporte de datos en tiempo real vía API.*

Otra forma relevante de capturar datos de forma remota es mediante la utilización de imágenes satelitales. Para esto, se ha promovido el uso de programas de observación de la Tierra de acceso libre, principalmente de Estados Unidos (Landsat[12]) y de la Comunidad Europea (Copernicus[13]), que permiten acceder a series temporales de imágenes ópticas y de radar para monitorear el estado de la vegetación, cuerpos de agua y la presencia de estructuras en el territorio, entre otros (Figura 18). El uso sostenido y permanente de este tipo de herramientas permite incrementar la cobertura de la vigilancia ambiental en el país, no solo en unidades fiscalizables y sus procesos, sino que también en diversas componentes ambientales a lo largo y ancho del territorio. Todo esto potenciado por las nuevas capacidades que incorpora el Sistema Nacional Satelital (SNSat), administrado por el Servicio Aerofotogramétrico (SAF) de la Fuerza Aérea de Chile (FACH).

Chile es considerado un laboratorio natural para observar la evolución y los impactos del cambio climático. La disponibilidad de información confiable y de alta frecuencia, puede contribuir sustantivamente a la generación de conocimiento para el diseño de soluciones a esta problemática global (MinCiencia, 2021). En este sentido, se considera el desarrollo de un esquema de transmisión de los datos de la PID vía API, con objeto de impulsar iniciativas público-privadas como el Observatorio de Cambio Climático o el Data Observatory, entre otros.

[11] https://portal.sma.gob.cl/index.php/portal-regulados/instructivos-y-guias/conexion-en-linea-a-la-sma/

[12] https://www.usgs.gov/core-science-systems/nli/landsat

[13] https://www.copernicus.eu/en

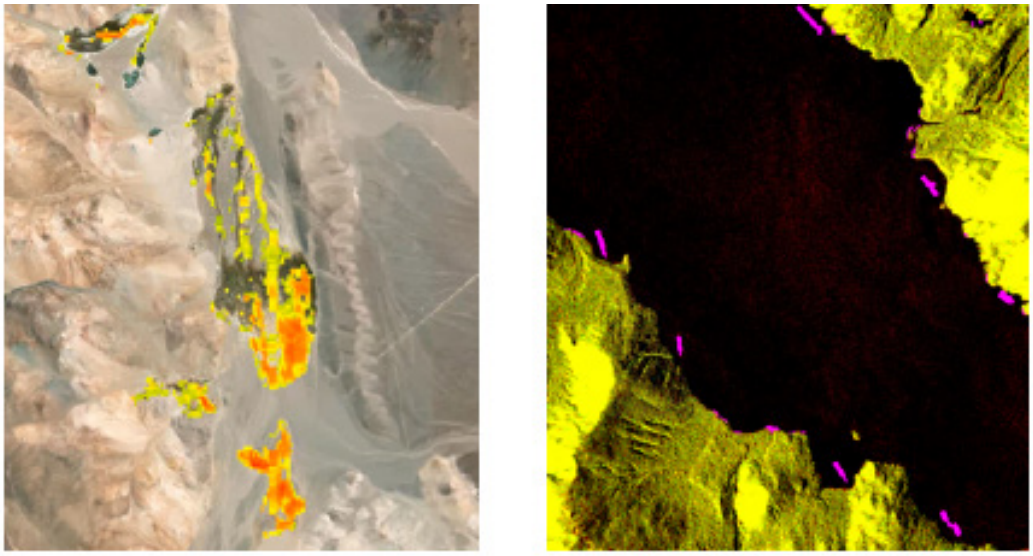


***Figura 18:*** *Imágenes satelitales para monitorear un humedal altoandino (izquierda) y Centros de Engorda de Salmones (derecha), de tipos óptica y radar de apertura sintética, respectivamente.*

## Laboratorio de Alta Complejidad:

La SMA está implementando un Laboratorio de Alta Complejidad (LAC), que será inaugurado a fines del 2021 y que permitirá que el país cuente con la capacidad analítica necesaria para identificar contaminantes ambientales de forma rápida, con altos estándares de precisión y exactitud, mediante la asesoría experta del Norwegian Institute for Air Research (NILU) de Noruega, centro de referencia mundial en temas relacionados con el monitoreo de contaminantes ambientales. En una primera etapa, los esfuerzos se han concentrado en la generación de las capacidades técnicas para la determinación de Compuestos Orgánicos Volátiles (COVs) en el aire. De esta forma, el Laboratorio de Alta Complejidad permitirá generar datos de calidad del aire y de otras matrices ambientales, los que serán utilizados por la PID para generar información y conocimiento (Figura 19).

***Figura 19:*** *Instalaciones del Laboratorio de Alta Complejidad de la SMA.*

## Consistencia y validación de datos

La implementación de la PID debe abordar una gran variedad de datos que son reportados en distintos formatos y mediante diversos canales, que serán integrados para generar un repositorio consolidado que permita la realización de análisis. Considerando que periódicamente se captura una gran cantidad de datos mediante planillas estandarizadas en formato CSV y/o XLSX (lo que sucede principalmente con respecto a las componentes ambientales de agua, aire, y biodiversidad) es necesario considerar procesos de preparación de los datos. En este se sentido, dependiendo de las etapas en el desarrollo de la PID se considera el uso de 2 paradigmas, que utilizan los mismos principios, pero en un orden distinto: (i) Extracción, Carga y Transformación (ELT, por sus siglas en inglés), que es un proceso en el cual los datos sin procesar se cargan directamente en un repositorio final y se transforman allí; y (ii) Extracción, Transformación y Carga (ETL, por sus siglas en inglés), en el cual la transformación tiene lugar en un servidor intermedio antes de la carga en el destino final (Figura 20), según se explica a continuación:

- **Extracción** de los datos desde sistemas de origen (por ejemplo, desde planillas reportadas*).*
- **Carga** de los datos en un repositorio de datos (por ejemplo, en un *Data Lake*).
- **Transformación** de los datos desde su formato de origen al formato requerido para analizarlos, para lo cual se pueden aplicar diversas reglas de validación, ya sea poniéndolos a disposición de las etapas siguientes de análisis o activando procedimientos para rechazar registros erróneos que no cumplen los estándares establecidos y solicitar la carga de los datos corregidos. Algunos ejemplos corresponden a la aplicación de funciones matemáticas, el reemplazo de códigos por valores y la combinación de datos de distintas fuentes.

Esta es una etapa fundamental para el control de calidad que debe ejecutarse recurrentemente y que proporciona retroalimentación para los procesos de análisis de los datos, incluyendo la incorporación de nuevas fuentes de datos como, por ejemplo, textos en formato PDF, imágenes y videos, entre otras.

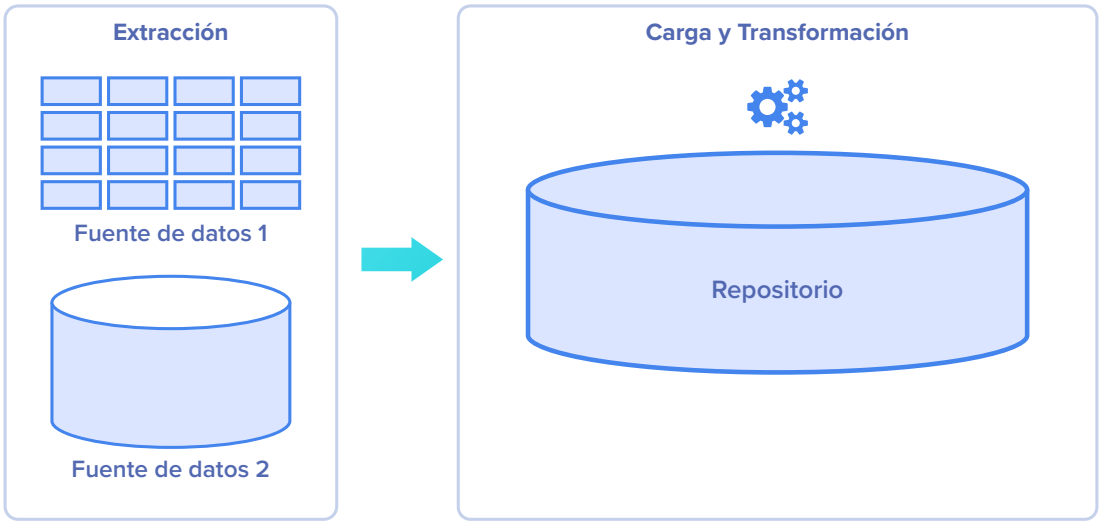



***Figura 20:*** *Proceso de Extracción, Carga y Transformación de los datos (ELT).*

Si bien se debe velar por la buena calidad de los datos consolidados, no es el único actor relevante. De hecho, las fuentes de información de origen son los primeros responsables de la consistencia de los datos entregados en la forma y el modo requerido, por lo que un rol importante de la PID es generar una constante retroalimentación a las fuentes de origen, con respecto a la calidad de los reportes recibidos, para así orientar la implementación de mejoras. La corrección de datos de manera conjunta no sólo permite el objetivo primario de contar con información confiable, sino que también promueve la solución de errores en el inicio de la cadena de datos, facilitando la depuración y consolidación de las bases de datos de forma previa a su análisis, lo que contribuye a evitar la presencia de valores erróneos e identificar datos anómalos (OECD, 2019b).

**¿A qué llamamos datos de calidad?**

Según la Organización para la Cooperación y el Desarrollo Económico (OECD, 2011), los datos de calidad deben cumplir con las siguientes características:

- **Relevantes:** que sean útiles para responder la interrogante planteada.
- **Exactos:** que representen correctamente lo que miden.
- **Creíbles:** que tengan un origen seguro y confiable.
- **Oportunos:** que reflejen el período temporal de interés.
- **Accesibles:** que estén disponibles para su uso.
- **Interpretables:** que sean entendibles y utilizables.
- **Coherentes:** que sean comparables y consistentes.

## 4.3 Analítica avanzada

El desarrollo de la habilidad de comprender es clave para dar un propósito útil a los datos capturados, con objeto de generar información y, a partir de ello, conocimiento sobre una materia especifica (Ackoff, 1989). Esto incluye entender patrones en los datos, considerando el contexto presente al momento en que estos se generan. De esta manera, la presente estrategia considera la implementación de diversos módulos analíticos integrados para optimizar la toma de decisiones, automatizar procesos y generar información y conocimiento sobre el cumplimiento de la normativa, la efectividad de la regulación y el estado del medio ambiente. En este sentido, se incluyen los diferentes niveles de capacidades analíticas, según lo planteado en el modelo ascendente de Gartner (Hostmann, 2012).

- **Analítica descriptiva**: Se enfoca en la obtención de métricas, indicadores y patrones de datos existentes, para entender el pasado. Es el enfoque de Inteligencia de Negocios (BI) tradicional. En el ámbito de la SMA, una aplicación al respecto sería verificar el nivel de cumplimiento de una unidad fiscalizable, a partir de los datos reportados, según los compromisos y requerimientos establecidos en la normativa.

- **Analítica de diagnóstico**: Su objetivo es entregar información sobre las causas de por qué sucedió algo en el pasado. Para ello, se utilizan técnicas tales como correlaciones, *data mining* y *drill-down*. A modo de ejemplo, una aplicación para la SMA sería determinar el efecto de los pozos de bombeo de una unidad fiscalizable sobre la vigorosidad de la vegetación en un humedal altoandino.

- **Analítica predictiva**: Se utiliza para la predicción de posibles resultados o escenarios, a través de modelos estadísticos y técnicas de *machine learning*. En este caso, una aplicación para la SMA sería el desarrollo de un modelo predictivo de calidad del aire, a partir de los datos de emisiones de unidades fiscalizables y las condiciones meteorológicas en una zona específica.

- **Analítica prescriptiva**: Se utiliza para recomendar o sugerir acciones, de modo de facilitar y automatizar la toma de decisiones. Para esto, es posible integrar distintas técnicas de las analíticas descriptivas y predictivas. En este caso, un ejemplo para la SMA sería un modelo que oriente la toma de decisiones en contingencias ambientales, ¿qué fuentes de contaminación deberían fiscalizarse, según las condiciones meteorológicas en un momento determinado?

De esta forma, la estrategia se enfoca en la generación de capacidades predictivas y prescriptivas, que permitan identificar desviaciones o irregularidades de forma temprana para activar alertas y medidas oportunas, mediante la ciencia de datos y la inteligencia artificial, potenciando las capacidades basales para análisis descriptivos y de diagnóstico, instaladas durante el período 2015-2019. De este modo, la SMA necesita desarrollar y fortalecer sus capacidades para todo tipo de analíticas, principalmente de aquellas relacionadas con la automatización de tareas y la predicción de escenarios para apoyar la toma de decisiones (Figura 21).

En el primer caso, por ejemplo, interesa automatizar el proceso de verificación del cumplimiento de los distintos instrumentos de carácter ambiental, con objeto de alcanzar una cobertura total de las Unidades Fiscalizables en el territorio, incluyendo normas de emisión, Resoluciones de Calificación Ambiental y Planes de Descontaminación Ambiental, entre otros. En el caso de los modelos predictivos, interesa identificar desviaciones o irregularidades de forma temprana, mediante distintos escenarios y proyecciones en el estado del medio ambiente, que permitan anticiparse a la materialización de impactos significativos. En este sentido, y siguiendo el modelo *FPEIR*, el desarrollo de modelos analíticos es fundamental para entender los impactos generados por las presiones antrópicas sobre el estado del medio ambiente, así como también para revisar si las respuestas generadas por la institucionalidad ambiental son efectivas en el tiempo y evaluar alternativas cuando corresponda.

Cabe destacar que el concepto de "*predicción*" en la ciencia de datos se refiere a la estimación de un valor desconocido, ya sea en el futuro, presente o pasado, mediante una formulación matemática o lógica, lo que contrasta con el enfoque "descriptivo", el cual busca entender un fenómeno y generar conocimiento a partir de los datos (Provost & Fawcett, 2013). En este sentido, el uso de métodos de aprendizaje automático (*machine learning*) ha tenido un importante crecimiento durante los últimos años. En líneas generales, esta técnica puede aportar significativamente a mejorar el trabajo de la SMA y de manera más amplia, a enfrentar problemas ambientales en general (Rolnick et al., 2019).

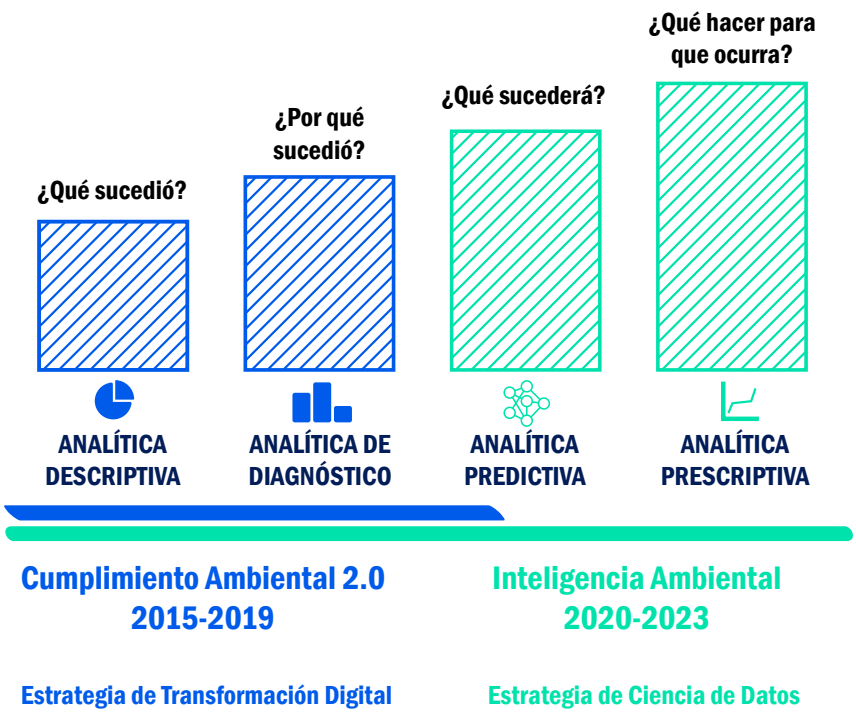



***Figura 21:*** *Desarrollo de capacidades analíticas en la SMA basado en Hosteman (2012).*

## Machine learning

Machine learning o aprendizaje automático comprende una gran familia de algoritmos que "aprenden" de nuevas observaciones a través de la experiencia, para lo cual existen distintos paradigmas de aprendizaje (supervisado, no-supervisado y reforzado) y dos formas generales de incluirlos en el proceso de toma de decisiones: (i) Sistemas de soporte, donde resultados generados por los modelos se utilizan como insumos para la toma de decisiones; y (ii) Sistemas de toma de decisión, donde los resultados obtenidos por los modelos se aplican directamente sin intervención humana directa (Jordan & Mitchell, 2015; González et al, 2020).

Existen distintos métodos para el desarrollo de modelos de aprendizaje automático, con diferentes arquitecturas y estructuras (Figura 22). Este tipo de modelos son capaces de descubrir la representación del problema más eficiente y, al mismo tiempo, realizar tareas como la clasificación de elementos, por medio de enfoques probabilísticos (modelos de markov, redes bayesianas, etc.), estadísticos (regresión lineal, regresión logística, etc.) y computacionales (redes neuronales, árboles de decisión, etc.), cuya selección depende del objetivo buscado, las capacidades existentes y los datos disponibles (Murphy, 2012; Mitchell, 1997; Bishop, 2006, Lepenioti et al, 2020). Cabe destacar que, si bien no aplica a la totalidad de los casos, los modelos de aprendizaje automático se fundan principalmente en la teoría estadística clásica y moderna (Wasserman, 2004; Hastie et al, 2009; Hastie et al, 2019; Shalizi, 2016).

***Figura 22:*** *Métodos para el desarrollo de modelos predictivos (adaptado de Lepenioti et al, 2020).*

**Paradigmas de aprendizajes para modelos de Machine Learning:**

Existen diferentes tipos de aprendizaje de los algoritmos (OECD, 2019b):

- **Aprendizaje supervisado:** los datos se encuentran etiquetados y los algoritmos aprenden a predecir los datos de salida a partir de los datos de entrada. Los principales algoritmos corresponden a regresión y clasificación.
- **Aprendizaje no supervisado:** los datos no se encuentran etiquetados y los algoritmos aprenden a agrupar los datos de entrada. Los principales algoritmos típicamente utilizados son clustering y reducción de dimensionalidad.
- **Aprendizaje reforzado:** es un área inspirada en la psicología conductista; donde la máquina aprende el comportamiento a seguir en base a recompensas y penalizaciones directamente con el entorno, mediante el ensayo y error. Se utiliza en gran medida para aplicaciones de robótica.

En la diversidad de algoritmos de aprendizaje automático, muchos de ellos son inspirados en mecanismos de computación propios del cerebro humano (Hinton, 1992), en técnicas de optimización sofisticadas (Sra et al, 2012) o en técnicas de regresión (Tibshirani, 1996). Recientemente, se ha popularizado el uso de algoritmos de aprendizaje profundo con estructuras matemáticas de alta complejidad, principalmente para el desarrollo de modelos supervisados (Goodfellow et al, 2016; LeCun et al, 2015). Este tipo de esquemas también ha sido utilizado en modelos no-supervisados, permitiendo reducir la complejidad de datos de alta dimensión a versiones que permiten apreciar patrones simples escondidos en la data (Hinton & Salakhutdinov, 2006).

Particularmente, para el desarrollo de modelos de clasificación, predicción y detección de anomalías a partir de series temporales, destaca el desempeño de las redes neuronales recurrentes denominadas LSTM (por sus siglas en inglés, *Long Short Term Memory*), que tienen una memoria que es capaz de recordar a través de largas secuencias (Hochreiter & Schmidhuber, 1997) y que, por lo tanto, son útiles para analizar datos de seguimiento ambiental en el tiempo, que es uno de los desafíos analíticos más importantes de la SMA (por ejemplo: calidad del agua, calidad del aire, biodiversidad, entre otros).

Para contar con sistemas confiables y robustos, éstos deben ser testeados en escenarios reales mediante pruebas de concepto, antes de ser puestos en servicio operativo (OECD, 2019b). Durante su operación, los sistemas desarrollados deben ser verificados adecuadamente, para asegurar que se cumplen las especificaciones formales de diseño y validados continuamente, garantizando que los requerimientos operacionales respondan al propósito establecido, lo que implica evaluar y diagnosticar fallas con precisión cuando se presentan, a objeto de reparar y ajustar los sistemas para operar dentro de límites seguros (NSTC, 2019).

Por lo tanto, considerando la gran variedad de alternativas de modelos y algoritmos existentes, es fundamental generar capacidades internas y fortalecer la gobernanza de datos. Para esto, la SMA requiere contar con un equipo de científicos de datos, que tengan experiencia y conocimientos en programación, estadísticas y temas ambientales. De esta manera, será posible sustentar el desarrollo de sistemas tanto de soporte, como de toma de decisiones, con los más altos estándares científicos, generando un espacio de trabajo creativo que aporte sustantivamente a la investigación científica, ya sea teórica o aplicada, y así proporcionar soluciones innovadoras a desafíos de políticas públicas.

## 4.4 Acciones

Potenciadas las capacidades de sentir y comprender el estado del medio ambiente y la efectividad de la normativa ambiental, es necesario priorizar acciones que permitan dar respuesta a situaciones que requieran mejoras y/o medidas oportunas. De esta manera, se considera la implementación de distintas acciones de la PID que contribuyan a la misión institucional, incluyendo el desarrollo de sistemas de automatización, sistemas de toma de decisiones, sistemas de soporte, sistemas predictivos y sistemas responsivos, orientados a facilitar respuestas adaptativas a un entorno dinámico. Lo anterior implica transmitir los resultados generados a todas las organizaciones y personas interesadas que puedan contribuir con la protección del medio ambiente, de forma oportuna, eficaz y eficiente, ya sea al interior de la SMA o en otras organizaciones. Para esto, se consideran las siguientes acciones.

### Optimizar la gestión interna

En primer lugar, se considera mejorar la gestión interna, a través de la optimización y automatización de procesos de denuncias, fiscalización, seguimiento, cumplimiento y sanción y mediante la generación de paneles informativos y herramientas que faciliten las actividades de la SMA. Esto incluye la verificación automática del cumplimiento normativo de los diversos instrumentos de carácter ambiental por parte de la comunidad regulada, generando informes de fiscalización e insumos para los procedimientos sancionatorios de forma directa. Por otro lado, también se considera el envío de reportes y antecedentes para apoyar la toma de decisiones en las distintas áreas de la SMA, en función de los resultados de los modelos analíticos, incluyendo la generación de indicadores, proyección de escenarios y predicciones de variables ambientales de interés.

**Proyecto destacado:**

### Gestión de denuncias

Uno de los objetivos estratégicos que la SMA estableció para el presente período fue mejorar la entrega de servicios a la comunidad y en este contexto el proyecto para optimizar la gestión de denuncias es prioritario. Históricamente, las denuncias eran presentadas en formato físico, siendo ingresadas por los denunciantes de forma presencial en oficinas de la institución. De esta forma, para facilitar y digitalizar la presentación de denuncias por parte de la ciudadanía, se desarrolló una plataforma electrónica que comenzó a funcionar en diciembre de 2020: https://denuncia.sma.gob.cl/

Adicionalmente, se ha considerado el desarrollo de distintas herramientas informáticas y analíticas con objeto de optimizar la gestión de denuncias, incluyendo un modelo de clasificación y priorización de éstas, en función de la información ingresada por los denunciantes y la experiencia de analistas de la SMA, y la automatización de los procesos relacionados con la norma de ruidos, principal motivo de denuncias.

## Facilitar información pública

Desde el año 2008, Chile cuenta con la Ley 20.285 sobre acceso a la información pública, que establece que toda persona tiene derecho a solicitar y recibir información de órganos de la Administración del Estado. En particular, el derecho de acceso a la información pública genera diversos beneficios, incluyendo la profundización de la democracia, un incentivo a la participación ciudadana y el acceso a otros derechos, por lo cual ha sido denominado “derecho llave” (CPLT, 2018). Adicionalmente, la Ley Orgánica de la SMA establece que la institución debe administrar el Sistema Nacional de Información de Fiscalización Ambiental (SNIFA), de acceso público y conformado por lo antecedentes y datos que maneja con respecto al cumplimiento de instrumentos de carácter ambiental, procesos de fiscalización y sancionatorios, entre otros (Figura 23). Sin embargo, más allá de las obligaciones establecidas, es importante reconocer que la transparencia incentiva el cumplimiento normativo, al ayudar a los regulados a comprender su impacto ambiental y permitir que la comunidad comprenda y actúe sobre esta información para promover la justicia ambiental (EPA, 2014). **La transparencia de la información genera un efecto disuasivo sobre la comunidad regulada y cumple un rol clave para asegurar el cumplimiento normativo.**

***Figura 23:*** *Sistema Nacional de Información de Fiscalización Ambiental (https://snifa.sma.gob.cl/).*

En este contexto, los resultados de la PID se deben publicar automáticamente en SNIFA, como una acción directa que servirá para promover el uso de los datos ambientales y el desarrollo científico, de acuerdo con los principios FAIR (del inglés FAIR: Findable, Accessible, Interoperable, Reusable). Estos principios fueron publicados en la revista Nature (Wilkinson et al. 2016) y ofrecen un conjunto de cualidades precisas y medibles para que los datos sean Encontrables, Accesibles, Interoperables y Reutilizables. De esta manera, los principios están concebidos para facilitar las capacidades de las máquinas o sistemas computacionales para encontrar, acceder, interoperar y reutilizar datos con ninguna o mínima intervención humana, considerando que las personas dependen sustancialmente del soporte computacional para manejar datos, como resultado del aumento en el volumen, complejidad y velocidad en la creación de ellos. Particularmente, se considera ampliar y fortalecer la publicación de todos los antecedentes que maneja la SMA en el SNIFA, poniéndolos a disposición de la ciudadanía, Universidades, Centros de Investigación y desarrolladores de tecnologías para que todos los sectores contribuyan a generar información y conocimiento sobre el estado del medio ambiente.

**Proyecto destacado:**

## Sistema de Seguimiento Ambiental 2.0

Se trata de un proyecto ícono para el desarrollo de la PID, que permitirá generar más información y conocimiento sobre el estado del medio ambiente en Chile. El Sistema de Seguimiento Ambiental (SSA) es el que utilizan los titulares de proyectos y/o actividades que cuentan con Resoluciones de Calificación Ambiental (RCA) vigentes para el reporte electrónico de sus compromisos y obligaciones asociadas a mediciones, análisis y datos de diversos parámetros ambientales (biodiversidad, agua, fungi, aire, patrimonio, etc.), como parte de planes y programas de vigilancia ambiental.

Desde el año 2013, la SMA ha recibido más de 100.000 reportes de seguimiento ambiental, principalmente en formato PDF, que se encuentran disponibles en SNIFA y que contienen millones de datos en su interior. Justamente, el proyecto consiste en la renovación completa del sistema, tanto en su arquitectura, como en su interfaz y en la definición de formatos de reporte estandarizados para facilitar la validación, almacenamiento y procesamiento de datos, lo que permitirá no solo automatizar la verificación del cumplimiento, sino que también generar una base de datos única sobre el estado del medio ambiente en permanente actualización gracias a los reportes de miles de empresas reguladas (Figura 24).

Con respecto a los avances, ya existe un formato estandarizado para datos relacionados con aguas superficiales, subterráneas y marinas (incluyendo sedimentos) y durante el segundo semestre del año 2021 se formalizará el formato para reportar datos de biodiversidad (flora, fauna y fungi). En tanto, para el desarrollo de algoritmos y modelos predictivos se contará con el apoyo de la Fundación Ciencia y Vida, gracias a un proyecto financiado por el Fondo de Fomento Ciencia y Tecnología (FONDEF) IdeA I+D 2021.

De esta forma, se generará una base de datos consolidada con respecto a los principales indicadores de los efectos del cambio climático (biodiversidad y agua) y poner a disposición estos datos para la academia y comunidad en general mediante SNIFA, contribuyendo al desarrollo de análisis de series temporales y modelos predictivos, entre otros.

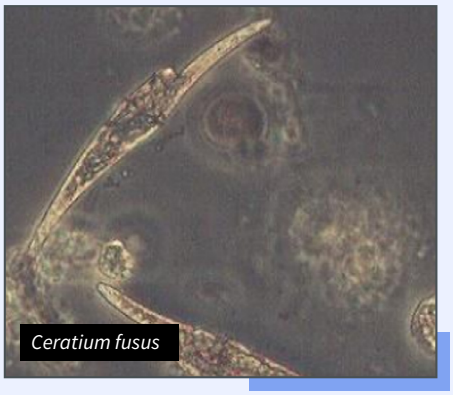



***Figura 24:*** *Algunas especies monitoreadas en el Programa de Vigilancia Ambiental de Bahía de Mejillones en verano 2019, reportado a la SMA mediante SSA y disponible en* ***SNIFA.***

## Promover el cumplimiento normativo

Uno de los ejes claves de esta estrategia es la entrega temprana de retroalimentación a la comunidad regulada, con respecto al nivel de cumplimiento normativo y al estado del medio ambiente, a través de reportes automatizados y el desarrollo de un Portal del Regulado (Figura 25) que incluya historial de comportamiento, alertas tempranas e indicadores ambientales en el territorio donde se emplazan los proyectos y/o actividades. Por lo tanto, más allá de perseguir infracciones a la regulación ambiental y activar los procesos sancionatorios que correspondan, también se plantea utilizar los resultados de la PID para incentivar el cumplimiento del sector privado, entregando información temprana para que adapten sus procesos y/o actividades de forma proactiva y así prevenir afectaciones futuras. La retroalimentación periódica busca reforzar significativamente el comportamiento del sector regulado, promoviendo el cumplimiento y evitando la materialización de impactos significativos.

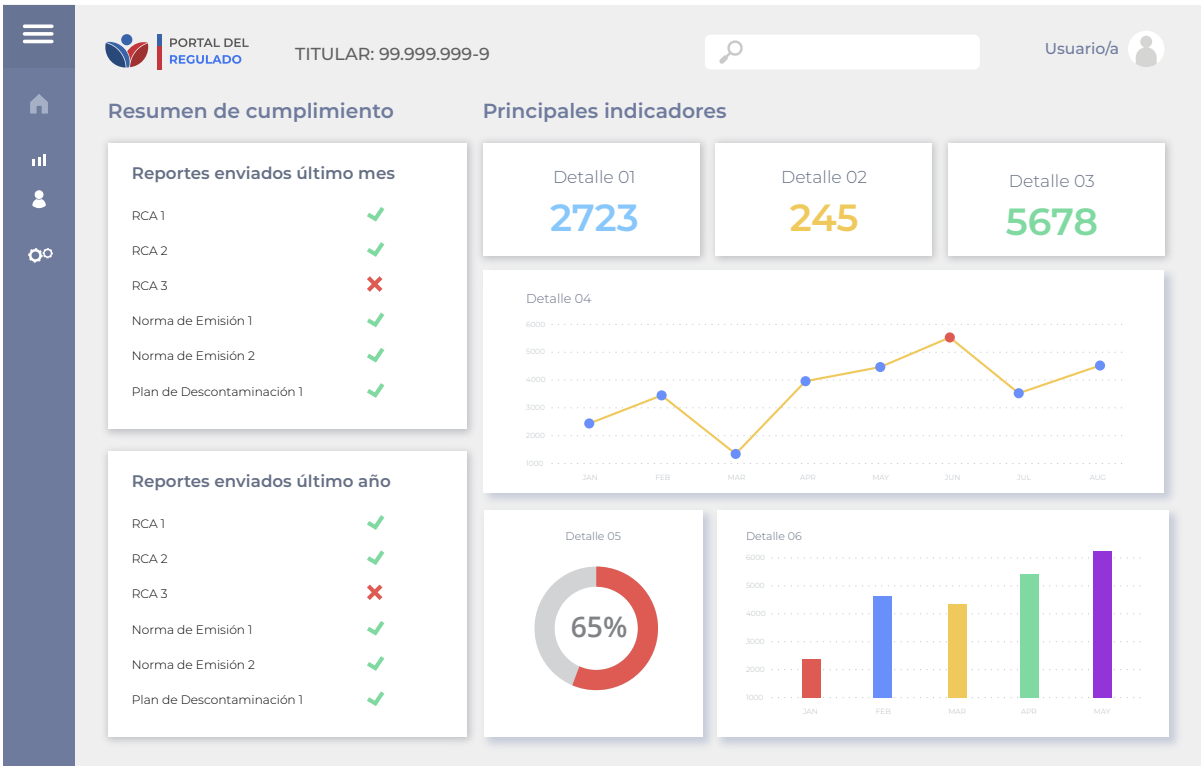



***Figura 25:*** *Maqueta de Portal del Regulado con información sobre el cumplimiento normativo de un proyecto y el estado del medio ambiente en su entorno cercano.*

En este sentido, se deben considerar elementos de las ciencias del comportamiento para lograr una comunicación efectiva que permita facilitar la toma de decisiones de las personas, lo cual es particularmente necesario para transmitir los efectos del cambio climático e influenciar cambios de conducta (Pidgeon & Fischhoff, 2011). En particular, para la SMA es clave facilitar decisiones de la comunidad regulada que permitan promover el cumplimiento de la normativa, comunicando los riesgos asociados en casos de incumplimiento, considerando tanto los efectos ambientales, como las sanciones aplicables correspondientes.

## Comunicación de riesgo

Cada vez es más popular en agencias de gobierno el uso de mensajes diseñados mediante técnicas de economía del comportamiento, debido a que son medidas altamente costo eficientes (Benartzi et al, 2017; Sunstein, 2014; Thaler, 2018; Loewenstein & Chater, 2017). Para la presente estrategia se propone realizar la comunicación de riesgos de manera no-persuasiva, reconociendo que los receptores de estos mensajes pueden optar por diferentes formas de respuesta (Fischhoff, 2007).

Si bien se busca incentivar una respuesta ambientalmente positiva, un potencial infractor puede preferir exponerse a una reacción sancionatoria. En esos casos puede ser necesario calibrar simultáneamente la potencial multa para dar una señal disuasiva que permita incentivar un comportamiento responsable (Fischhoff, 2007). Para construir una comunicación de riesgo efectiva, se propone integrar conocimientos de las ciencias sociales y de las decisiones en su diseño (Wong-Parodi et al, 2016).

Esto permitirá una comprensión mayor de cómo las personas interpretan la información presentada, qué opciones prefieren tomadores de decisión en presencia de información completa, cómo se realizan elecciones al experimentar sesgos cognitivos o en escenarios adversariales, y qué intervenciones realizar para reducir la diferencia entre el comportamiento ideal y la realidad (Wong-Parodi et al, 2016). Con esto, se busca facilitar decisiones informadas, permitiendo la comprensión adecuada de riesgos, beneficios e incertidumbres, para que las decisiones adoptadas reflejen preferencias individuales (Wong-Parodi, 2016; Fischhoff & Davis, 2014).

Con esta información y de manera iterativa se debe depurar el mensaje, utilizando la experimentación y aplicando técnicas de análisis cualitativo y cuantitativo, determinando con precisión los modelos mentales existentes y el atractivo de las diferentes alternativas disponibles (Wong-Parodi et al, 2016). Para esto, un estándar de experimentación altamente utilizado corresponde a los ensayos de control aleatorio, conocidos como pruebas A/B, que permiten comparar 2 versiones de un producto para identificar cual tiene una mayor efectividad, en función de los resultados deseados (Kohavi & Longbotham, 2017). Por ejemplo, en la SMA se pueden probar variaciones de mensajes, ya sea en correos electrónicos o en sistemas de reporte electrónico, que generen una mejor respuesta por parte de la comunidad regulada.

**Proyecto destacado:**

## Verificación del cumplimiento de las normas de RILES

Un caso emblemático es la intervención realizada para promover el cumplimiento de las normas de emisión de descargas de Residuos Líquidos Industriales, RILES (D.S. 90/2000 y D.S. 46/2002 MINGSEGPRES). Para esto, se automatizó el proceso de verificación normativa y la generación de reportes de cumplimiento que se envían mensualmente a las empresas, lo que ha permitido alcanzar una cobertura total de las fuentes afectas (700 aproximadamente, con más de 800 ductos) y, lo más importante, un nivel de cumplimiento histórico. En efecto, desde el año 2017, la cantidad de ductos en conformidad y no conformidad (líneas azul y naranja -Figura 26-, respectivamente), generándose un aumento marcado del cumplimiento normativo de forma posterior al envío de reportes realizado por primera vez en abril del 2020.

De esta forma, la entrega de retroalimentación logró cambios significativos y permanentes en el comportamiento de las empresas, sin perjuicio de las acciones sancionadoras que se ejecutaron de forma posterior y que, sin duda, incrementaron los incentivos para alcanzar un nivel de cumplimiento aún mayor. Esto demuestra que estas acciones basadas en datos generan beneficios directos para el medio ambiente y las personas, lo que constituye un modelo de referencia para aplicar en la gestión de todos los instrumentos de carácter ambiental.

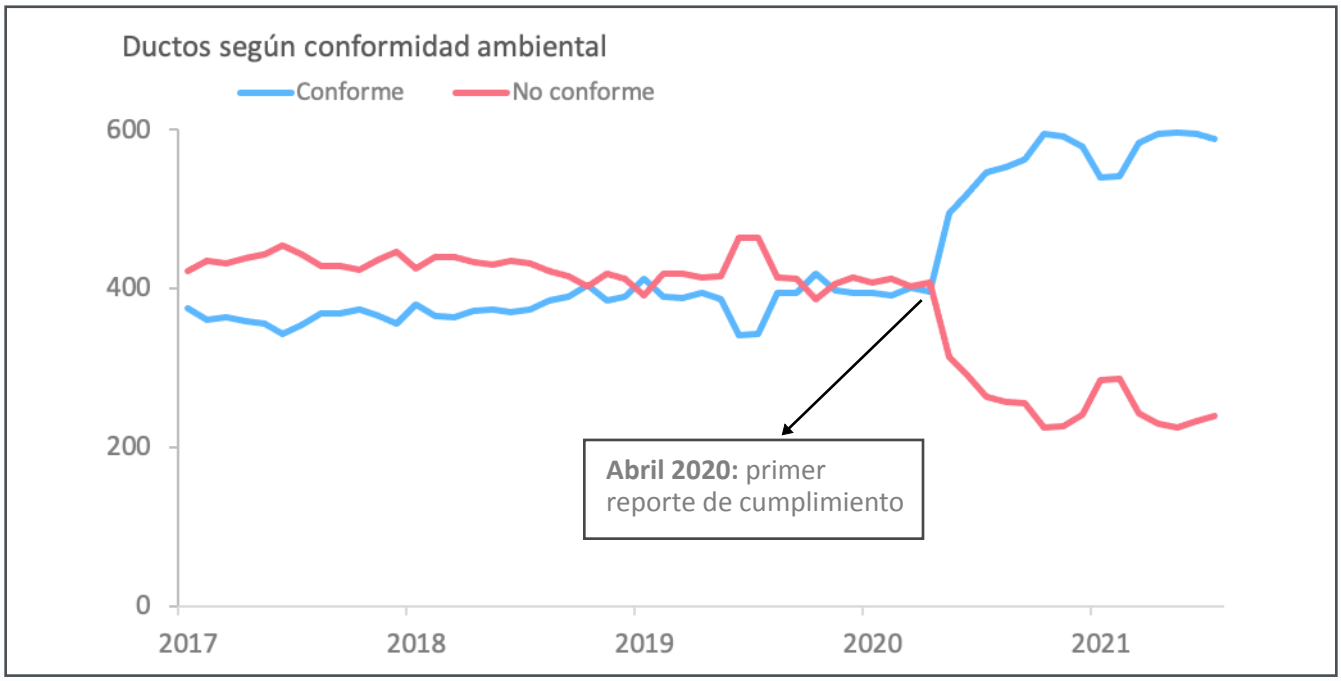



***Figura 26:*** *El nivel de cumplimiento de las normas de RILES aumentó significativamente el año 2020, desde la automatización del proceso de verificación y el envío de reportes periódicos a las empresas realizado por primera vez en el mes de abril.*

## Informar a otros servicios

También interesa proporcionar antecedentes al Servicio de Evaluación Ambiental (SEA), al Ministerio del Medio Ambiente (MMA) y Organismos Sectoriales con competencia ambiental, con respecto a la eficiencia y eficacia de la regulación existente, y sobre el estado del medio ambiente y la evolución de los ecosistemas. De este modo, será posible revisar y fortalecer el diseño normativo de manera permanente, ya sea mediante la adaptación de instrumentos existentes y/o la priorización de nuevas regulaciones, considerando la información disponible.

En particular, la generación de una base de datos ambientales consolidada y en permanente actualización (agua, aire, biodiversidad, etc.), junto con la automatización de la verificación del cumplimiento de los distintos instrumentos de carácter ambiental, permitirán avanzar en la comprensión de si efectivamente se cumplen los objetivos de protección trazados. Se trata de información fundamental para apoyar la toma de decisiones, potenciando tanto el Sistema de Evaluación de Impacto Ambiental (SEIA), como el diseño normativo por parte del Ministerio del Medio Ambiente. Para esto, se considera la interoperabilidad e integración de datos automatizada con otros organismos (vía APIs), junto con la coordinación de espacios para presentar los resultados de la PID y sus modelos analíticos.

Proyecto destacado

## Monitoreo satelital de humedales

Se encuentra en desarrollo un programa de monitoreo satelital de los distintos tipos de humedales existentes en el país, con énfasis en los humedales altoandinos y urbanos en una primera etapa, considerando la aprobación de la Ley 21.202 que busca proteger estos últimos. De esta forma, el proyecto consiste en desarrollar un análisis histórico de los humedales en Chile, en función de la disponibilidad de imágenes satelitales de acceso libre durante los últimos 30 años, identificando los quiebres en las series temporales de espejos de agua y de la vigorosidad de la vegetación (Figura 27).

Adicionalmente, también interesa identificar afectaciones en el tiempo presente, apenas comiencen a materializarse, de modo de activar alertas tempranas cuando corresponda, por medio del proyecto SAMSARA (Sistema de Alerta y Monitoreo Satelital de Áreas de Relevancia Ambiental), el cual se desarrollará en conjunto con la Universidad Adolfo Ibáñez, gracias al Fondo de Fomento Ciencia y Tecnología (FONDEF) IdeA I+D 2021. Se trata de una iniciativa que generará información relevante no solo para los procesos de la SMA, sino que también para diversos organismos, incluyendo el Ministerio del Medio Ambiente y Municipalidades, entre otros.

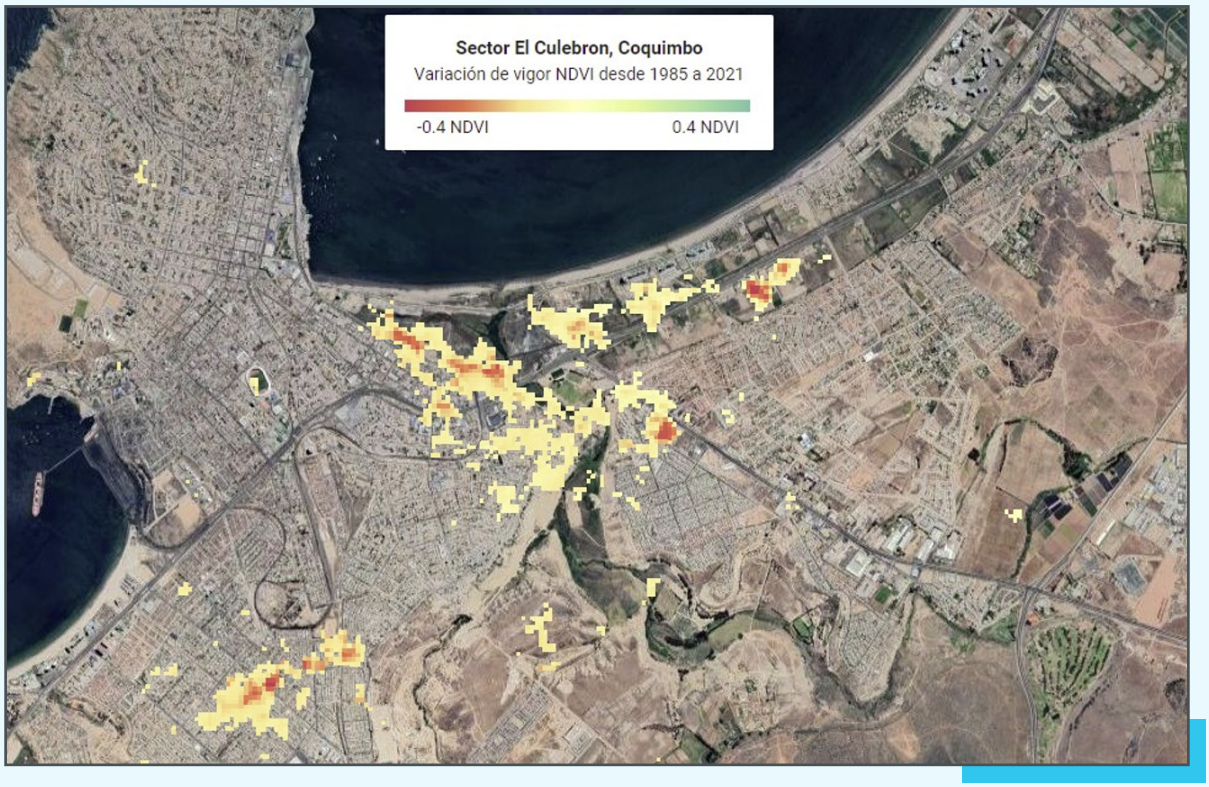



***Figura 27:*** *Monitoreo satelital mediante índice NDVI en humedales de Coquimbo.*

## 4.5 Aprendizaje

Como ya fue descrito, una de las principales características que identifica la presencia de inteligencia es la capacidad de aprender en base a la experiencia. Así como los seres humanos aprenden, adaptando su comportamiento en el tiempo y las máquinas se entrenan permanentemente al observar nueva información mediante el aprendizaje automático (OECD, 2019b), la organización misma también aprende ante experiencias exitosas y fallidas.

En este contexto, el desarrollo de proyectos de ciencia de datos e inteligencia artificial en la SMA se sustenta en la metodología CRISP-DM (Chapman et al, 2000), consistente en las seis fases cíclicas mostradas en la Figura 28, que incluyen desde la comprensión de las necesidades de la institución, al entendimiento y preparación de datos, para luego modelar, evaluar e implementar soluciones analíticas, con una retroalimentación permanente en sus distintas fases y un foco en la identificación de oportunidades para mejorar la calidad de los datos existentes y capturar nuevos datos. Por lo tanto, se trata de un enfoque que releva la importancia de un trabajo colaborativo y multidisciplinario entre las distintas áreas y personas de la institución, que puede ser apalancado mediante iniciativas conjuntas con otras organizaciones.

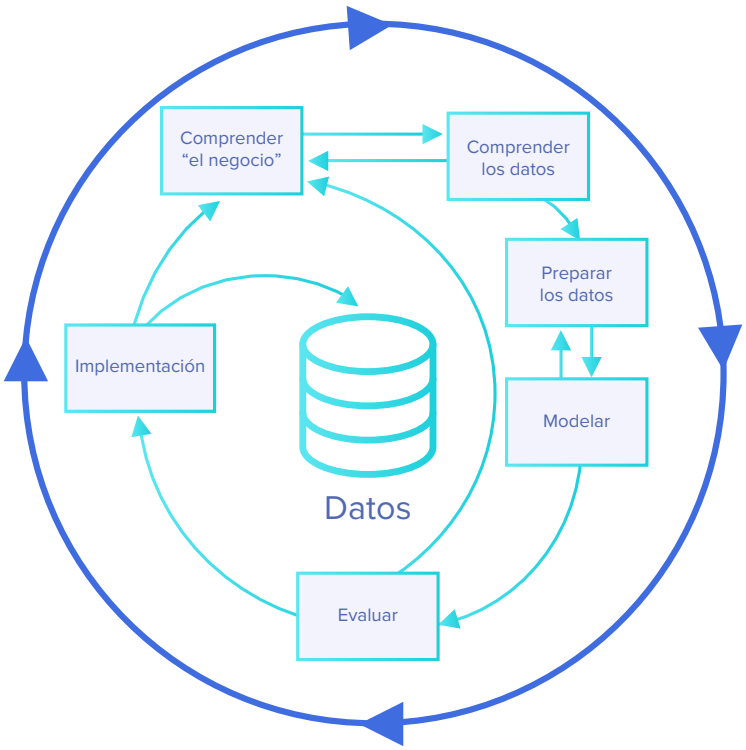



***Figura 28:*** *Metodología de desarrollo de proyectos CRISP-DM (adaptado de Chapman et al, 2000)*

Tecnologías cognitivas como las que se propone desarrollar en la SMA requieren de un reconocimiento de los elementos esenciales del aprendizaje en base a la experiencia, especialmente en lo relacionado con los procesos dinámicos de toma de decisiones en la organización. De acuerdo con Simon (1991a), mientras el ser humano aprende a nivel individual, una organización aprende mediante el aprendizaje de sus miembros o integrando nuevas personas que tengan conocimientos adicionales.

Por lo tanto, el aprendizaje de individuos depende de los conocimientos presentes en la organización, accesibles por sus miembros mediante interacciones robustas (Krackhardt et al, 2003) y almacenados en la memoria institucional, la cual aborda equipos multidisciplinarios de alta experiencia profesional, bases de datos y bibliotecas de documentos, o instructivos procedimentales (Simon, 1991a). A partir de este conocimiento, es posible responder de manera rápida e intuitiva a diversas situaciones que se presentan por analogía, codificando las reglas construidas en base a la experiencia en sistemas computacionales (OECD, 2019b; Laird et al, 1987; Newell, 1990; Anderson, 1996).

El ser humano usa incrementalmente todo el conocimiento previo acumulado para tomar decisiones (González et al, 2003; Langley & Simon, 1981). Así, una decisión comienza por el reconocimiento de las alternativas disponibles en un contexto determinado, las cuales son evaluadas y clasificadas según su similitud con experiencias previas almacenadas en la memoria, para luego elegir la mejor alternativa y proceder a ejecutar la acción correspondiente. A medida que se van enfrentando nuevas decisiones en el tiempo es posible refinar la evaluación de alternativas y retroalimentar decisiones futuras (Figura 29) (González et al, 2003).

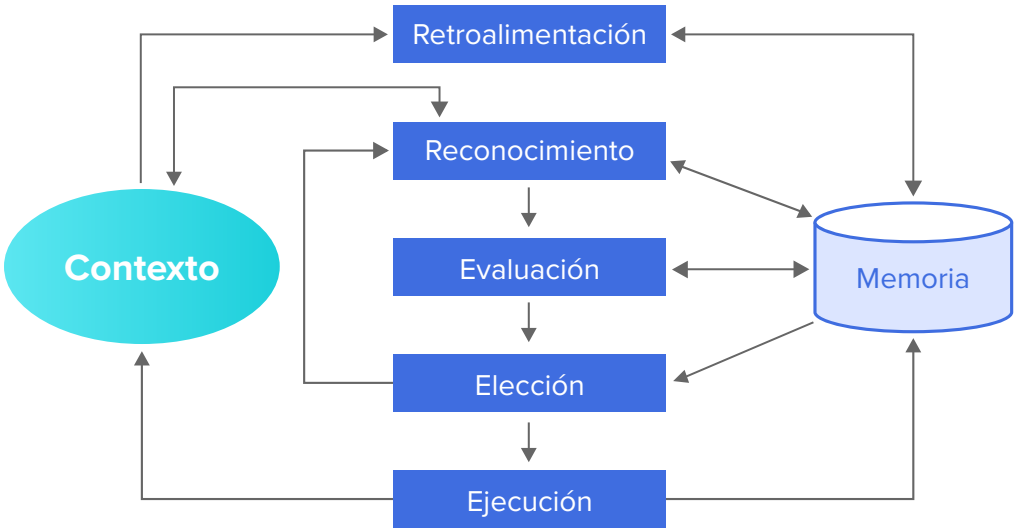



***Figura 29:*** *Principales mecanismos de la toma de decisiones en base a la experiencia (basado en González et al, 2003)*

Por lo tanto, es clave mantener activa la memoria de individuos, grupos de trabajo, y de toda la organización, con políticas de selección y retención de talento, manteniendo bases de datos y sistemas internos de documentación fácilmente accesibles y generando mecanismos de retroalimentación para revisar y aprender de situaciones pasadas, contribuyendo a la activación de la memoria y la recuperación de información almacenada (NSTC, 2019).

De este modo, se plantea que la presente estrategia sea actualizada continuamente y retroalimentada a partir de la experiencia. En particular, se propone seguir el método de estándares abiertos (Russell, 2014), incorporando nuevas propuestas en el tiempo y actualizando la estrategia de manera dinámica. Para esto, anualmente se desarrollará un proceso interno de planificación estratégica para definir las actividades y productos que deben ser priorizadas para su desarrollo y se mantendrán guías técnicas con las mejores prácticas que permitan asegurar la integridad de los productos desarrollados, la prueba de conceptos, la selección de las mejores tecnológicas disponibles y el escalamiento en base a soluciones probadas (Eggers et al, 2019).

# 5. Hoja de ruta

La hoja de ruta para materializar esta estrategia se enfoca en el desarrollo de las distintas capas que conforman la Plataforma de Inteligencia de Datos, incluyendo: **Captura de datos, Analítica avanzada y Acciones;** según lo planteado en secciones anteriores de este documento.

Para esto, se considera el desarrollo de diversos proyectos y líneas de trabajo, que tienen distintos niveles de avance y consolidación. De este modo, algunos proyectos incipientes se enfocan en la captura de datos, mientras que otros se orientan a la explotación de éstos, por medio de la implementación de modelos analíticos. La Figura 30 resume la hoja de ruta y objetivos generales de cada año. Destaca un avance progresivo desde la captura de los datos, para posteriormente potenciar la analítica avanzada, lo que permitirá finalmente concentrar los esfuerzos en generar acciones disuasivas basada en la evidencia reunida.

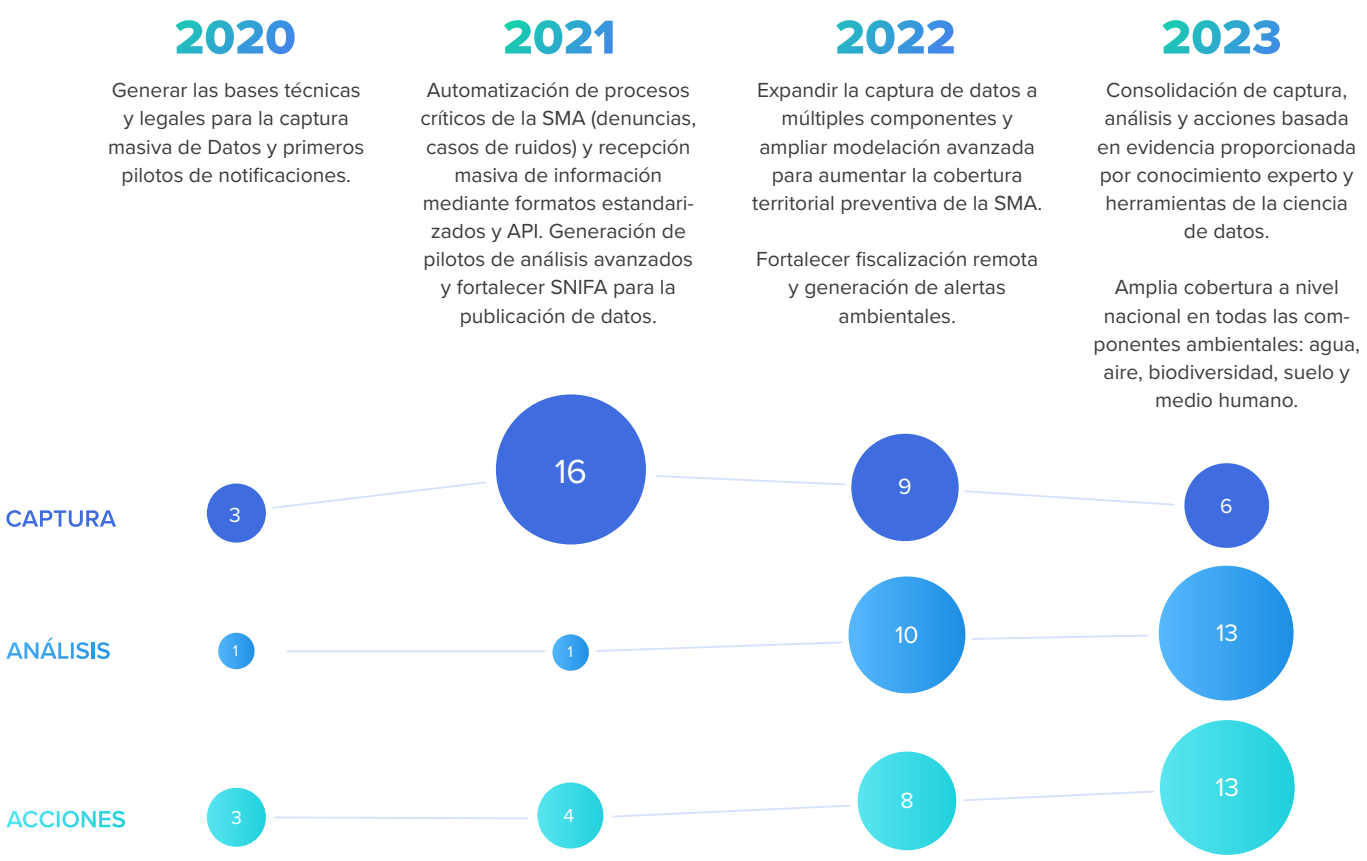



***Figura 30:*** *Total de proyectos considerados por año en la hoja de ruta.*

Cabe señalar que, sin perjuicio de lo señalado a continuación, esta hoja de ruta podría experimentar modificaciones en función de las necesidades de la institución, la disponibilidad de recursos y las recomendaciones del Comité Asesor de la Estrategia, entre otros factores, lo cual se difundirá oportunamente.

## 5.1 Captura de datos

Se plantea el siguiente esquema de captura de datos de presiones, variables de estado y respuestas, considerando las normativas ambientales y las instrucciones de reporte establecidas por la SMA. Esta planificación incluye datos de diferente naturaleza y orígenes, múltiples componentes ambientales e instituciones públicas como privadas, respectivamente. Con esto se propone intensificar el reporte periódico de datos, vía API, como alternativa al reporte manual.

| Indicador FPEIR | Componente o materia | Proyecto o iniciativa | Normativa aplicable | Tipo de registro | Instrucción SMA | Plazo |
|---|---|---|---|---|---|---|
| Fuerzas motrices | Unidades Fiscalizables | Catastro de unidades fiscalizables sujetas a la normativa ambiental | General | Discreto | Norma interna | 2021 |
| | Resoluciones de Calificación Ambiental | Catastro de proyectos y/o actividades, según antecedentes del Servicio de Evaluación Ambiental. | RCA | Discreto | Norma interna | 2021 |
| | Establecimientos del RETC | Establecimientos registrados en RETC del Ministerio del Medio Ambiente | General | Discreto | Norma interna | 2021 |
| Presiones | Descargas de RILES | Descargas en aguas superficiales y subterráneas a nivel nacional (caudales y concentraciones) | DS 90/2000, DS 46/2002 MINSEGPRES | Discreto | 222/2017 | 2020 |
| | Residuos | Disposición de toneladas de residuos domiciliarios en rellenos sanitarios y vertederos | RCA | Continuo | 431/2021 | 2021 |
| | Residuos | Monitoreo de variables de flujo y paráme-tros fisicoquímicos en relaves | RCA | Continuo | Pendiente | 2021 |
| | Residuos | Generación, reutilización, reciclaje y valorización de residuos | Ley REP | Continuo y discreto | Pendiente | 2022 |
| | Emisiones Atmosféricas | Sistemas de monitoreo continuos (CEMS) comprometidos en distintos instrumentos | General | Continuo | 1574/2019, 680/2021, 862/2020 | 2020 |
| | Emisiones Atmosféricas | Reporte estandarizado de estimaciones, muestreos y mediciones comprometidas en distintos instrumentos | General | Discreto | Pendiente | 2022 |
| | Emisiones Atmosféricas | Conexión en línea y reporte periódico de variables operacionales de calderas, turbinas y procesos con combustión, para la estimación de emisiones | General | Discreto y continuo | 2452/2020, 1814/2021 | 2022 |
| | Ruidos | Reporte de monitoreo de fuentes emisoras en formato estandarizado | DS 38/2011 MMA | Discreto | Pendiente | 2022 |
| | Ruidos | Monitoreo permanente de fuentes emisoras relevantes vía API | RCA | Continuo | Pendiente | 2022 |
| | Contaminación lumínica | Fuentes emisoras en regiones afectas a la norma | DS 43/2013 MMA | Discreto | 434/2019 | 2021 |
| | Olores | Catastro de fuentes emisoras de olores y emisiones | RCA | Discreto y continuo | Pendiente | 2022 |
| | Uso de recursos naturales | Caudales de extracción de salmuera y aguas subterráneas en la cuenca del Salar de Atacama | RCA | Continuo y discreto | 1314/2020 | 2021 |

| Indicador FPEIR | Componente o materia | Proyecto o iniciativa | Normativa aplicable | Tipo de registro | Instrucción SMA | Plazo |
|---|---|---|---|---|---|---|
| | Uso de recursos naturales | Caudal utilizado en generación eléctrica en centrales de pasada | RCA | Continuo | Pendiente | 2023 |
| | Áridos | Líneas base de topografía mediante drones para cuencas priorizadas con extracción de áridos | RCA | Discreto | Norma interna | 2022 |
| **Estado** | Biodiversidad | Reporte de monitoreo de biodiversidad a nivel nacional | RCA | Discreto | Pendiente | 2021 |
| | Agua | Niveles de agua subterránea y salmuera en la cuenca del Salar de Atacama | RCA | Continuo | 1314/2020 | 2021 |
| | Agua | Monitoreo de calidad y cantidad de agua superficial y subterránea en proyectos con compromisos ambientales de RCA | RCA | Discreto | 894/2019 | 2020 |
| | Agua | Calidad en aguas marinas donde se ubican Centros de Engorda de Salmones. | RCA | Continuo | 2019/2020 | 2021 |
| | Aire | Monitoreo de calidad del aire considerando monitoreo en línea y monitoreo discreto | General | Continuo y discreto | 126/2020, pendiente | 2022 |
| | Meteorología | Monitoreo de variables climáticas como viento, humedad, presión, temperatura, precipitaciones, nieves, hielos, entre otros | General | Continuo | 126/2020 | 2021 |
| | Suelo | Monitoreo de calidad del suelo | General | Discreto | Pendiente | 2023 |
| **Impacto** | Salud humana | Datos de eventos de morbilidad y mortalidad en centros de salud | General | Discreto | Pendiente | 2023 |
| **Respuestas** | Denuncias | Mejoras de captura de datos, mediante Portal Ciudadano de Denuncias (SIDEN) | LOSMA | Discreto | Norma interna | 2021 |
| | Fiscalizaciones | Mejoras de captura de datos, mediante el Sistema de Fiscalización (SISFA) | LOSMA | Discreto | Norma interna | 2021 |
| | Procesos sancionatorios | Mejoras de captura de datos, mediante el Sistema de Procesos Sancionatorios (SIPROS) | LOSMA | Discreto | Norma interna | 2021 |
| | Medidas provisionales | Mejoras de captura de datos, mediante el Sistema de Fiscalía (SIF) | LOSMA | Discreto | Norma interna | 2021 |
| | Normativa ambiental | Actualización de catastro, según antecedentes del Ministerio del Medio Ambiente | General | Discreto | Norma interna | 2021 |
| | Resoluciones de Calificación Ambiental | Registro de compromisos y obligaciones de reporte en el Sistema de Seguimiento Ambiental (SSA) | RCA | Discreto | Pendiente | 2022 |
| | Jurisprudencia | Fallos ilustres Tribunales Ambientales | General | Discreto | Norma interna | 2023 |
| | Jurisprudencia | Fallos Cortes de Apelaciones y Corte Suprema | General | Discreto | Norma interna | 2023 |
| | Dictámenes | Dictámenes de Contraloría General de la República | General | Discreto | Norma interna | 2023 |

*Instrucción SMA “pendiente” se refiere a documentos en proceso de elaboración o planificados.

## 5.2 Analítica avanzada

A continuación, se presentan los proyectos o iniciativas consideradas en el ámbito de la analítica avanzada, las cuales tienen por objetivo sentar las bases para generar acciones basadas en datos. Las iniciativas incluyen diferentes grados de complejidad, desde verificaciones automatizadas de normas ambientales, con el objeto principal de disminuir tiempos de procesos internos, hasta la generación de modelos predictivos o prescriptivos que guíen el accionar de la institución en los diferentes ámbitos de su competencia.

| Indicador FPEIR | Aplicación | Proyecto o iniciativa | Plazo |
|---|---|---|---|
| **Analítica descriptiva** | Fiscalización | Verificación del cumplimiento de la norma de ruidos, principal causa de denuncias | 2021 |
| | | Verificación del cumplimiento de las normas de descargas de RILES | 2020 |
| | | Verificación del cumplimiento de la norma de emisión de Centrales Termoeléctricas | 2022 |
| | | Verificación del cumplimiento de la norma de emisión de Fundiciones de cobre | 2022 |
| | | Verificación del cumplimiento de los reportes establecidos en Planes de Prevención y Descontaminación Ambiental | 2023 |
| | | Verificación del cumplimiento de los reportes establecidos Resoluciones de Calificación Ambiental | 2023 |
| | | Verificación del cumplimiento de Centros de Engorda de Salmón (RAMA y RCA) | 2022 |
| | | Verificación de cumplimiento automatizado de normas primarias de calidad del aire | 2022 |
| | | Verificación de cumplimiento automatizado de normas secundarias de calidad del agua | 2022 |
| **Analítica Diagnóstico** | | Modelación de Meteorología y Calidad del Aire en territorios prioritarios | 2022 |
| | | Modelación Olores Molestos en territorios prioritarios | 2023 |
| | | Monitoreo de operación de tranques de relaves mediante imágenes satelitales y variables en línea | 2023 |
| | | Monitoreo de deforestación submarina (bosques de huiro) mediante imágenes satelitales | 2023 |
| | | Monitoreo de turberas mediante imágenes satelitales | 2023 |
| **Analítica predictiva** | Seguimiento | Predicción de calidad del aire en zonas de interés, a través de modelos que incluyan meteorología y emisiones atmosféricas | 2023 |
| | Seguimiento | Módulos analíticos focalizados en la componente biodiversidad permitan entender la relación entre presiones y estado, incluyendo variables climáticas | 2023 |

| Indicador FPEIR | Aplicación | Proyecto o iniciativa | Plazo |
|---|---|---|---|
| | Seguimiento | Módulos analíticos focalizados en la componente agua permitan entender la relación entre presiones y estado, incluyendo variables climáticas | 2023 |
| | Seguimiento | Modelos para detectar cambios en series temporales de humedales y realizar proyecciones, utilizando imágenes satelitales ópticas | 2022 |
| | Seguimiento | Modelos para identificar ubicación de Centros de Engorda de Salmones, utilizando imágenes satelitales de tipo radar | 2022 |
| | Seguimiento | Modelo de pronóstico de bloom de algas en lago Villarrica | 2023 |
| | Fiscalización | Modelos que permitan clasificar e identificar anomalías en emisiones reportadas por distintas fuentes (fundiciones, calderas, termoeléctricas, etc.) | 2023 |
| **Analítica prescriptiva** | Fiscalización | Modelo que oriente el programa de fiscalización y que recoja las experiencias y preferencias de fiscalizadores, por medio de escenarios | 2022 |
| | Denuncia | Modelo que clasifique denuncias y oriente acciones para su gestión, en función de los datos históricos | 2022 |
| | Sanción | Modelo que clasifique hallazgos y oriente acciones para su gestión, en función de los datos históricos | 2023 |
| | Cumplimiento | Modelo que evalúe efectividad de incentivos al cumplimiento como programas de cumplimiento, en función de datos históricos | 2023 |

## 5.3 Acciones

La siguiente lista la componen las acciones realizadas y planificadas en el período 2020-2023 dirigidas principalmente hacia actores claves externos a la SMA. Estas se conforman de acciones generadas por la PID de forma directa (sistemas de toma de decisiones), para apoyar la toma de decisiones (sistemas de soporte), publicación de información a la ciudadanía (desde datos brutos a plataformas de monitoreo), así como acciones para informar a organismos públicos, de modo de colaborar activamente en la mejora de la normativa ambiental y mejorar la coordinación interinstitucional.

| Componente o materia | Proyecto o iniciativa | Plazo |
|---|---|---|
| **Optimizar la gestión interna** | Generar documentación automatizada para la gestión de denuncias, incluyendo derivación a organismos sectoriales y archivo de expedientes, entre otros | 2021 |
| | Crear alertas sobre plazos de denuncias y de otros procesos | 2022 |
| | Crear alertas sobre superación de umbrales de emisiones y calidad del aire en zonas prioritarias | 2022 |
| | Generar informes de fiscalización estandarizados masivamente para distintos instrumentos, incluyendo RCA, normas de emisión, Planes de Descontaminación, entre otros | 2022 |
| | Generar documentos automatizados para apoyar procesos de sanción, cumplimiento y medidas provisionales | 2023 |
| | Generar herramienta de planificación del ciclo de programación | 2023 |
| **Facilitar información pública** | Publicar reportes de seguimiento ambiental de titulares de proyectos y/o actividades que cuentan con Resoluciones de Calificación Ambiental | 2020 |
| | Publicar Infraestructura de Datos Espaciales (IDE) con información de la SMA y otras coberturas ambientales relevantes | 2020 |
| | Publicar reportes asociados al impuesto verde, norma de fundiciones y Planes de Descontaminación | 2021 |
| | Actualizar sección de estadísticas en SNIFA que aborde integralmente los procesos de la SMA y aspectos ambientales específicos | 2022 |
| | Publicar resultados del programa de monitoreo de humedales, basado en el uso de imágenes satelitales | 2022 |
| | Publicar Geoportal en SNIFA, sitio web que integre todas las herramientas espaciales de la SMA | 2022 |
| | Publicar datos de presiones ambientales en tiempo real, capturados mediante sensores remotos | 2023 |
| | Publicar datos estado del medio ambiente en tiempo real, capturados mediante sensores remotos | 2023 |
| | Publicar base de datos consolidada de componentes ambientales reportados por titulares de proyectos y/o actividades que cuentan con Resoluciones de Calificación Ambiental | 2023 |
| | Publicar nivel de cumplimiento por unidad fiscalizable y titular, de forma consolidada, incluyendo los distintos instrumentos de carácter ambiental aplicables | 2023 |
| | Habilitar APIs para compartir datos de variables ambientales y procesos de la SMA | 2023 |

| | | |
|---|---|---|
| **Promover el cumplimiento normativo** | Entregar retroalimentación permanente a las fuentes afectas a normas de RILES | 2020 |
| | Crear alertas ante caídas de conexiones en línea para el reporte de variables a través de API | 2022 |
| | Entregar retroalimentación permanente a los proyectos y/o actividades con RCA vigente | 2023 |
| | Entregar retroalimentación permanente a las fuentes emisoras sujetas a Planes de Prevención y Descontaminación Ambiental | 2023 |
| | Habilitar Portal del Regulado, que permita consolidar compromisos de reporte y facilitar información sobre cumplimiento normativo y estado del medio ambiente | 2023 |
| | Habilitar Geoportal interno, sitio web que integre todas las herramientas espaciales de la SMA | 2021 |
| **Informar a otros servicios** | Desarrollar y compartir aplicaciones web de datos en líneas y espaciales con organismos públicos con competencias ambientales | 2021 |
| | Convocar seminario anual para compartir los resultados y lecciones de esta estrategia con otros organismos públicos y miembros de la Red Nacional de Fiscalización Ambiental (RENFA) | 2022 |
| | Publicar reporte consolidado anual al Ministerio del Medio Ambiente sobre los niveles de cumplimiento y efectividad de las medidas consideradas en distintos instrumentos de carácter ambiental | 2023 |
| | Publicar reporte consolidado anual al Servicio de Evaluación Ambiental sobre los niveles de cumplimiento de las medidas y condiciones establecidas en las Resoluciones de Calificación Ambiental | 2023 |
| | Habilitar APIs para compartir información de la PID con otros organismos públicos | 2023 |

# 6. Referencias


Ackoff, R. L. (1989). From data to wisdom. Journal of applied systems analysis, 16(1), 3-9.

Aghina, W., De Smet, A., Lackey, G., Lurie, M., & Murarka, M. (2018). The five trademarks of agile organizations. McKinsey & Company Report.

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American psychologist*, *51*(4), 355.

Bataller, C. & Harris, J. (2016). Turning Artificial Intelligence into Business Value. Today. Accenture.

Bell, S. (2005). Lean enterprise systems: using IT for continuous improvement (Vol. 33). John Wiley & Sons.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., Tucker-Ray, W., Congdon, W., & Galing, S. (2017). Should governments invest more in nudging? Psychological science, 28(8), 1041-1055.

Bishop, C. M. (2006). Pattern recognition. Machine learning, 128(9).

CPLT (2018). El Derecho de Acceso a la Información Pública como Derecho Llave para el Acceso a otros derechos fundamentales. Consejo para la Transparencia.

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., & Ebel, P. (2021). The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems. Hawaii international conference on system sciences (HICSS).

Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. Business & Information Systems Engineering, 61(5), 637-643.

Denis, Hermosilla, Aracena, Sánchez, González, & Pombo (2021). Uso responsable de IA para política pública: Manual de formulación de proyectos. Banco Interamericano de Desarrollo.

Eggers, Mendelson, Chew & Kishnani (2019). Crafting an AI strategy for government leaders. Does your agency have a holistic AI strategy? Deloitte Insights.

EPA (2014). Next Generation Compliance: Strategic Plan 2014-2017. US Environmental Protection Agency.

Fischhoff, B. (2007). Nonpersuasive communication about matters of greatest urgency: Climate change: 7204-7208

Fischhoff, B., & Davis, A. L. (2014). Communicating scientific uncertainty. Proceedings of the National Academy of Sciences, 111(Supplement 4), 13664-13671.

Gobierno Digital (2019). Estrategia de transformación digital del Estado: Estado al servicio de las personas. Ministerio Secretaria General de la República (SEGPRES).

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. Cognitive Science, 27(4), 591-635.

González, F., Ortiz, T., & Sánchez, R. (2020). "IA Responsable". Manual técnico – Ciclo de vida de la inteligencia artificial". Banco Interamericano de Desarrollo.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

Hamming, R. (1962). *Numerical Methods for Scientists and Engineers*, McGraw-Hill.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning. New York: Springer

Hastie, T., Tibshirani, R., & Wainwright, M. (2019). Statistical learning with sparsity: the lasso and generalizations. Chapman and Hall/CRC.

Hinton, G. E. (1992). How neural networks learn from experience. Scientific American, 267(3), 144-151.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. science, 313(5786), 504-507.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

Hostmann, B. (2012). Best practices in analytics: Integrating analytical capabilities and process flows. Gartner.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review, 93*(5), 1449-75.

Kahneman, D., Slovic, P. & Tversky, A. (Eds.). (1982). Judgment under uncertainty: Heuristics and biases. Cambridge university press.

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). Strategy, not technology, drives digital transformation. MIT Sloan Management Review and Deloitte University Press, 14(1-25).

Kim, G., Humble, J., Debois, P., & Willis, J. (2016). The DevOps Handbook: How to Create World-Class Agility. Reliability, and Security in Technology Organizations.

Kohavi, R., & Longbotham, R. (2017). Online Controlled Experiments and A/B Testing. Encyclopedia of machine learning and data mining, 7(8), 922-929.

Krackhardt, D., Nohria, N., & Eccles, B. (2003). The strength of strong ties. Networks in the knowledge economy, 82.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. Artificial intelligence, 33(1), 1-64.

Langley, P., & Simon, H. A. (1981). The central role of learning in cognition. Cognitive skills and their acquisition, 361-380.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

Lepenioti, K., Bousdekis, A., Apostolou, D., & Mentzas, G. (2020). Prescriptive analytics: Literature review and research challenges. International Journal of Information Management, 50, 57-70.

Loewenstein, G., & Chater, N. (2017). Putting nudges in perspective. Behavioural Public Policy, 1(1), 26-53.

Maturana, H. R., & Varela, F. J. (1987). The tree of knowledge: The biological roots of human understanding. New Science Library/ Shambhala Publications.

MinCiencia (2021). Política Nacional de Inteligencia Artificial. Borrador, Ministerio de Ciencias, Tecnología, Conocimiento e Innovación, Santiago de Chile.

Mitchell, T. M. (1997) Machine Learning. McGraw Hill.

Murphy, K. P. (2012). Machine learning: a probabilistic perspective. MIT press.

Newell, A. (1990). Unified theories of cognition. Cambridge, MA: Harvard University Press.

NITI Aayog (2018). National strategy for artificial intelligence. Reporte, NITI Aayog, Government of India New Delhi, India.

NSTC (2019). The national artificial intelligence research and development strategic plan: 2019 update. National Science and Technology Council (US). Select Committee on Artificial Intelligence, Washington DC, USA.

OECD (2018). Digital Government Review of Brazil: Towards the Digital Transformation of the Public Sector

OECD (2019a). The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies, OECD Publishing, Paris.

OECD (2019b). Hello, World: Artificial intelligence and its use in the public sector. OECD Working Papers on Public Governance, No. 36.

OECD (2019c). Artificial Intelligence in Society, OECD Publishing, Paris.

OECD (2020). "The OECD Digital Government Policy Framework: Six dimensions of a Digital Government", OECD Public Governance Policy Papers, No. 02.

OECD. (2011). Quality framework and guidelines for OECD statistical activities.

Palmer, Shelly (2015). Data Science for the C-Suite. New York: Digital Living Press.

Peeters, Van Diggelen, Van den Bosch et al. (2021). Hybrid collective intelligence in a human–AI society. AI & Soc 36, 217–238.

Pidgeon & Fischhoff (2011). The role of social and decision sciences in communicating uncertain climate risks. Nature climate change, 1(1), 35-41.

Provost & Fawcett (2013). Data science for business: what you need to know about data mining and data-analytic thinking. O'Reilly.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Slavin, A., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, N., Luccioni, A., Maharaj, T., Sherwin, E., Mukkavilli6, K., Kording, k., Gomes, C., Ng, A., Hassabis, D., Platt, J., Creutzig, F., Chayes, J., & Bengio, Y. (2019). Tackling climate change with machine learning. arXiv preprint arXiv:1906.05433.

Russell, A. L. (2014). Open standards and the digital age. Cambridge University Press.

Schwaber & Sutherland (2020). The Scrum Guide: The Definitive Guide to Scrum.

Settles, B. (2009). Active Learning Literature Survey. Science, 10(3), 237-304.

Shalizi, C. R. (2016). Advanced Data Analysis from an Elementary Point of View. Cambridge, England, UK: Cambridge University Press.

Shearer, C. (2000). The CRISP-DM model: the new blueprint for data mining. Journal of data warehousing, 5(4), 13-22.

Simon, H. A. (1991a). Bounded rationality and organizational learning. Organization science, 2(1), 125-134.

Simon, H. A. (1991b). Artificial intelligence: where has it been, and where is it going? IEEE Transactions on Knowledge and Data Engineering, 3(2), 128-136.

Simon, H. A. (1997). Models of bounded rationality: Empirically grounded economic reason (Vol. 3). MIT press.

Smeets & Weterings (1999). Environmental indicators: Typology and overview.

Smeets, E., & Weterings, R. (1999). Environmental indicators: Typology and overview. European Environment Agency, Technical report No 25.

Sra, S., Nowozin, S., & Wright, S. (2012). Optimization for machine learning. MIT Press.

Sunstein, C. R. (2014). Why nudge?: The politics of libertarian paternalism. Yale University Press.

Thaler, R. H. (2018). From cashews to nudges: The evolution of behavioral economics. *American Economic Review, 108*(6), 1265-87.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58(1), 267-288.

Turing, A. M. (1950). Computing machinery and intelligence. Mind, 59(236), 433-460.

UNESCO (2019). Artificial intelligence for sustainable development: challenges and opportunities for UNESCO's science and engineering programmes.

Van der Aalst, W. M. P. (2011). Process Mining: Data Science in action. Springer.

Wasserman, L. (2004). All of statistics: a concise course in statistical inference (Vol. 26). New York: Springer.

Wong-Parodi, G., Krishnamurti, T., Davis, A., Schwartz, D., & Fischhoff, B. (2016). A decision science approach for integrating social science in climate and energy solutions. Nature Climate Change, 6(6), 563-569.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686-688.

SMA

Superintendencia del Medio Ambiente
Gobierno de Chile

**Teatinos 280, pisos 7, 8 y 9, Santiago de Chile**

Fono: +56 2 2617 1800

Oficina de partes: Teatinos 280, piso 8.

Horario de atención: Lunes a viernes de

9:00 a 13:00 horas, piso 9.

**Sitio web: portal.sma.gob.cl**