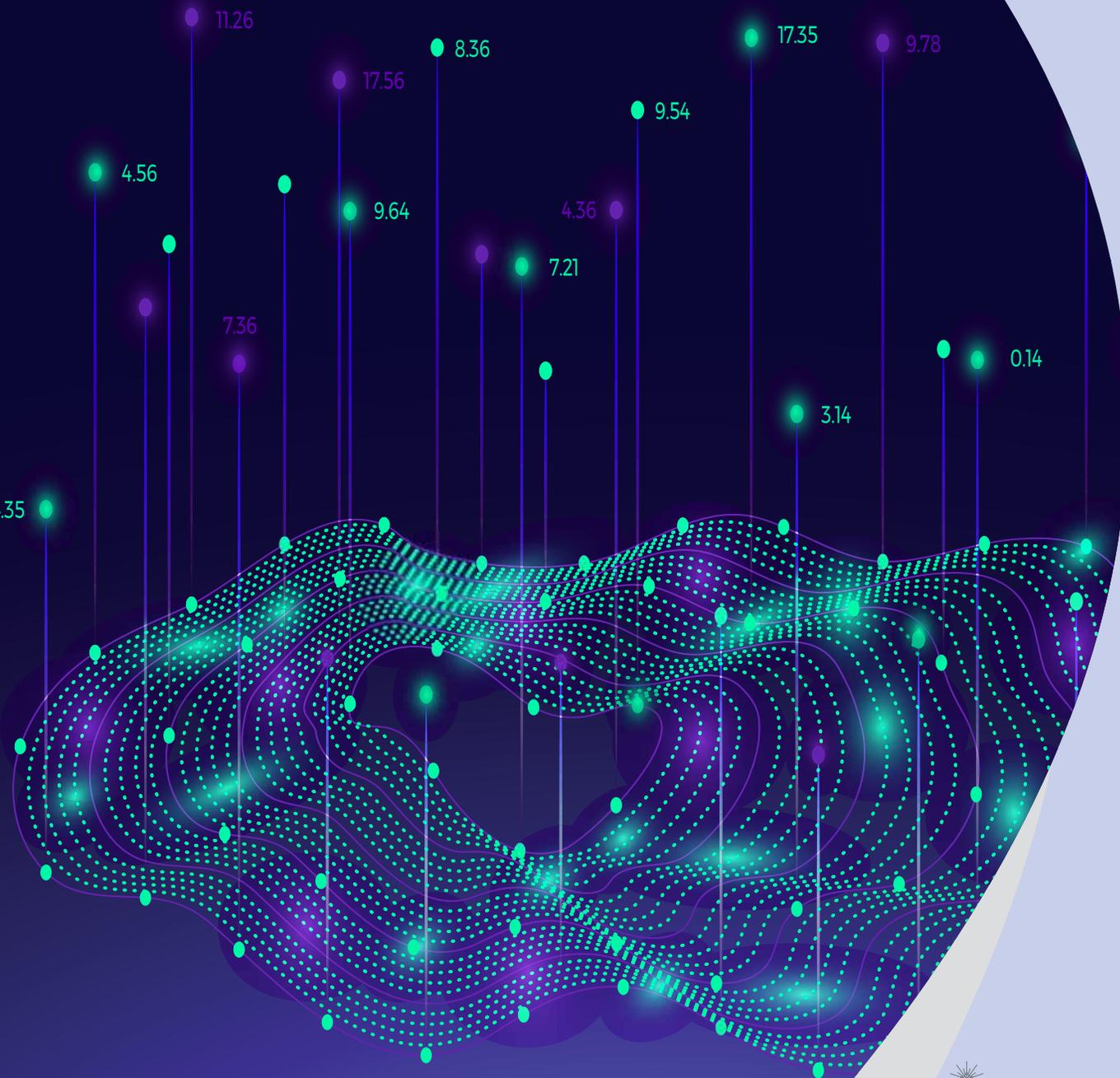


# BIG DATA EN MINERÍA

Editor: Javier Ruiz del Solar





## **BIG DATA EN MINERÍA**

**ISBN: 978-956-19-1186-4**

**Editor General:**  
Javier Ruiz del Solar

---

**Director Ejecutivo de Beauchef Minería:**  
Willy Kracht

---

**Comité Editorial:**  
Patricio Meller  
Gabriela Novoa

---

**Colaboradores:**  
Álvaro Egaña  
Alejandro Ehrenfeld  
Humberto Estay  
Bárbara Salinas  
Richard Weber

---

[contacto@beauchefmineria.cl](mailto:contacto@beauchefmineria.cl)

<http://www.beauchefmineria.cl/>

Av. Beauchef 850, Santiago, Chile

**Producción gráfica:**  
<http://www.publisiga.cl/>

**Impresión:**  
Gráfica Andes

**Septiembre de 2020**

**BIG  
DATA**

# Contenido



<b>CARTA DE PRESENTACIÓN</b>	<b>4</b>
<b>PRÓLOGO</b>	<b>6</b>
<b><i>Big Data en minería</i></b>	<b>8</b>
1. Digitalización, big data y minería	8
2. Modelo conceptual y desafíos	13
3. Captura y transmisión de datos	14
4. Uso de datos e información obtenida	16
5. Conocimiento generado	17
6. Capital humano	17
7. Ciberseguridad	18
8. Desafío organizacional	20
Referencias	21
<b><i>Big Data, principales características y desafíos</i></b>	<b>22</b>
1. Elementos metodológicos y tecnológicos	27
2. Definición del problema	28
3. Organización y comprensión de los datos	29
4. Procesamiento de los datos	30
5. Modelamiento y análisis de los datos	31
6. Agradecimientos	33
Referencias	34
<b><i>Big Data en exploración minera</i></b>	<b>36</b>
1. Introducción	37
2. ¿Qué se hace hoy?	40
3. Desafíos y oportunidades	43
3.1 Incorporación de nuevas fuentes de información	45
3.2 Tomografía sísmica y sus potenciales alcances	49
4. Visión y comentarios finales	52
Referencias	55
<b><i>Big Data en operaciones mineras</i></b>	<b>56</b>
1. Introducción	57
2. ¿Qué se hace hoy?	61
3. Desafíos y oportunidades	67
4. Visión y comentarios finales	71
5. Agradecimientos	71
Referencias	72
<b><i>Big Data en el procesamiento de minerales y metalurgia extractiva</i></b>	<b>74</b>
1. ¿Qué se hace hoy?	75
1. Una primera etapa de procedimientos simplificados que busquen limpiar los datos removiendo valores implausibles (data conditioning)	81
2. Trazabilidad espacio-temporal y consistencia entre los resultados de diferentes mediciones	81
3. Análisis estadístico de los datos	82
2. ¿Qué está cambiando?	87
3. Visión: Ideas y propuestas	94
Referencias	98
<b>Casos aplicados: <i>Big Data</i> en geometalurgia</b>	<b>102</b>
1. Introducción	103
2. Casos Aplicados	106
3. Visión: Ideas y propuestas	121
4. Conclusiones	124
Referencias	126

# Carta de Presentación

Me es grato presentar el informe titulado "*Big Data* en Minería", el cual ha sido elaborado gracias al esfuerzo y compromiso de académicas, académicos y profesionales de la Universidad de Chile y del Advanced Mining Technology Center (AMTC), en su constante labor en el descubrimiento de nuevos conocimientos y tecnologías puestas al servicio de la comunidad y de la sociedad en su conjunto.

Vivimos en un contexto en el cual la transformación digital ha impactado múltiples dimensiones que comprenden no solo el espacio académico, sino también a las esferas de la toma de decisiones tanto en lo público como en lo privado. En ese contexto, la Universidad de Chile ha sido pionera en estos procesos, y goza de un rol destacado en el desarrollo del *big data* y en la incorporación de nuevas metodologías en el campo, asumiendo una responsabilidad formativa en lo que respecta a la docencia e investigación a la hora de formar profesionales que se insertan en la investigación aplicada multidisciplinar.

Un rol clave le compete en este camino al papel que a Chile le toca jugar como país líder en minería y a los potenciales desarrollos asociados a la misma. Por ello, el estudio "*Big Data* en Minería" busca expandir el conocimiento en el área a partir de los análisis que se presentan, en miras de superar desafíos y contribuir al desarrollo tecnológico en



términos de productividad y competitividad del sector minero. A su vez, busca fomentar escenarios dinámicos en esta industria, en la cual la innovación y el análisis interdisciplinar confluyan en un espacio profesional y de acción colectiva a la hora de incorporar nuevas tecnologías que transforman de manera virtuosa a estas organizaciones.

Sin perjuicio de la especificidad en torno al sector minero, el trabajo que nos reúne debe servir el propósito de incorporarse como un elemento útil y transversal en las discusiones y desafíos que el país enfrenta en una multiplicidad de dimensiones, en particular en lo que refiere al ámbito público, la implementación de políticas, y en especial, a las y los tomadores de decisiones en ese campo.

Asimismo, el carácter gratuito y universal en la distribución de este estudio, nos recuerda la misión de las universidades y centros de inves-

tigación en su compromiso con la sociedad en poner a disposición el conocimiento y nuevas herramientas, disminuyendo las barreras de acceso y posibilitando una red integrada entre academia y sociedad civil y productiva.

Finalmente, no me cabe más que relevar y agradecer el constante compromiso de académicas/os, investigadores y profesionales que con arduo trabajo y dedicación hicieron este informe posible.

**Richard Weber**

Académico y actualmente  
Director del Departamento de  
Ingeniería Industrial de la FCFM de  
la Universidad de Chile.



## Estudio del *Big Data*

**Javier Ruiz del Solar**

AMTC y Departamento de Ingeniería Eléctrica,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile



6

# Prólogo

La actual tendencia hacia la digitalización de dispositivos y procesos genera día a día una creciente cantidad de datos (fotografías, videos, emails, mensajes de texto, publicaciones digitales, transacciones comerciales, etc.), los cuales influyen no solo en la vida diaria de las personas, sino también en los sectores productivos de las distintas economías del mundo.

Los datos generados han alcanzado un volumen acumulado tan grande, que se ha debido buscar métodos y tecnologías específicas para gestionarlos. Es en este escenario que nace el término

no *big data*, el cual se utiliza para denominar la existencia y uso de estos grandes volúmenes de datos, así como también para designar métodos y tecnologías para transmitir, almacenar, procesar, analizar, visualizar y administrar estos datos. Complementariamente, el término *big data analytics* se refiere al análisis de datos que califican en la categoría de *big data*, utilizando generalmente técnicas de ciencia de los datos e inteligencia artificial.

En los últimos años, todas las industrias se han visto enfrentadas al desafío de la transformación digital y a ser capaces de obtener el máximo beneficio que esta pueda generar. Naturalmente, el valor no se encuentra en los datos mismos, sino que en la información contenida en estos y en el conocimiento que puede ser generado a partir de su análisis. Por lo tanto, es

fundamental contar con las metodologías y tecnologías para extraer automáticamente información y generar conocimiento a partir de estos grandes volúmenes de datos.

La industria minera es uno de los sectores productivos que se ha visto inmerso en este nuevo desafío y ya estamos siendo testigos de cómo está enfrentando estos retos. En algunos casos, lo ha hecho utilizando metodologías y tecnologías desarrolladas en otras industrias. Sin embargo, se debe tener en cuenta que cada sector productivo posee sus particularidades y necesidades, lo que significa que la aplicación del *big data* y de sus tecnologías asociadas va a depender de las características de cada industria. A modo de ejemplo, el sector de servicios, como el retail o la banca, no poseen los mismos requerimientos que la salud y, a la vez, esta tampoco tiene los requerimientos de las industrias de recursos naturales como la minería. Por lo mismo, seleccionar los datos a ser capturados, transmitidos, analizados y gestionados, así como las metodologías y tecnologías a ser empleadas para extraer la información y el conocimiento relevantes, dependerá de las particularidades de cada industria.

En este contexto, el presente estudio aborda el desafío del uso del *big data* en minería, considerando sus particularidades, como también la de sus distintos procesos productivos. Para aquellos no familiarizados con el quehacer del manejo de datos y el *big data*, se presenta una perspectiva histórica de su desarrollo y el de sus disciplinas relacionadas. Asimismo, se definen términos de común uso, se identifican desafíos generales y se presentan herramientas computacionales para el manejo de datos. Luego, se conceptualiza un modelo para el uso del *big data* en minería y se presentan ideas y desafíos para la aplicación actual y futura de estas tecnologías en distintos procesos unitarios del sector minero.

En los últimos años,  
todas las industrias se  
han visto enfrentadas  
al desafío de la  
transformación digital y  
a ser capaces de obtener  
el máximo beneficio que  
esta pueda generar.

Posteriormente, se expone la visión de distintos especialistas que analizan la utilización del *big data* en diversas etapas del negocio minero: exploración, planificación, operación, procesamiento de minerales y metalurgia extractiva. La visión de los especialistas puede ser de suma importancia para los expertos en ciencia de los datos, que no necesariamente conocen las particularidades de los distintos procesos mineros.

Considerando la multiplicidad de usos del *big data* y sus tecnologías asociadas, así como el desarrollo acelerado que estas tienen, el presente estudio no busca realizar un análisis exhaustivo de las tecnologías y sus usos, sino más bien mostrar tendencias y presentar conceptos que esperamos sean de utilidad, tanto para los futuros estudios que se realicen respecto de esta materia, como también para la implementación de estas tecnologías en la industria minera.



# *Big Data* en minería

## **Javier Ruiz del Solar**

AMTC y Departamento de Ingeniería Eléctrica,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

### **1. Digitalización, *big data* y minería**

Hoy en día todas las industrias se ven enfrentadas al desafío de la transformación digital y a ser capaces de capturar el máximo beneficio de los grandes volúmenes de datos que esta puede generar. En todo el mundo la digitalización avanza a una velocidad vertiginosa y se está produciendo una separación importante entre aquellas industrias y países que están preparados para capturar los beneficios de la digitalización y aquellos que no (McKinsey, 2015).



En todo el mundo la digitalización avanza a una velocidad vertiginosa y se está produciendo una separación importante entre aquellas industrias y países que están preparados para capturar los beneficios de la digitalización y aquellos que no (McKinsey, 2015).

La digitalización no se refiere solo a la compra de equipos y sistemas, sino que a la forma en que los datos digitales generados se almacenan, gestionan y utilizan. El valor no está en los datos mismos, sino que en la información contenida en estos y en el conocimiento que puede ser generado a partir de su análisis. Por lo tanto, es fundamental contar con las tecnologías apropiadas para extraer automáticamente información y generar conocimiento a partir de estos grandes volúmenes de datos. Diversos estudios señalan que la digitalización y el uso apropiado de los datos generados puede tener importantes efectos en el aumento de la productividad y en la forma en que muchas industrias se organizan (McKinsey, 2015).

En el caso de la minería, se visualiza que la digitalización contribuirá a abordar los desafíos específicos que la industria

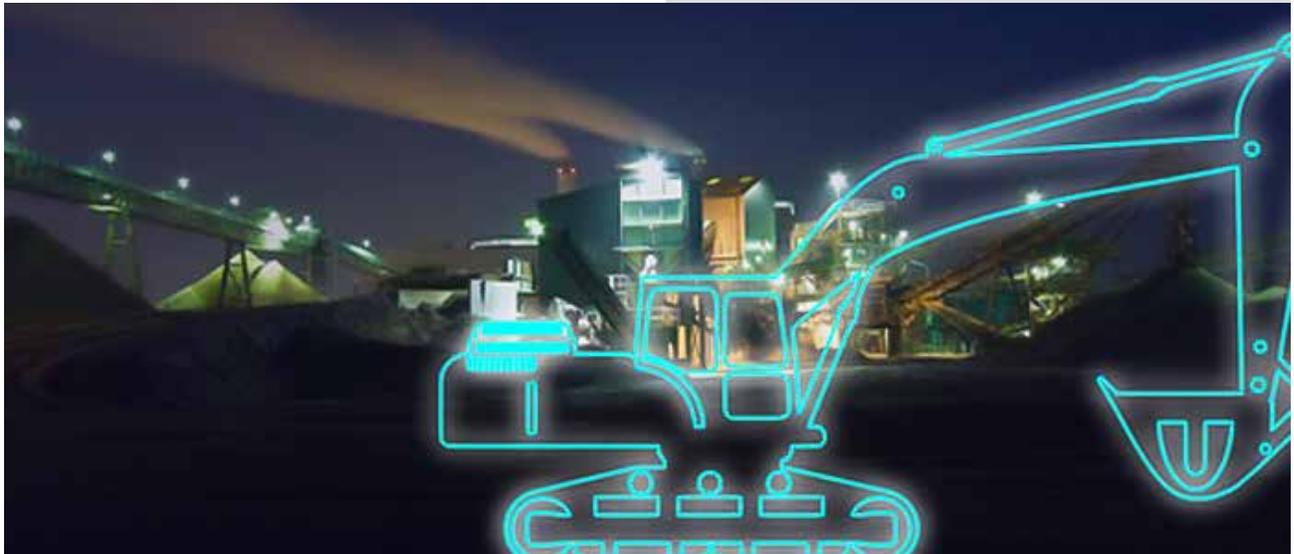
enfrenta a nivel global (Durrant-Whyte et al., 2015; World Economic Forum, 2017), tales como yacimientos antiguos con leyes de mineral decrecientes; necesidad de disminuir la huella de carbono y el impacto ambiental de la actividad; necesidad de tener mayor trazabilidad de los procesos y de los productos generados, y necesidad de aumentar la seguridad de los trabajadores.

En la minería, al igual que en otros sectores productivos de recursos naturales, se observa un rezago en la adopción de la digitalización y sus tecnologías asociadas como *big data*, respecto a otras industrias como el retail o la industria financiera. De acuerdo a McKinsey (2015), el año 2015 en Estados Unidos el índice de digitalización de la industria minera era uno de los más bajos, comparable al de la industria de la construcción y al de la agricultura. Sin embargo, en los últimos años

se observa una preocupación de la industria minera a nivel global, en particular de sus grandes corporaciones, por disminuir esta brecha y acelerar el proceso de digitalización. Sin embargo, dentro de la industria existe una brecha interna importante entre los esfuerzos de las grandes mineras, respecto a las de tamaño medio y pequeño.

Estudios recientes han cuantificado a nivel global el beneficio económico potencial de la digitalización en la industria minera. Por ejemplo, de acuerdo también a McKinsey (Durrant-Whyte et al., 2015), el beneficio de una mayor digitalización en términos de un mejor

Los beneficios de la digitalización se pueden obtener solamente si los datos generados son apropiadamente almacenados, gestionados y utilizados.



manejo de las operaciones; mantenimiento de equipos; salud y seguridad laboral; diseño y compra de equipos, y productividad de las personas, se calcula al año 2025 en 370 mil millones de dólares a nivel global. Por otra parte, el World Economic Forum (2017) es aún más optimista en sus predicciones y en su estudio del 2017 indica que una mayor digitalización traerá los siguientes beneficios en los próximos 10 años: 425 mil millones de



dólares en utilidades para la industria minera, sus clientes, la sociedad y el medio ambiente; una reducción de 610 millones de toneladas de emisiones de CO<sub>2</sub>, y una disminución en 1000 del número de fallecidos en accidentes fatales en el periodo.

Se debe destacar que los beneficios de la digitalización se pueden obtener solamente si los datos generados son apropiadamente almacenados, gestionados y utilizados. Para esto se requiere el uso del *big data* y sus tecnologías asociadas de tal forma de extraer la información contenida en los datos y generar conocimiento a partir de esta, por ejemplo modelos predictivos para una adecuada toma de decisiones.

Al igual que en otras industrias, el uso del *big data* en minería posee beneficios importantes en las áreas de gestión, logística y comercialización de productos. En el ámbito operacional, los beneficios del uso del *big data* han sido relevados en diversos estudios (por ejemplo Durrant-Whyte et al.

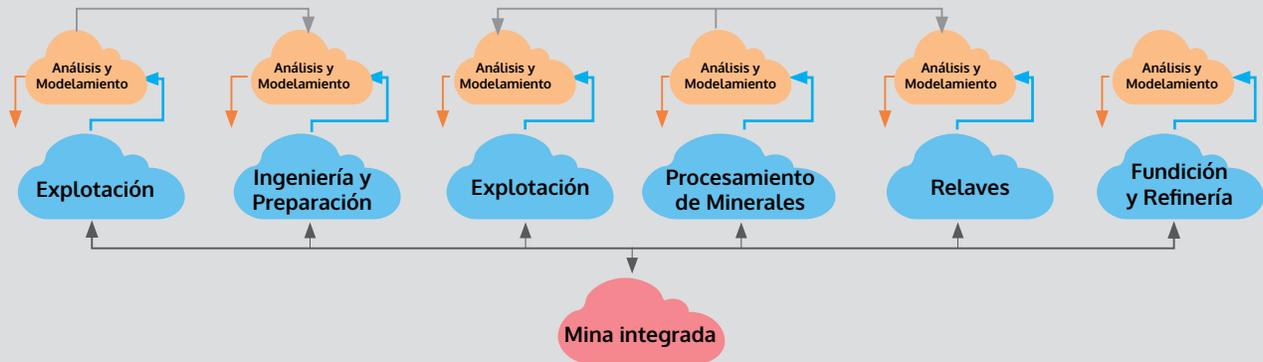
(2015) y World Economic Forum (2017)). A continuación se destacan algunos de los principales beneficios operacionales, los cuales serán abordados en más detalle en los capítulos siguientes del presente estudio:

- Mejor entendimiento de los recursos geológicos mediante la integración y el modelamiento de datos obtenidos desde diversos tipos de sensores, correspondientes a distintas fuentes de información. A modo de ejemplo, el uso de nuevos métodos geofísicos como la tomografía sísmica permite obtener proxies para la localización de yacimientos no superficiales.
- Monitoreo y toma de decisiones en tiempo real mediante el análisis de los datos provenientes de sensores que monitorean continuamente equipos, procesos e incluso el comportamiento de las personas, a través de dispositivos como teléfonos inteligentes o tags RFID. Este monitoreo permite mejorar la seguridad de las operaciones, observar desviaciones respecto a lo planificado y verificar el grado de utilización de los equipos.
- Mantenimiento más eficiente de los equipos a través del análisis de sus datos sensoriales internos ("signos vitales"), lo cual permite transitar desde una política de mantenimiento preventivo (basada en estadísticas de fallas), a una de mantenimiento predictivo basada en el estado actual de los distintos componentes.



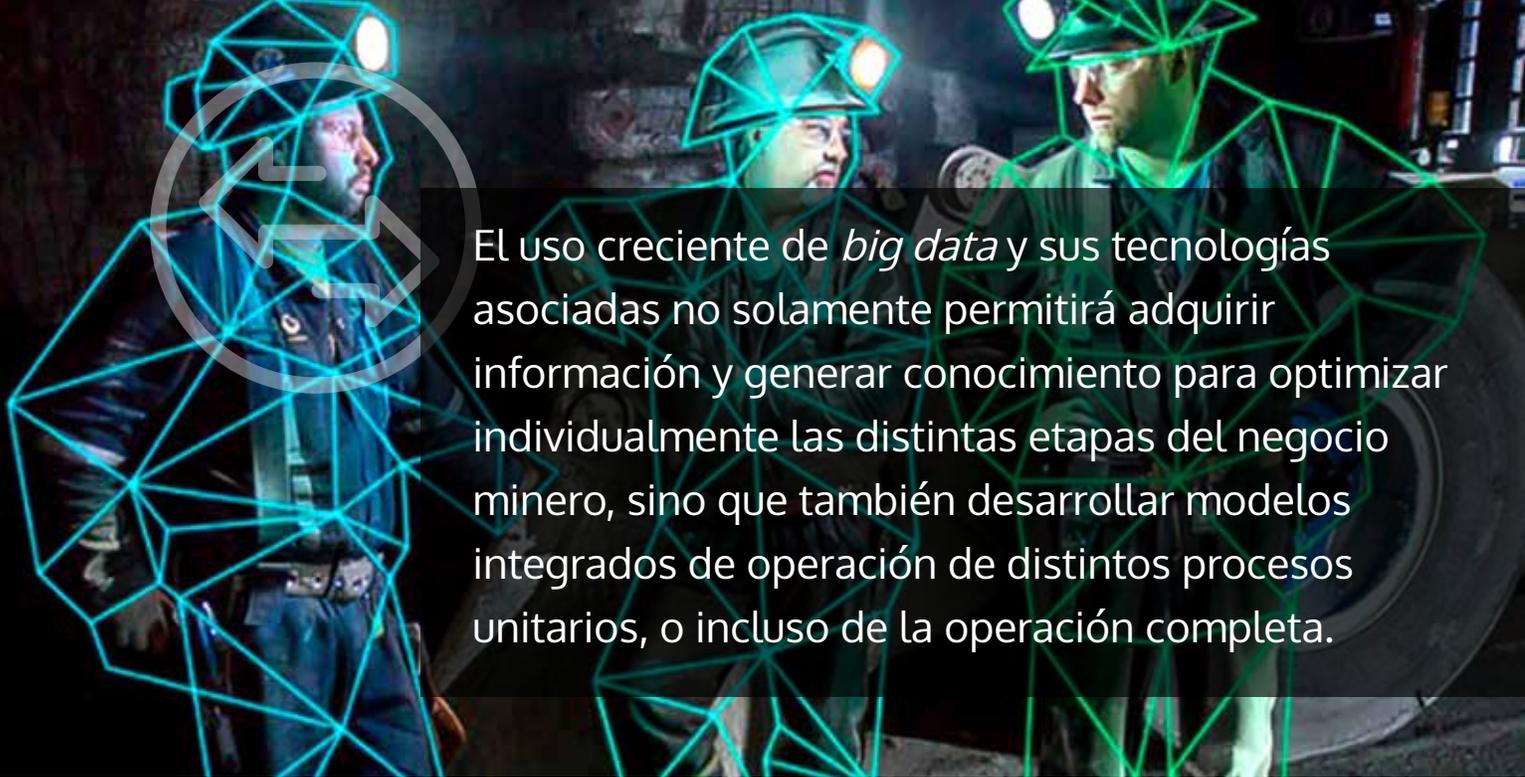
**Figura 1.**

Ejemplo de modelo integrado de aplicación del *big data* en el negocio minero.



- Monitoreo en tiempo real del estado de salud de equipos clave como molinos, chancadores y correas transportadoras, de tal forma de minimizar la ocurrencia de fallas y de planificar apropiadamente su mantenimiento.
- Modelamiento predictivo de procesos, lo cual permite predecir el comportamiento de estos y la variabilidad de su salidas (por ejemplo del flujo de material), pasando desde una política de monitoreo de valores medios a una que además incluya desviaciones respecto a dichos valores.
- Optimización del flujo de material y equipos en tiempo real, de tal forma de optimizar el rendimiento de los procesos, mediante el análisis en tiempo real de las variables operacionales, y de equipos, a través del análisis en tiempo real de su estado (por ejemplo, posición y velocidad en el caso de equipos móviles).
- Optimización del uso de insumos importantes como energía y agua mediante la sensorización de los procesos en los cuales estos se utilizan, de tal forma de gestionar su uso eficiente. Por ejemplo, de acuerdo a McKinsey (2017), el uso combinado de sensores, sistemas de control y modelamiento de datos puede reducir entre 10% y 30% el consumo de energía en procesos industriales.

Más aún, el uso creciente de *big data* y sus tecnologías asociadas no solamente permitirá adquirir información y generar conocimiento para optimizar individualmente las distintas etapas del negocio minero, sino que también desarrollar modelos integrados de operación de distintos procesos unitarios, o incluso de la operación completa. Los modelos integrados de operación han sido propuestos desde hace unos 20 años bajo distintos paradigmas, dentro de los que destacan el mine-to mill (McKee, 2013) y la geo-minero-metalurgia (Departamento de Ingeniería de Minas de la



El uso creciente de *big data* y sus tecnologías asociadas no solamente permitirá adquirir información y generar conocimiento para optimizar individualmente las distintas etapas del negocio minero, sino que también desarrollará modelos integrados de operación de distintos procesos unitarios, o incluso de la operación completa.

Universidad de Chile, 2019). Se espera que la digitalización de los distintos procesos unitarios y el uso del *big data* y sus tecnologías asociadas faciliten la implementación de estos paradigmas. De esta forma, en los próximos años veremos una creciente digitalización, modelamiento y toma de decisiones integrada de distintos procesos unitarios. La figura 1 esquematiza esta idea, aunque naturalmente los niveles de integración entre los distintos procesos pueden ser implementados de maneras muy diversas.

Actualmente los Centros Integrados de Operaciones (CIO) pueden desempeñar un rol fundamental en este aspecto. A modo de ejemplo, CODELCO indica en su sitio web institucional que sus salas CIO permitirán registrar y analizar “información fidedigna y on line de los procesos mineros, logrando de esta forma una visión en tiempo real de la operación en su conjunto, lo cual posibilita una toma de decisiones de gestión en forma integrada y con una oportunidad nunca antes

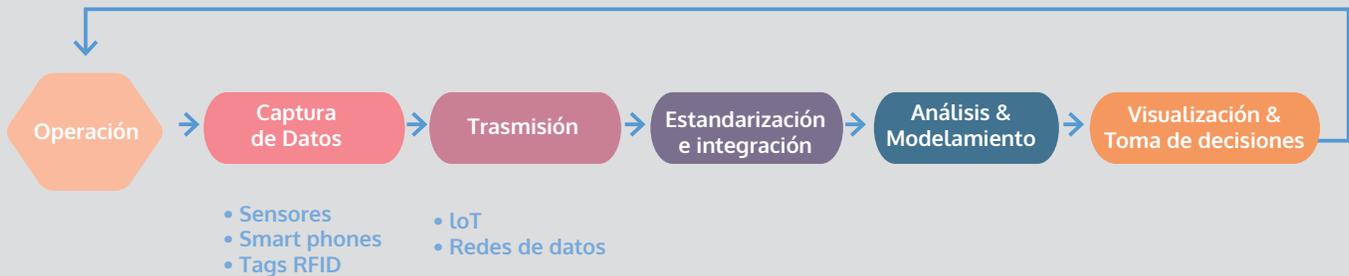
lograda”. CODELCO espera asimismo que este modelo de gestión integrada permita “capturar grandes volúmenes de valiosos datos (*big data*) los cuales facilitan una planificación colaborativa, focalizada en aspectos críticos para el negocio”. Actualmente CODELCO opera salas CIO en varias de sus divisiones (El Teniente, Andina, DMH).

## 2. Modelo conceptual y desafíos

En términos generales, la aplicación del *big data* y sus tecnologías asociadas posee importantes desafíos de implementación, como la necesidad de un apropiado modelamiento y diseño de los proyectos de transformación digital y *big data*, los cuales deben identificar claramente cuál es el beneficio a ser obtenido y las metodologías a ser utilizadas; la escalabilidad de las soluciones a ser implementadas; la disponibilidad del capital humano requerido, y la privacidad de los datos. Estos desafíos son transversales a todas las industrias.

**Figura 2.**

Modelo conceptual de la aplicación del *big data* en procesos productivos de la minería.



En el caso de la industria minera se agregan desafíos específicos, los cuales se relacionan con las características del negocio minero: operaciones localizadas en lugares geográficamente aislados (desierto, alta montaña, etc.), faenas de difícil acceso (minas profundas, etc.) y entornos adversos para la tecnología digital (presencia de polvo y humedad, materiales abrasivos, entre otros).

En este contexto cobra especial relevancia la captura, almacenamiento y transmisión de los datos a ser analizados. En la figura 2 se presenta un modelo conceptual de la aplicación del *big data* en minería, donde se identifican las distintas etapas del *pipeline* de procesamiento típico.

A continuación se describen algunos de los desafíos que existen para la aplicación de *big data* y sus tecnologías asociadas en la operación minera.

### 3. Captura y transmisión de datos

Tal como fue mencionado anteriormente, el ambiente minero es, en términos generales hostil para los dispositivos electrónicos y la obtención de datos mediante sensorización es muchas veces compleja. Por ejemplo, en minería subterránea el uso de sensores dentro de la mina para observar condiciones ambientales tales como humedad o estabilidad geotécnica, se enfrenta al problema de que los sensores a ser instalados no sólo deben ser resistentes al polvo y a la humedad, sino que también a los distintos procesos unitarios, como la tronadura, la cual puede dañar estos dispositivos. En el caso de los procesos metalúrgicos, muchas variables no pueden ser sensadas directamente, pues los procesos químicos involucrados pueden dañar a los sensores. Por lo mismo, un área importante de investigación y desarrollo actual es el desarrollo de sensores para



aplicaciones mineras que sean robustos frente a las condiciones de operación en las cuales serán utilizados.

Por otra parte, si los datos capturados deben ser utilizados para tomar decisiones en tiempo real, es necesario poder transmitirlos en forma efectiva. Para lograr esto, un desafío es el limitado ancho de banda de las redes de comunicación utilizadas, tanto en minería subterránea como en minería a cielo abierto. Estas redes muchas veces funcionan a plena capacidad por los requerimientos de transmisión de señales de video y de comandos de control requeridos por equipos teleoperados o autónomos operando bajo supervisión. Estos desafíos pueden ser abordados mediante el uso de redes de datos de mayor capacidad (por ejemplo 5G) y de protocolos de fácil comunicación entre los sensores y las redes de datos (IoT industrial), pero también mediante el uso del paradigma

Se hace necesario tener sistemas interoperables basados en formatos de datos estándar.

Diversos estudios señalan que una parte importante de los datos que se generan en los distintos procesos mineros no es utilizada en la toma de decisiones y que incluso ni siquiera es almacenada.



*edge computing*, el cual busca realizar la mayor parte del procesamiento en el mismo sensor, de tal forma de minimizar la cantidad de datos a ser transmitidos, mediante la extracción in situ de la información relevante.

Un tercer desafío se refiere a la propiedad de los datos. Esta situación ocurre en el caso de vehículos y grandes equipos mineros, los cuales tienen muchos sensores internos y por lo mismo generan una importante cantidad de datos (signos vitales). Sin embargo, muchos de los contratos de compra de estos equipos establecen que los datos que el equipo genere al entrar en operación, no pueden ser utilizados por la empresa dueña de los equipos para su análisis y para toma de decisiones (por ejemplo en aplicaciones de mantenimiento predictivo), pues pertenecen al fabricante. Naturalmente este desafío no es de carácter técnico, sino contractual.

Un último desafío se refiere al formato de los datos y su compatibilidad. Se hace necesario

tener sistemas interoperables basados en formatos de datos estándar. Existen iniciativas como el “Programa tecnológico para la creación y adopción de estándares internacionales para interoperabilidad minera” de la Corporación Alta Ley (n.d.) y el *Global Mining Guidelines Group* (GMG) (2019) que se encuentran abordando esta problemática.

#### 4. Uso de datos e información obtenida

Diversos estudios señalan que una parte importante de los datos que se generan en los distintos procesos mineros no es utilizada en la toma de decisiones y que incluso ni siquiera es almacenada. A modo de ejemplo, en Durrant-Whyte et al. (2015) se señala que menos de un 1% de los datos que se generan en los procesos mineros serían utilizados en procesos de toma de decisión. Naturalmente, considerando la gran variedad de operaciones mineras que existen en el mundo y sus distintos procesos unitarios, es muy difícil hacer una estimación del porcentaje de datos



efectivamente utilizados. Lo que sí podemos afirmar con certeza es que la fracción de datos generados en operaciones mineras que son utilizados en la toma de decisión es baja y que esta situación debe ser revertida si se quiere aprovechar plenamente los beneficios de la digitalización.

Por otra parte, usualmente se confunde el concepto de dato con el de información. Se olvida que los datos pueden ser redundantes, estar correlacionados o simplemente no aportar información. La información corresponde a “datos que han sido procesados de tal forma de ser útiles” (Bellinger et al., 2003). La Teoría de la Información provee metodologías formales que permiten cuantificar la cantidad de información contenida en los datos. De hecho, utilizando estas metodologías los datos pueden ser comprimidos sin perder información. Un desafío importante es utilizar algoritmos que permitan extraer información relevante de los datos disponibles en distintas etapas

del *pipeline* de procesamiento, de tal forma de almacenar, transmitir y analizar solo los datos que contengan información relevante.

## 5. Conocimiento generado

El conocimiento se genera a partir de datos y de información, y permite entender patrones de los datos y generar modelos. En muchas aplicaciones la generación de modelos predictivos permite una toma de decisión automatizada o asistir a procesos de toma de decisión de humanos. Sin embargo, la generación de modelos que sean útiles requiere un entendimiento fenomenológico de los procesos que se están modelando, por lo que el desafío para la generación de estos modelos es la efectiva colaboración entre los especialistas en *big data*, o en ciencia de los datos e inteligencia artificial, con los profesionales que entienden los distintos procesos a ser modelados (ingenieros metalúrgicos, ingenieros de minas, ingenieros hidráulicos, geólogos, geotécnicos, etc.).

## 6. Capital humano

El uso del *big data* y sus tecnologías asociadas requiere contar con técnicos y profesionales que entiendan estas nuevas tecnologías y que puedan generar valor a partir del análisis y procesamiento de los datos. En el caso del personal técnico a cargo de las tareas operativas los desafíos son diversos. Por una parte se requerirá que estos puedan operar apropiadamente dispositivos y maquinaria con nuevas interfaces y nuevas filosofías de funcionamiento; lo cual va desde máquinas que se usan al interior de las operaciones,



hasta equipos que deben ser teleoperados remotamente y equipos autónomos cuya operación debe ser supervisada, también de forma remota. Por otra parte se requerirá nuevo personal técnico, experto en las nuevas tecnologías digitales, que pueda instalar y reparar sensores, redes de datos, servidores y componentes digitales de los equipos mineros.

A nivel profesional se requerirán nuevos perfiles de profesionales, expertos en ciencia de los datos, que sean capaces de implementar los proyectos de transformación digital y *big data* al interior de la empresa, definiendo requerimientos, implementando

los sistemas de gestión de los datos, modelando los datos, e incorporando datos, información y predicciones en los sistemas de toma de decisión.

Para una industria tradicional como la minería el capital humano es un importante desafío a abordar. Por una parte el personal actual en operación debe ser capacitado y por otra parte profesionales con nuevos perfiles deben ser contratados y entrenados.

## 7. Ciberseguridad

Todas las organizaciones deben preocuparse hoy en día de la ciberseguridad, es decir, de tener procedimientos efectivos para defender sus computadores, servidores, redes de datos y en general todos sus sistemas electrónicos, de ciberataques. En una industria productiva como la minera, esta preocupación se extiende a la seguridad de los equipos operacionales.

Todas las organizaciones deben preocuparse hoy en día de la ciberseguridad, es decir, de tener procedimientos efectivos para defender sus computadores, servidores, redes de datos y en general todos sus sistemas electrónicos, de ciberataques. En una industria productiva como la minera, esta preocupación se extiende a la seguridad de los equipos operacionales.

Las políticas de ciberseguridad incluyen definición y uso de procesos, herramientas computacionales de protección, análisis de riesgo, entrenamiento de personas, adopción de buenas prácticas de otras industrias y definición de procedimientos alternativos de operación frente a contingencias. Por ejemplo, cómo operar equipos autónomos y teleoperados frente a un ataque que deje inutilizada la red de datos o los servidores.

Por otra parte, debido al aumento del uso de las tecnologías de la información y sobre todo, debido al sostenido crecimiento del uso de la automatización de procesos, del incremento en el uso de sistemas autónomos y del incremento de las interconexiones de equipos, se hace necesario que las organizaciones tengan un área dedicada a los temas de seguridad informática, cuyo responsable reporte directamente al máximo responsable de la organización.

Un aspecto importante a ser abordado es la cultura organizacional en temas de ciberseguridad. La mayoría de las empresas no capacita a todos sus trabajadores en temas de seguridad informática, por lo cual no genera una cultura organizacional en este ámbito. De hecho, muchas organizaciones han sido víctimas de ciberataques porque algún trabajador se descuidó, dejando sin querer una "puerta informática" abierta por la cual se produjo el ataque.

Finalmente, es importante considerar que existen estándares para construir software y hardware seguro, también hay estándares para implementar redes de datos seguras, y por sobre todo, es muy importante utilizar las herramientas de automatización de procesos que existen en el mercado, porque estas serán de mucha ayuda para prevenir ataques, para descubrir intrusos y para recuperarse de ataques.

## 8. Desafío organizacional

La digitalización de los procesos y el uso del *big data* es una transformación mayor de la industria minera. Por lo mismo existe un desafío organizacional importante a abordar considerando que las barreras para la adopción digital son físicas, financieras, legales y culturales (McKinsey, 2017). Existen importantes beneficios, los cuales solo podrán ser capturados si las empresas adecuan sus estructuras organizacionales y gestionan apropiadamente su transformación digital. De acuerdo al World Economic Forum (2017) las empresas pueden ser clasificadas en tres categorías según su madurez digital: líderes (*first movers*), seguidores rápidos (*fast followers*) y conservadores (*business as usual*). Las empresas en las primeras dos categorías son las que tendrán las mejores posibilidades de capturar los beneficios de la transformación digital.

En el mismo estudio se presentan recomendaciones a seguir para implementar un proceso exitoso de transformación digital.

Una de las principales es tener una estrategia focalizada que incorpore la transformación digital y la alinee con el modelo de negocios, los procesos y la organización.

La transformación digital requiere aplicar, adaptar o desarrollar nuevas tecnologías y metodologías de análisis y procesamiento de datos, tareas en las cuales los proveedores de la minería pueden jugar un importante rol. A modo de ejemplo, en CSIRO (2017) se presenta una hoja de ruta para los *METS (Mining Equipment, Technology and Services)* australianos, en la cual se identifican las áreas de toma de decisiones basadas en el análisis de datos (*data driven mining decisions*), y de robótica y automatización, como dos de las áreas principales de desarrollo futuro para estas empresas.





# Referencias

Alta Ley. (2017). Programa Tecnológico para la creación y adopción de estándares internacionales para interoperabilidad minera. Recuperado de: <https://corporacionaltaley.cl/proyectos/programa-tecnologico-para-la-creacion-y-adopcion-de-estandares-internacionales-para-interoperabilidad-minera/>

Bellinger, G., Castro, D., & Mills, A. (2003). Data, Information, Knowledge, and Wisdom. Recuperado de: <https://homepages.dcc.ufmg.br/~amendes/SistemasInformacaoTP/TextosBasicos/Data-Information-Knowledge.pdf>

CSIRO. (2017). Mining Equipment, Technology and Services: A Roadmap for unlocking future growth opportunities for Australia. Departamento de Ingeniería de Minas de la Universidad de Chile. (2019). Línea de Investigación en Geo-Minero-Metalurgia. Recuperado de: <http://www.minas.uchile.cl/investigacion/lineas-de-investigacion/geo-minero-metalurgia>

Durrant-Whyte, H., Geraghty, R., Pujol, F., & Sellschop, R. (2015). How digital innovation can improve mining productivity. McKinsey & Company.

Global Mining Guidelines Group. (2019). No Title. Recuperado de: <https://gmgroup.org/>

McKee, D. J. (2013). Understanding Mine to Mill.

McKinsey. (2015). Digital America: A Tale of the Haves and Have-Mores.

McKinsey. (2017). Beyond the Supercycle: How Technology Is Reshaping Resources.

World Economic Forum. (2017). Digital Transformation Initiative: Mining and Metals Industry.

# *Big Data*, principales características y desafíos



## Álvaro Egaña

AMTC, Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## Javier Ruiz del Solar

AMTC y Departamento de Ingeniería Eléctrica,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## Richard Weber

Departamento de Ingeniería Industrial,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

El creciente desarrollo y uso de la tecnología computacional y de los dispositivos digitales (Internet, cámaras fotográficas digitales, teléfonos celulares, *streaming* de video, etc.), los cuales generan importantes cantidades de datos, motivó a la definición del concepto del *big data*, término que se refiere a conjuntos de datos que se caracterizan por "las 3 V" (Marr, 2017):

- Volumen: Se tienen que preparar y analizar grandes volúmenes de datos.
- Velocidad: Los datos se generan a alta velocidad (ejemplo: *streaming* de datos).
- Variedad: Se tienen que analizar diferentes tipos de datos, como por ejemplo textos, imágenes, videos, entre otros.



Los grandes volúmenes de datos generados deben ser transmitidos, almacenados, procesados, visualizados y analizados. En la literatura se usa el término *big data analytics* para referirse a técnicas de análisis de grandes volúmenes de datos. El *big data analytics* se basa en el uso de técnicas de aprendizaje de máquinas o de ciencia de los datos.

Desde su definición inicial, el concepto de *big data* se ha ampliado, considerando actualmente cuatro componentes (De Mauro et

al. 2015): la información y cómo esta se genera en forma digital, las tecnologías requeridas para su utilización (almacenamiento, procesamiento, etc.), los métodos utilizados para su análisis y el estudio del impacto en la sociedad y las empresas. De esta forma, hoy en día podemos definir al *big data* como "activos de información caracterizados por tener un gran volumen, velocidad y variedad y que requieren tecnologías y métodos analíticos específicos para su transformación en valor" (De Mauro et al. 2015).

### Definiciones básicas desde una perspectiva histórica

En el siglo XVIII se formalizó la disciplina de la estadística (Hald, 2003), con el objetivo de recolectar, resumir y analizar datos para encontrar información útil, principalmente en el sector público de los "estados". Más adelante, esta disciplina fue utilizada en diferentes áreas en las que se podían recolectar y analizar datos, como la medicina, la ciencia, los negocios, entre otros. Con el desarrollo de los computadores y otros dispositivos digitales creció el volumen de datos disponibles (Williams, 1985) y surgieron nuevas disciplinas que aportaron a su análisis.

Una de ellas es la *inteligencia artificial*, término que se acuñó en una conferencia de Dartmouth (EE. UU.) en 1956 (Bellman, 1978), con el objetivo de denominar la resolución de problemas que hasta el momento estaban reservados a los humanos. En paralelo, una comunidad científica independiente desarrolló lo que hoy conocemos como *redes neuronales artificiales*, que son modelos matemáticos que se inspiran en el funcionamiento del cerebro humano. En aquella época la comunidad científica trabajaba hacia una máquina para resolver problemas generales, el llamado "General Problem Solver" (Newell & Simon, 1963). Sin embargo, al no cumplirse esta expectativa en la década del 70, se provocó el abandono casi total de las investigaciones en este ámbito durante quince años, hasta que en los años 80 una nueva ola de desarrollo de la inteligencia artificial y de las redes neuronales prometió (nuevamente) la creación de máquinas inteligentes. Los focos de la investigación apuntaron, por ejemplo, a áreas como los sistemas expertos, la representación formal de conocimiento, la robótica y el *aprendizaje de máquinas* (*machine learning*), entre otros. La investigación de esa época empezó a poner énfasis en desarrollar sistemas para reconocer patrones en conjuntos de

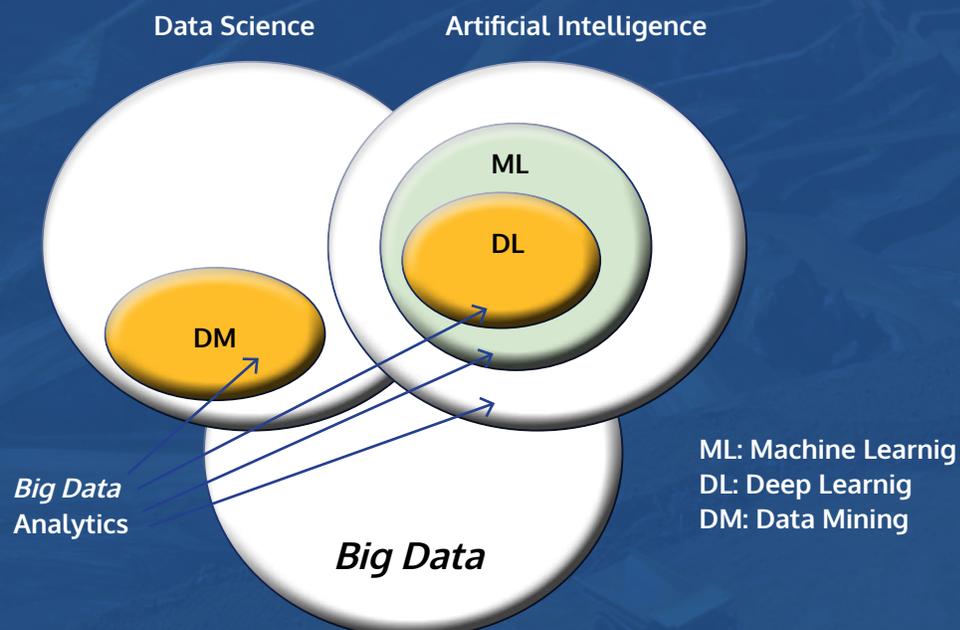
datos (disciplina conocida como *pattern recognition*). Sin embargo, tanto los modelos matemáticos, como también las capacidades de sensores y computadores de esa época no fueron suficientes para cumplir con las expectativas creadas, y hacia finales de los años 90 se produjo un nuevo invierno de la inteligencia artificial (*AI Winter*).

Durante la década del 90 la minería de datos (*data mining*) partió en forma menos ambiciosa. Este enfoque no intenta emular la inteligencia humana, pero sí pretende descubrir conocimiento en bases de datos (según la sigla en inglés KDD: *Knowledge Discovery in Databases*) (Han et al. 2011). Muchos modelos de minería de datos fueron desarrollados o mejorados durante los últimos 25 años. Por otra parte, la llamada ciencia de los datos (*data science*) (Kelleher & Tierney, 2018), se refiere en términos generales a métodos para extraer conocimiento y entender los datos. Mientras la minería de datos pone énfasis en el proceso del análisis de los datos, la ciencia de los datos es una disciplina mucho más amplia que también incluye los desafíos técnicos y científicos de dicho análisis.

Por otro parte, con el invento de las redes neuronales profundas (*deep learning*) (Skansi, 2018) se logró un nuevo salto hacia la resolución de problemas, al construir redes de muchas capas y con un mayor número de conexiones. Estas redes profundas son capaces de resolver problemas de alta complejidad, como por ejemplo, la interpretación de imágenes o la toma de decisiones en juegos de estrategia como ajedrez, go y shogi. Estos avances despertaron (nuevamente) altas expectativas hacia la inteligencia artificial, cuyos desarrollos se concentran principalmente en el aprendizaje de máquinas con el propósito de analizar los datos que se generan en diferentes ámbitos.

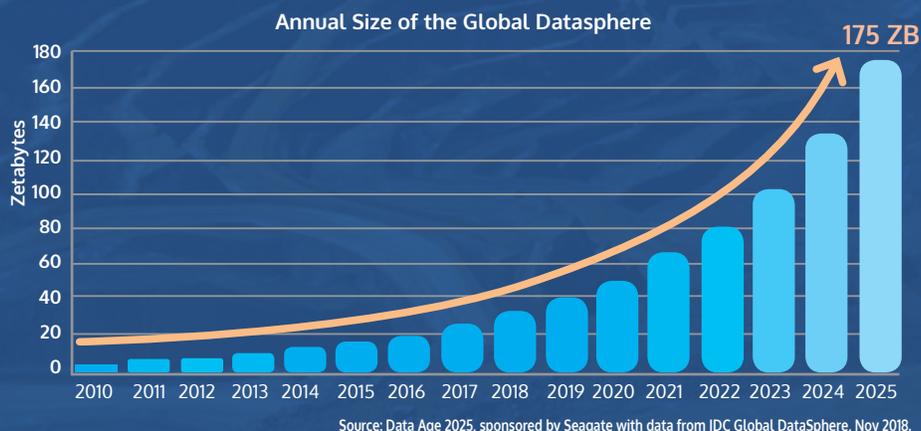
En la figura 1 se muestra la relación entre algunas de las disciplinas anteriormente mencionadas.

**Figura 1.**  
Diagrama que muestra la relación entre disciplinas relacionados al *big data*. Se utilizan los nombres originales en inglés.



**Figura 2.**

Crecimiento de Datos en el mundo, desde 2010 con proyección a 2025.



Fuente: Coughlin (2018).

Para poder comprender el fenómeno del *big data* y lo que impulsó su uso en la última década, es necesario entender el desarrollo de las tecnologías que permiten generar, capturar y almacenar datos masivos. Diferentes tipos de dispositivos que manejan datos digitales han sido desarrollados o mejorados considerablemente en los últimos años, generando cantidades crecientes de datos. A continuación, se presentan brevemente ejemplos de las tecnologías más relevantes en la generación de datos.

- En 2015 había más de 5.000 millones de teléfonos móviles en el mundo que significa una penetración por sobre el 65% (Taylor & Silver, 2019).
- El número de cámaras para diferentes tipos de uso ha aumentado significativamente. Es llamativo el caso de Londres donde se han instalado más de 500.000 cámaras CCTV que captan al ciudadano londinense un promedio 300 veces cada día<sup>1</sup>.

- A fines de 2017, más de la mitad de la población mundial (4.156 millones de personas) navegaba por Internet.
- El llamado Internet de las Cosas (IoT en inglés) se refiere a elementos o dispositivos conectados a Internet (Greengard, 2015). Este incluye elementos de la vida cotidiana como ropa inteligente, electrodomésticos inteligentes, casas inteligentes, brazaletes inteligentes, entre otros, pero también a dispositivos de uso industrial como sensores, robots, máquinas, vehículos y sistemas ciber-físicos en general.

En la Figura 2 se muestra una proyección del crecimiento de datos en el mundo, hasta el año 2025. Según esta proyección, el volumen de datos digitales (*"datasphere"*) crecerá de 33ZB en 2018 a 175ZB en 2025 donde un Zettabyte (ZB) equivale a 1012GB.

<sup>1</sup> <https://www.caughtoncamera.net/news/how-many-cctv-cameras-in-london/>

La llegada y el análisis de los datos masivos han cambiado nuestra vida y seguirán cambiándola en los próximos años. Esta tendencia requiere enfrentar una serie de desafíos (Alharthi et al. 2017). Desde el punto de vista técnico habrá que resolver, entre otros, los siguientes:

- De los datos masivos (*big data*) hay que elegir los datos relevantes (*right data*) para una tarea específica, los cuales corresponden a conjuntos mucho más pequeños (*small data*). Métodos avanzados de selección de atributos o características, y reducción de dimensionalidad son fundamentales para esta tarea.
- Crecientes volúmenes de datos requieren plataformas escalables que permiten un manejo adecuado frente a aumentos en volumen o tráfico de datos.
- La formación de capital humano especializado (técnicos y profesionales con las capacidades y habilidades adecuadas para analizar datos masivos) constituye un desafío para las instituciones de educación técnica y superior, ya que la falta de personal calificado es una real una amenaza para el crecimiento del país.

Pero también existen un sinnúmero de retos no técnicos que tendremos que abordar:

- Las tecnologías para capturar, almacenar y analizar datos masivos han acelerado la globalización de varios modelos de negocio, generando incluso desafíos legales. Empresas como Google, Uber, Amazon, entre otras, operan internacionalmente, por lo que se han presentado brechas legales respecto del uso de los datos masivos de las personas.
- Aspectos de la privacidad reciben cada vez más atención, tanto en la práctica como en la literatura académica (Baesens, 2014). En mayo del 2018 entró en vigencia la *General Data Protection Regulation* (GDPR), que “involucra una serie de principios, derechos y obligaciones, los que tienen el objetivo de resguardar los datos de los usuarios. De esta manera, el usuario deberá dar su consentimiento para que las empresas puedan usar sus datos y ellas deberán informar cuando sean ocupados y con qué finalidad” (Heselaars, 2018).





El fenómeno *big data* exige no solo tener que soportar grandes volúmenes de datos por extensos períodos de tiempo, si no también ser capaces de aumentar (o “escalar”) su capacidad en intervalos de tiempo muy cortos, debido a que las distintas fuentes de datos ahora pueden producir más y más datos en cada vez menos tiempo.



- La ética en el uso de datos masivos es un tema poco estudiado todavía. Al igual que en el caso de la bioética, por ejemplo, habrá que dar respuestas a preguntas como la siguiente: ¿Es éticamente correcto hacer con los datos masivos lo que técnicamente podemos hacer? Al respecto, se puede mencionar el ejemplo de la empresa Cambridge Analytica, acusada de haber recopilado y usado datos de millones de usuarios de una red social, en campañas electorales.

### 1. Elementos metodológicos y tecnológicos

Dentro de este nuevo contexto tecnológico, un error común proviene de pensar que es necesario cambiar radicalmente la infraestructura y los requerimientos básicos, que las tecnologías de información han venido desplegando desde su masificación a mediados de los años 90. Ciertos conceptos, tales como la seguridad, la alta disponibilidad,

la tolerancia a fallas, la consistencia, el almacenamiento redundante, la eficiencia y eficacia en la transmisión y comunicación de datos - siguen siendo igualmente relevantes dentro del nuevo escenario generado por el *big data*. Por lo tanto, cualquier infraestructura orientada en este ámbito, debería contemplar estos mismos requerimientos básicos a la hora de ser desplegada en cualquier ambiente productivo. Todos ellos, por supuesto, aumentados y desafiados ahora por la presencia de una cantidad gigantesca de datos y la necesidad de procesarlos para extraer la información que contienen.

Dicho lo anterior, cabe preguntarse cuál es la estructura típica de una metodología de *big data*, cuáles son sus etapas y sus requerimientos tecnológicos específicos. Uno de los modelos más utilizados para esquematizar la metodología de *big data* se muestra en la Figura 3. Una observación

interesante del modelo es la naturaleza cíclica del proceso de análisis de los datos, que parece calzar con la intuición razonable de que toda cadena de generación de información o conocimiento es iterativa e incremental. Es decir, se aprende más a medida que se profundiza más en el fenómeno en los diferentes ciclos de acercamiento y análisis. Esta es una de las razones por las cuales las *metodologías ágiles*, tales como SCRUM, XP y Kanban, han tenido un acercamiento natural a la metodología de *big data*, ya que ellas se basan en ciclos iterativos e incrementales de acercamiento al problema.

## 2. Definición del problema

En esta etapa de la metodología hay dos elementos tecnológicos importantes a considerar:

- Manejo de la variedad en la captura y fuentes de datos. Se refiere a cuáles van a ser los dispositivos o elementos genera-

dores de datos que, por supuesto, deben estar perfectamente alineados con los objetivos del problema. Pueden incluir desde distintos tipos de sensores, ubicados de tal manera de medir eventos o variables a partir de diferentes etapas de la cadena productiva, hasta sistemas de encuestas, de categorización o incluso monitores de opinión, que capturen el comportamiento subjetivo humano dentro de la misma cadena productiva.

- Escalabilidad de las tecnologías de almacenamiento. El fenómeno *big data* exige no sólo tener que soportar grandes volúmenes de datos por extensos períodos de tiempo (meses o años), si no también ser capaces de aumentar (o "escalar") su capacidad en intervalos de tiempo muy cortos (semanas, días o incluso horas), debido a que las distintas fuentes de datos ahora pueden producir más y más datos en cada vez menos tiempo.

**Figura 3.**

Diagrama que muestra el ciclo de vida de un proceso de Ciencia de Datos típico. Modificado a partir de Cady (2017, p.9).





Cada tecnología de almacenamiento utiliza su propio esquema para definir lo que se conoce como *data set* o conjunto de datos. La popularización del uso de Apache Hadoop y su ecosistema como herramienta para *big data*, ha ido introduciendo una idea un poco más amplia que el antiguo concepto de *data warehouse*.

### 3. Organización y comprensión de los datos

Una vez que los datos están disponibles, a través de distintas fuentes, y almacenados en medios confiables y altamente escalables, surge la necesidad natural de llevarlos a un soporte común y amigable para los diversos métodos y herramientas de análisis. Cada tecnología de almacenamiento utiliza su propio esquema para definir lo que se conoce como *data set* o conjunto de datos. La popularización del uso de Apache Hadoop y su ecosistema como herramienta para *big data*, ha ido introduciendo una idea un poco más amplia que el antiguo concepto de *data warehouse*, donde los datos se encuentran filtrados, pre-procesados y listos para ser utilizados en procesos de análisis de negocios. Este es el concepto de *data lake*, que al igual que el *data warehouse* es un intento por abstraerse de las tecnologías

de almacenamiento. El *data lake* se orienta a mantener todos los datos en su estado más crudo, permitiendo que los distintos métodos de análisis puedan acceder eficientemente a ellos.

En la sub etapa de *análisis exploratorio* de los datos se busca visualizar -a través de una gran variedad de gráficos- reportar y transformar los datos de diferentes formas, hasta poder detectar patrones que motiven el planteamiento de hipótesis concretas respecto a los procesos que se intentan modelar. Un ejemplo relevante de este proceso en minería se encuentra en el análisis geometalúrgico, donde se busca relacionar procesos de naturaleza muy distinta a lo largo de la cadena de valor minera. Para ello hay muchas herramientas en el mercado. Dentro de las más destacables se encuentran Tableau y MS Power BI, que proveen interfaces gráficas bastante fáciles de usar.

#### 4. Procesamiento de los datos

En esta etapa se incluyen los llamados pre-procesos de limpieza y filtrado de datos espurios o valores atípicos, e imputación de datos. Esto último se entiende como el proceso de manejo de datos faltantes, que puede incluir su estimación, siguiendo diversos métodos, o simplemente su eliminación.

Sin embargo, el aspecto más importante en el procesamiento de los datos es lo que se conoce como ingeniería de características (*feature engineering*) y que está compuesta por tres etapas principales:

- Selección de características: Consiste en la selección de las características más adecuadas para el modelo en base a criterios estadísticos/numéricos respecto a la característica en estudio (variable de respuesta) u objetivo del proceso de análisis. Un ejemplo muy básico es el método que selecciona las características con mejor correlación con la variable de respuesta.
- Extracción de características. Se refiere a la transformación de datos crudos en características informativas o numéricas que puedan ser incorporadas en él o los modelos posteriores.

Para el entrenamiento de modelos con datos históricos se han ideado muchas estrategias. Una de las más exitosas está basada en el viejo concepto computacional de "dividir para reinar".

- Construcción o síntesis. Consiste en la creación, o síntesis, de nuevas características a partir de funciones de las características ya existentes, también buscando maximizar ciertos criterios respecto a la característica en estudio.

## 5. Modelamiento y análisis de los datos

Lo relevante en esta etapa, respecto al contexto *big data*, es tener en cuenta que cualquiera sea el modelo de análisis de los datos escogido, necesitará pasar por una etapa en la que tendrá que aprender los parámetros del modelo a partir de una cantidad masiva de datos, conocida como *fase de entrenamiento del modelo*. El término "aprender", en este contexto, se refiere a la aplicación de una serie de métodos estadísticos y matemáticos que exigirán al máximo las capacidades de cómputo (CPU,

Desde el punto de vista temporal, los datos pueden ser *históricos* o *en tiempo real*. En términos prácticos, entendiendo que cada caso particular mueve estas definiciones hacia caminos levemente distintos, cualquier evento que no esté ocurriendo ahora (tiempo real) es histórico.

Para el entrenamiento de modelos con datos históricos se han ideado muchas estrategias. Una de las más exitosas está basada en el viejo concepto computacional de "dividir para reinar": una gran tarea se divide en muchas pequeñas subtarefas, que se resuelven ya sea



GPU) y de almacenamiento, virtual (RAM) o físico (discos), de la infraestructura elegida. Aquí surge nuevamente la necesidad de que una infraestructura para *big data* sea altamente escalable, tanto en capacidad de cómputo como en capacidad de almacenamiento, como ya se mencionó.

en paralelo (en muchas CPU/GPU dentro de una misma máquina) o en forma distribuida (en muchas máquinas), coordinadas bajo un esquema eficiente de sincronización y balance de carga - cuantas sub tareas se destinan a determinada unidad de cómputo. En cualquiera de los dos casos, el *big data*,

aumenta fuertemente los requerimientos de comunicación, transporte, eficiencia y consistencia de los datos. En este ámbito, un modelo de programación muy conocido -y extendido en su uso- es el llamado MapReduce (Dean & Ghemawat, 2008). Este modelo fue ideado hace más de quince años en Google y su objetivo original era calcular los índices para cantidades masivas de documentos en la Web. Su carácter general y la aparición de Hadoop - uno de los primeros paquetes de software en proveer una implementación genérica de MapReduce - lo hizo rápidamente aplicable a otro tipo de tareas, como el procesamiento de datos o el entrenamiento de modelos.

Actualmente, Hadoop es en realidad un ecosistema completo de plataformas, alojadas por la fundación Apache, que proveen soporte al espectro de requerimientos de una infraestructura de cómputo paralela y distribuida (Hadoop, 2018). Dentro de los más destacados se puede mencionar: HDFS (sistema de archivos distribuido), MapReduce, YARN (administración de recursos distribuidos), Pig (lenguaje de consulta para componentes distribuidas), HBase (base de datos NoSql distribuida),



*Mahout (Machine Learning sobre Hadoop), Sqoop (para importar datos desde distintos sistemas de almacenamiento en Hadoop), Ambari (seguridad) y Oozie (para manejo de workflows).*

Una gran limitación de estas estrategias “dividir para reinar”, es que el proceso de entrenamiento del modelo debe ser susceptible de ser dividido en sub tareas completamente independientes entre sí. En minería hay dos importantes ejemplos de modelos en donde esto no es posible: el problema del cálculo del *pit* final en planificación, porque cada componente (o bloque) tiene una fuerte dependencia estructural, espacial y temporal, de las otras componentes; y el problema de la estimación geoestadística en la evaluación de recursos, debido a la correlación espacial que debe ser considerada entre todas las componentes. En estos casos, la situación se vuelve aún más desafiante y debe ser abordada mediante otras estrategias, como por ejemplo, la utilización de arquitecturas de hardware con memoria compartida.

Cuando los datos fluyen en tiempo real (*data streaming*), *Hadoop* no es la mejor alternativa, principalmente porque su diseño original no consideró modelos posibles de ser entrenados incrementalmente, en línea, a medida que los datos fueran llegando. Por otra parte, dado que el modelo puede ir cambiando constantemente con los nuevos





datos, la infraestructura de base debe considerar principalmente un sistema de transporte de datos que garantice eficiencia y consistencia en todo momento. A pesar de ser un área de desarrollo reciente, hoy existen varios paquetes de software disponibles para esta tarea, siendo los principales Spark, Kafka, Flink, Storm y Samza.

Finalmente, dado que, en el contexto de *big data*, la demanda de recursos computacionales puede crecer indefinidamente, para cualquier industria es difícil planificar y sostener una estrategia de crecimiento constante en infraestructura que la soporte. En respuesta a esta demanda nació el concepto de "arquitecturas de cómputo en la nube" (*cloud computing architectures*). La idea central es que los requerimientos de cómputo y almacenamiento sean consumidos como servicios, provistos por un tercero, a través de Internet. Por ejemplo, si se necesitan diez máquinas para entrenar o procesar un modelo bajo un esquema distribuido, ya no hay que comprarlas físicamente, si no "comprar" el

acceso a ellas por Internet, sin importar donde estén ubicadas y como estén administradas. Y sencillamente usarlas. Los requerimientos centrales bajo este esquema son: a) el ancho de banda en el acceso a Internet, para no limitar la capacidad en el transporte de los datos y b) que el diseño del software sea capaz de soportar llamadas a servicios remotos. Para ambos requerimientos hoy existen excelentes soluciones: las súper carreteras digitales para lo primero, y protocolos bien establecidos para lo segundo, tales como Web Services y REST; además de proveedores que han ido acumulando valiosas experiencias, ofreciendo servicios eficientes y confiables, tales como Amazon Web Services, MS Azure, Google Cloud, entre muchos otros.

## 6. Agradecimientos

Se agradece el aporte de las siguientes personas, cuyas ideas contribuyeron a la escritura de este capítulo: Rodrigo Cortés, Rocío Nuñez, Juan Velásquez y Pablo Estévez.



# Referencias

34

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to *big data*. *Business Horizons*, 60(3), 285–292.

Baesens, B. (2014). *Analytics in a big data World: The Essential Guide to Data Science and its Applications*. Wiley.

Bellman, R. (1978). *An Introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser.

Cady, F. (2017). *The Data Science Handbook*. John Wiley & Sons.

Coughlin, T. (2018). 175 Zettabytes by 2025. Recuperado de: <https://www.forbes.com/sites/tomcoughlin/2018/11/27/175-zettabytes-by-2025/#71b84a4a5459>

De Mauro, A., Greco, M., & Grimaldi, M. (2015). What is big data? A consensual definition and a review of key research topics. In *AIP Conference Proceedings* (Vol. 1644, pp. 97–104).

Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107–113.

Greengard, S. (2015). *The Internet of Things*. MIT Press Essential Knowledge series.

Hadoop. (2018). MapReduce Tutorial. Apache Hadoop. Recuperado de: [https://hadoop.apache.org/docs/r1.2.1/mapred\\_tutorial.html#Purpose](https://hadoop.apache.org/docs/r1.2.1/mapred_tutorial.html#Purpose)

Hald, A. (2003). *A History of Probability and Statistics and Their Applications Before 1750*. Hoboken, NJ: John Wiley & Sons.



Han, J., Kamber, M., & Pei, J. (2011). *Data Mining: Concepts and Techniques* (3ra edición). Morgan Kaufmann.

Heselaars, T. (2018, May 25). A chile le afecta: Lo que debes saber sobre las nuevas reglas europeas de control de datos digitales. Recuperado de: <https://www.emol.com/noticias/Tecnologia/2018/05/25/907330/Lo-que-debes-saber-sobre-el-reglamento-de-proteccion-de-datos-europeo-que-comienza-a-regir.html>

Kelleher, J. D., & Tierney, B. (2018). *Data Science*. MIT Press.

Marr, B. (2017). *Data strategy: how to profit from a world of big data, analytics and the internet of things*. Kogan Page Publishers.

Newell, A., & Simon, H. A. (1963). GPS: A Program that Simulates Human Thought. In E. A. Feigenbaum & J. Feldman (Eds.), *Computers and Thought*. New York: McGraw-Hill.

Skansi, S. (2018). *Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence*. Springer-Verlag.

Taylor, K., & Silver, L. (2019). Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Recuperado de: <https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/>

Williams, M. R. (1985). *A History of Computing Technology*. Prentice-Hall Series in Computational Mathematics.

# *Big Data en exploración minera*

## **Álvaro Egaña**

AMTC, Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## **Felipe Navarro**

AMTC, Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## **Brian Townley**

AMTC y Departamento de Geología,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## **Diana Comte**

AMTC y Departamento de Geofísica,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## **Rocío Núñez**

Beauchef Minería,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile



## 1. Introducción

Hasta hace unos pocos años, tal vez esta era una de las áreas en la cadena minera donde la aplicación de la ciencia de datos en general, y el *big data* en particular, estaba prácticamente inexplorada y desarrollada a través de aplicaciones aisladas de métodos estadísticos espaciales (Lisitsin 2015). A pesar de que la situación ha ido cambiando paulatinamente y que algunos trabajos iniciales incorporan técnicas de *machine learning* tales como *support vector machine* y *random forest* para definir y categorizar objetivos de exploración (Abedi et al. 2012, Rodríguez-Galiano et al. 2014, Sadeghi et al. 2015, Gandhi & Sarkar 2016, McCuaig & Hronsky 2017), aún queda mucho por hacer. Lo que puede ser visto como una oportunidad. Conducir buenos estudios, podría impactar de manera significativa en el proceso mismo. Si se considera que los blancos de exploración fáciles de detectar ya han sido encontrados y agotados, es importante, de manera de poder encontrar los blancos difíciles, que el proceso esté dirigido fuertemente desde la experiencia y que esté sustentado en habilidades técnicas importantes. Estas habilidades técnicas pueden ser enormemente potenciadas por el uso de herramientas de análisis de datos, sobre todo si se cuenta con una gran cantidad de ellos.



En el caso de la exploración básica, normalmente se cuenta con información pública (imágenes satelitales, en espectro visible, multi e hiperespectrales, mapas geológicos, etc.) que es aprovechada cualitativamente para generar blancos de exploración (Gandhi & Sarkar 2016, McCuaig & Hronsky 2017). Este proceso tiene una dinámica bastante consolidada a través del reconocimiento geológico y el análisis geoquímico de superficie, lo que ha dado como resultado el descubrimiento de grandes yacimientos que llevan años de explotación. Es aquí donde surge la gran oportunidad de aprovechar el análisis de los datos del proceso de exploración original, en conjunto con el conocimiento experto y los datos históricos acumulados a través de la operación en el proceso minero. Es decir, se cuenta con el conocimiento y los datos en retrospectiva, en superficie y en profundidad.

Esto permitiría construir y alimentar modelos que sean capaces de ayudar a definir y localizar nuevos blancos de manera más precisa, permitiendo cuantificar y reducir la incertidumbre, y por tanto, los costos inherentes al proceso.

Esta misma visión se potencia aún más para las etapas de exploración intermedia y avanzada al observar que el conocimiento experto, el volumen y la calidad de la información solo tienden a crecer. En este caso, la creación de modelos que permitan un análisis

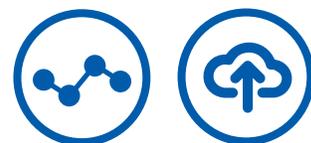
profundo de estos volúmenes de datos permitiría no solo la disminución de la incertidumbre en la construcción del conocimiento del yacimiento, sino también la generación de herramientas que permitan optimizar y reducir la incertidumbre en el proceso mismo de obtención de estos nuevos datos, como lo son las campañas de sondajes. Con estos nuevos modelos se podrían responder preguntas como: ¿dónde perforar los primeros sondajes después de la exploración básica?, ¿dónde perforar los siguientes sondajes de tal manera de aumentar las posibilidades de tocar cierto nivel de mineralización?, ¿dónde perforar de tal manera de disminuir la incertidumbre respecto al conocimiento actual del modelo geológico, geoquímico o geomecánico? Como se ve, son muchas las posibilidades y oportunidades que se vislumbran.



Un error común que se comete al considerar el análisis de datos en *big data* y el aprendizaje de máquinas, es pensar que al incorporar sistemas expertos automatizados, estos podrían reemplazar el criterio de un ser humano.



Un error común que se comete al considerar el análisis de datos en *big data* y el aprendizaje de máquinas, es pensar que al incorporar sistemas expertos automatizados, estos podrían reemplazar el criterio experto de un ser humano (por ejemplo, un geólogo de exploración). Esto está muy lejos de conseguirse, tanto en la práctica como en la teoría. De la misma manera que el incorporar el uso de un GPS a un vehículo no reemplaza a su conductor, si no que más bien lo convierte en un conductor más eficiente. El tener más información disponible le otorga al experto la posibilidad de enfocar el proceso de toma de decisiones en los puntos críticos de su negocio.



## 2. ¿Qué se hace hoy?

El desarrollo del negocio minero descansa en dos aspectos fundamentales: la demanda por productos minerales y la existencia natural de recursos (Fig. 1). Para el desarrollo del negocio minero la primera etapa consiste en la búsqueda de yacimientos minerales, la cual se inicia con exploraciones mineras, desde básicas hasta avanzadas (Fig. 1). Es necesario señalar que cada empresa minera tiene su propia metodología para hacer exploración minera y una forma particular de desarrollarla, por lo tanto, pretender describir un proceso estandarizado no es realista. No obstante, si se piensa en la gran minería, el estilo es por etapas, como sigue a continuación:

- ① Determinar qué explorar.
- ② Etapa de campo y resonancia magnética.
- ③ Mapeo, muestreo geoquímica y geofísica.
- ④ Planificación de targets.
- ⑤ Perforación diamantina inicial.
- ⑥ Perforación diamantina avanzada.
- ⑦ Evaluación de recursos.

Lo primero es identificar si el blanco de exploración es *greenfield* o *brownfield*. Esta diferencia no solo tendrá impacto en la cantidad de información asociada, sino también en el objetivo de la exploración, la complejidad de las distintas etapas a realizar y la inversión económica asociada.

Una vez determinado el tipo de blanco a explorar, se hace un reconocimiento y una recopilación de información para definir un área favorable a dónde ir y qué ir a explorar.

La exploración *greenfield* se refiere a los proyectos que parten desde cero con respecto al terreno, ya que se desarrollan en áreas nuevas. La exploración *brownfield*, por otra parte, es una exploración que se desarrolla en terrenos existentes, donde existe un yacimiento previamente descubierto y en operación.

Una vez determinado el tipo de blanco a explorar, se hace un reconocimiento y una recopilación de información (antecedentes) para definir un área favorable a dónde ir y qué ir a explorar. En este punto, algunas empresas poseen bases de datos con información proveniente de diferentes fuentes, la cual debe ser correctamente analizada y sometida a un control de calidad. Una vez determinada el área, se realizan mapeos geológicos, de

mineralización y alteración (dependiente del ambiente geológico) en diferentes escalas, tales como 1:100000; 1:20000; 1:5000. Estos mapeos se asocian a una campaña de muestras geoquímicas para definir un blanco de exploración. En este blanco, se harán campañas de geofísica (gravimetría, magnetometría, radiometría) y se hará uso de sensores remotos (satelitales y aerotransportados), las cuales permiten generar un modelo 3D inicial de interpretaciones a partir de secciones. Con esta información, se define un objetivo de exploración distrital. En este punto se considera una primera campaña de sondajes.





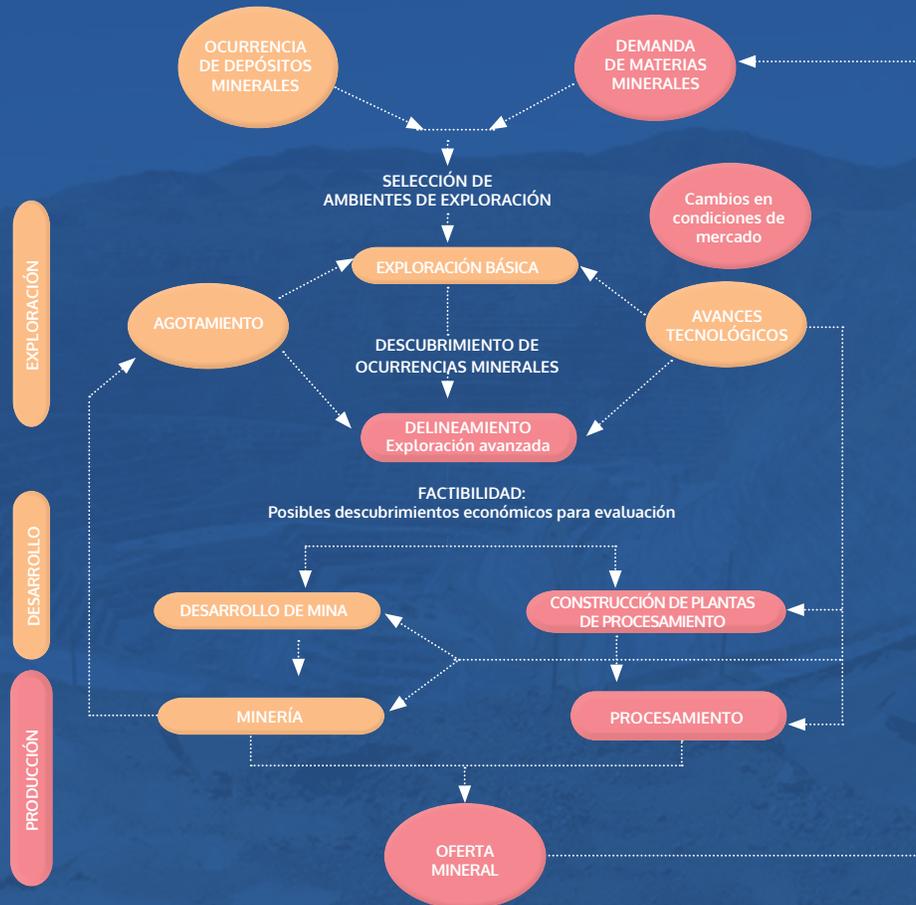
Si la perforación resulta exitosa, se busca definir cuál es el área mineralizada y la variabilidad de leyes. Estos potenciales blancos corresponden normalmente a áreas de 20x20km. Toda la información que es levantada en los pasos mencionados se traspasa a alguna herramienta de software, donde se hacen modelos del mineral y del tipo de recursos. En el caso de que la campaña de sondajes no sea exitosa, se abandona el lugar o se vuelven a analizar los datos para definir un nuevo blanco.

Además de los diferentes tipos de información recopilada, sea esta de carácter cualitativo o cuantitativo, hay diferentes actores que participan del proceso completo de exploración, lo cual lo complejiza aún más. Solo por nombrar algunos: geólogos de exploración, contratistas de ingeniería y construcción.

**Figura 1.**

Etapas principales y sucesión de actividades del negocio minero, gatilladas por la existencia de demanda de recursos minerales y la ocurrencia y disponibilidad de estos recursos en la corteza terrestre

Etapas de desarrollo del negocio minero.



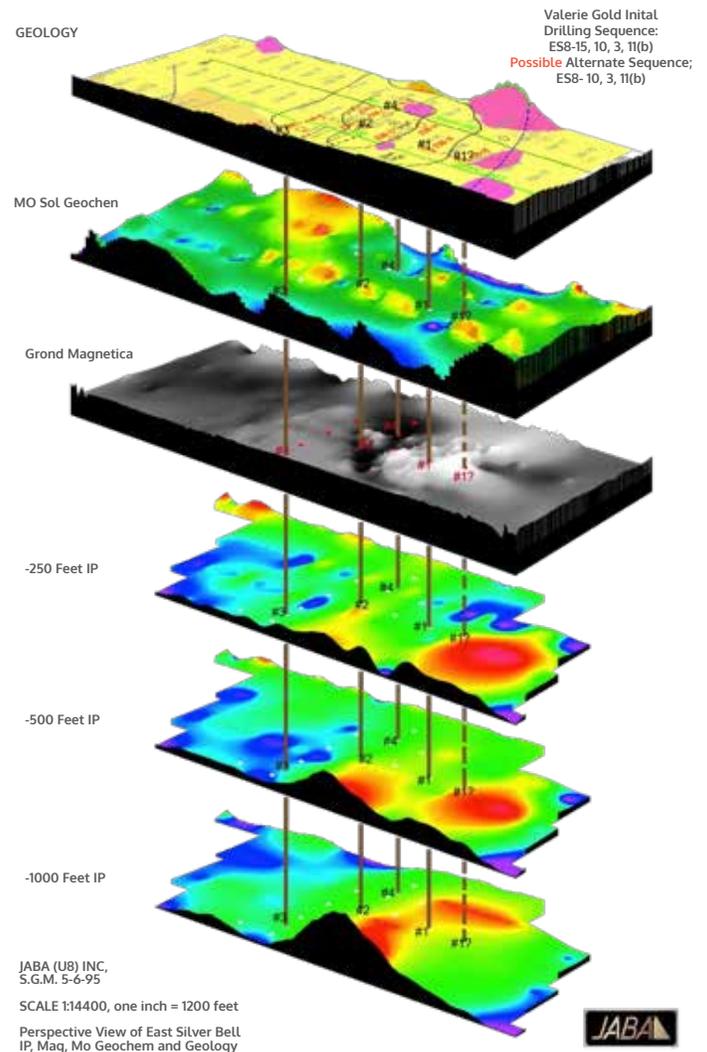
### 3. Desafíos y oportunidades

Las actividades progresivas descritas en la etapa de exploración minera recopilan información de diverso tipo, escala y resolución en la medida en que avanza de exploraciones básicas (selección de ambientes) a exploraciones avanzadas y evaluación y delineación de yacimientos.

La información empleada incluye variables categóricas (ej. geología) y numéricas (ej. geofísica, geoquímica), donde estas últimas pueden ser continuas o discretas, y donde el soporte o resolución de la información puede ser de diversas escalas. Lo anterior plantea desafíos importantes al momento de procesar, visualizar e interpretar toda la información, sobre todo si dicha interpretación se desea realizar de forma integral y transversal.

Para asistir el procesamiento de toda esta información existen, al presente, plataformas digitales con Sistemas de Información Geográfica (SIG) (McCuaig & Hronsky 2017), las cuales permiten almacenar, actualizar, revisar, visualizar, recuperar, procesar, manipular e integrar diferentes datos e información geoespaciales, ayudando a manejar la información en forma de capas. Estas capas de información pueden tener distintas dimensiones, escalas y resolución o soporte de datos (Fig. 2). Si se considera que en estos sistemas cada capa representa una base de datos –sean estos continuos, discontinuos, discretos, categóricos o numéricos– se presenta la oportunidad y el desafío, por cierto, de cruzar transversal y espacialmente toda la información, independiente de su tipo o resolución. Si bien los SIG permiten realizar comparaciones y análisis cruzados de información, cuando existen muchas capas de información y enormes volúmenes de datos (*big data*), se vuelve extremadamente complejo hacer todo en forma manual, sin existir en la actualidad herramientas que determinen qué aspectos cruzar y por qué. En estos sistemas, si bien se puede analizar la varianza de la distribución, por ejemplo, solamente se logra capa por capa. La comparación transversal, a través de las distintas capas sobre un soporte común, es realizada típicamente de forma visual, dependiendo exclusivamente de la habilidad del experto. Es aquí donde se presenta la oportunidad de modelar y automatizar el despliegue de relaciones no convencionales entre los datos, de tal manera que el experto pueda enfocarse en el análisis de información disponible en distintos niveles de abstracción.

**Figura 2.** Ejemplo de manejo de información en exploraciones mineras. Geología representa la base de información (superior), integrada espacialmente con geoquímica, magnetometría terrestre, e información de resistividad (IP) a 200, 500 y 1000 pies de profundidad.



En la actualidad, el desarrollo progresivo de tecnologías de exploración, directa e indirecta, permite levantar de forma rápida una enorme cantidad de información en terreno.

### 3.1 Incorporación de nuevas fuentes de información

En la actualidad, el desarrollo progresivo de tecnologías de exploración, directa e indirecta, permite levantar de forma rápida una enorme cantidad de información en terreno, la cual se puede integrar con información pre-existente con el fin de elaborar bases de datos integrales y fácilmente desplegables en algún SIG. En la medida en que se acumula información se vuelve progresivamente



más complejo y lento procesar, interpretar y aprovechar al máximo toda la información, y más aún el determinar las interrelaciones que puedan existir entre las distintas capas de información y el cómo vincular información de distinto tipo, escala y resolución.

Por lo anterior, se han desarrollado sistemas de manejo e integración de bases de datos. También los SIG se han modernizado y aumentado sus capacidades, convergiendo lentamente a sistemas digitales de *big data*, donde los datos son ahora procesados mediante técnicas de *data science*, ya sea con métodos supervisados (usuario experto) o bien incorporando métodos no supervisados, propendiendo a sistemas de *Machine Learning*.

En las bases de datos de exploraciones se combina información geológica, que representa la base sobre la cual se pueden incorporar capas de información indirecta. Entre ellas, información geofísica (ej. magnetometría, gravedad, tomografía sísmica, etc.), geoquímica (ej. geoquímica de rocas, regolito, suelo, sedimentos, etc.), sensores remotos (ej. imágenes hiperespectrales, radar, etc.), entre muchos otros tipos. Cabe destacar que en el caso de la información indirecta, lo que se determina o mide es una respuesta a una propiedad mineral y/o geológica. Por ejemplo, la información geofísica representa una respuesta física de los minerales; las propiedades magnéticas se verán reflejadas en la magnetometría, la densidad en su gravimetría, la porosidad/permeabilidad y rigidez en una tomografía sísmica, etc. Por otro lado, en geoquímica se refleja la composición química de las rocas, suelos o sedimentos, lo que es función de la litología, la alteración y la mineralización, siendo por tanto un reflejo indirecto de las rocas en

Dado que las propiedades físicas y químicas de los minerales serán las que determinen las respuestas de los diversos tipos de información, la convergencia de propiedades es la que puede otorgar mayor confianza en la interpretación y determinación de blancos de exploración y/o delineación de potenciales yacimientos.



cuestión. En sensores remotos, por su parte, la reflectancia a una diversidad de bandas del espectro electromagnético, refleja la composición mineral, entre otros aspectos.

Se puede ver que todas las herramientas e información levantadas deben corresponder a propiedades geológicas presentes en la geología estudiada. Los mapas geológicos representan así las propiedades en superficie.



A partir de las herramientas indirectas se busca interpretar la continuidad en subsuelo, e idealmente detectar aspectos no determinables en la geología superficial que puedan indicar la presencia de posibles yacimientos en profundidad. Dado que las propiedades físicas y químicas de los minerales serán las que determinen las respuestas de los diversos tipos de información, la convergencia de propiedades es la que puede otorgar mayor confianza en la interpretación y determinación de blancos de exploración y/o delimitación de potenciales yacimientos. El manejo de grandes volúmenes de información y las dificultades de cruzarlo, integrarlo e interpretarlo de forma rápida y eficiente, representa un desafío que se puede abordar precisamente mediante técnicas de *big data*, *data science* y *machine learning*.

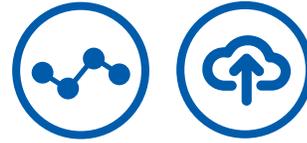
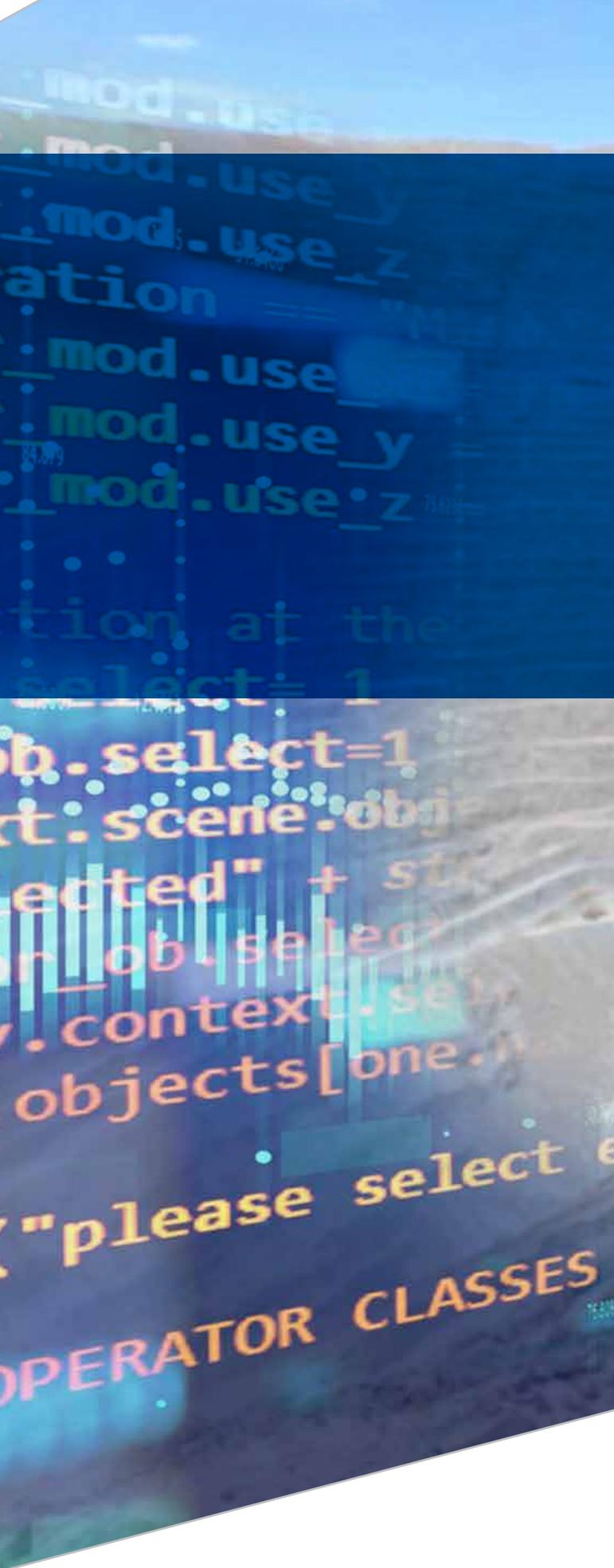


*Big data* en las Ciencias de la Tierra se caracteriza por la gran variedad de observables en los procesos asociados. Ejemplos de ello son las observaciones satelitales, las estaciones meteorológicas y de boyas, las estaciones GPS (permanentes y temporales), y las redes de estaciones sismológicas.

48

El uso más eficiente de los métodos indirectos, como la geofísica o tomografía sísmica, cruzando datos con observaciones, permitirán caracterizar mejor los sectores estudiados, junto con análisis químicos más sistemáticos, petrografía y análisis de imágenes hiperespectrales que ayuden a la interpretación. Por otro lado, la oportunidad de adquirir información en el corto plazo mejorará todo el proceso de exploración, dado que algunos análisis pueden tomar semanas. Uno de los usos actuales de este tipo de tecnología son las pistolas XRF, las cuales son usadas en minería de lantánidos, facilitando la interpretación de tipos de arcillas (imperceptible al ojo del geólogo). Además, facilitan la identificación del tipo de alteración al que se enfrentan. Por otra parte, el hacer uso de todas las variables que se generan en un estudio de exploración y determinar las relaciones entre ellas hará más realista los modelos de yacimientos.





Finalmente, el uso de *big data* ayudará a crear una base de datos integrada, que permitirá acceder fácilmente a los estudios previamente realizados en una región o alimentar modelos de aprendizaje usando diferentes fuentes de datos.

### 3.2 Tomografía sísmica y sus potenciales alcances

*Big data* en las ciencias de la Tierra se caracteriza por la gran variedad de observables en los procesos asociados. Ejemplos de ello son las observaciones satelitales, las estaciones meteorológicas y de boyas, las estaciones GPS (permanentes y temporales), y las redes de estaciones sismológicas. Estos datos ayudan a abordar una amplia gama de preguntas geocientíficas, en escalas de tiempo tan breves como minutos (terremotos, por ejemplo) y procesos variables en el tiempo sobre escalas de tiempo humanas, que son directamente relevantes para la existencia de la humanidad, como el calentamiento global. Se emplean en grandes simulaciones con modelos que poseen millones de parámetros para explicar estos datos y desentrañar los sistemas físicos gobernantes y sus propiedades (Comte et al. 2016, Chi-Durán et al. 2017).



En la actualidad, los yacimientos cuyas huellas superficiales eran más claras ya han sido explotados, por lo que la profundidad se ha transformado en un enemigo del geólogo de exploración, dando más importancia a métodos indirectos, como la geofísica. Esto se suma a que las inversiones en exploración son consideradas un gasto para las mineras, las cuales no apoyan los proyectos a menos que se trate de un lugar con información relativamente segura. Como es fácil caer en un argumento circular en este punto, ya que la única manera de obtener esa información relativamente segura es precisamente explorando, surge la necesidad de desarrollar nuevas técnicas de exploración cuyos costos sean menores a los que se enfrentan en la actualidad.

Un área de las ciencias de la tierra donde la investigación algorítmica toma un papel clave en la minería de datos y la interpretación, es la sismología. Los científicos de la tierra no pueden observar el interior de ella directamente. Todas las observaciones se hacen de manera indirecta y la forma dominante en la que aprendemos sobre el interior de la tierra es a través de las ondas sísmicas. Las ondas generadas por sismos y terremotos se propagan con velocidades que dependen del medio y las variaciones de estas velocidades nos enseñan sobre la composición del núcleo y manto, la tectónica de

El uso principal de la sismicidad es ofrecer una oportunidad para obtener imágenes del interior de la tierra. La industria adquiere datos sísmicos a través de hasta 1.000.000 sensores distribuidos de manera extremadamente densa en una pequeña área de la superficie de la tierra, con el objetivo de apoyar la exploración y producción de recursos subterráneos.

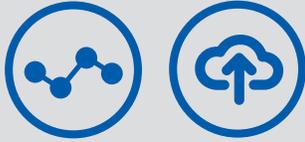
placas, las presiones, los tipos de rocas, las zonas de fracturas, los fluidos, y en particular, los recursos minerales y campos geotermales. La mejor fuente conocida son los terremotos, pero hay muchas más fuentes, donde se incluye el viento, el tráfico, los volcanes, las explosiones y los deslizamientos de tierra. El uso principal de la sismicidad es ofrecer una oportunidad para obtener imágenes del interior de la tierra.

La industria adquiere datos sísmicos a través de hasta 1.000.000 sensores distribuidos de manera extremadamente densa en una pequeña área de la superficie de la tierra, con el objetivo de apoyar la exploración y producción de recursos subterráneos (inicialmente hidrocarburos, pero actualmente también se aplica a depósitos minerales y campos geotermales).

El software utilizado para el procesamiento de datos en sismología debe combinar aspectos de extracción de datos, reconocimiento de patrones y reducción de mapas, para manejar datos heterogéneos de manera eficiente, veraz y automática. Los elementos del aprendizaje automático desempeñan un papel clave en la superación de estos retos. Debido a que los datos sísmicos se utilizan para restringir los procesos de la tierra que son poco conocidos, es un desafío procesar los datos sin saber exactamente qué señales esperar. Además, aunque empleamos sistemas sofisticados de ecuaciones diferenciales parciales para explicar los datos, gran parte de nuestros datos no se ajustan a nuestros modelos simplificados, por ejemplo, las formas de onda complejas, el diferente contenido de frecuencias en las ondas sísmicas y el registro continuo de ruido sísmico.

En los últimos años, muchos gigantes de recursos naturales han agregado datos para identificar las formas más rápidas, más seguras y más rentables de descubrir y extraer minerales de la tierra y prevenir eventos de riesgo. Mirando en el futuro cercano, las compañías mineras también aplicarán el análisis de *big data* a la fase de exploración.

Cuando los científicos e ingenieros buscan depósitos no descubiertos, deben adquirir, procesar e interpretar conjuntos de datos masivos, incluyendo la información geológica, geofísica y geoquímica derivada del área de exploración específica.



#### 4. Visión y comentarios finales

Dave Lowell, uno de los descubridores de Escondida, en su libro “Intrepid Explorer – The Autobiography of the World’s Best Mine Finder” comparte nueve reglas que él ha seguido en sus muchos años como experto en exploración minera (Lowell 2014). Aunque el título nos pudiera parecer un poco exagerado, vale la pena reflexionar sobre algunas de estas reglas:

- *“Mines are found in the field, not the office”*. Para todo geólogo de exploración esto debiera ser de sentido común. Sin embargo, la complejidad que representa hacer nuevos hallazgos, principalmente en profundidad, exige cada vez más un fuerte componente de análisis de datos, como hemos dicho, de distintas fuentes de información. Aún más, considerando que la mayoría de las reservas y recursos en Chile fueron descubiertos antes del año 2010, considerando exploración en áreas expuestas y subaflorantes. El verdadero desafío será explorar en áreas cubiertas.

Por lo tanto, la pasión por el trabajo de campo para los actuales y futuros geólogos de exploración debiera estar sustentada en sólidos conocimientos de ciencia y análisis de datos. Lowell lo expresa, de cierta forma, al decir *“Mines are now almost always found by drilling holes, so if no part of the budget is spent on drilling there is almost no chance of success. (Pierina was the rare exception)”*, ya que las campañas de sondajes son, hoy por hoy, la principal fuente de extracción de datos e información.

- *“Exploration is a cost/benefit business*. En contextos como la mediana y pequeña minería, la definición de “bajo costo” es bastante más dinámica y crítica, por cierto, que para la gran minería. Pero en todos los casos, es vital optimizar el uso de recursos en las campañas de sondajes. Para ello es crucial contar con modelos que permitan disminuir el riesgo y la incertidumbre de tal manera que, idealmente, cada sondaje realizado aporte el mayor valor posible al proceso de exploración.



Aquí la importancia del desarrollo de nuevas tecnologías más económicas y amigables con el medioambiente, como lo es la Tomografía Sísmica.

- *“High-tech devices and geophysical surveys are very rarely of value in mine discovery.* A pesar de que, como hemos dicho, la intuición y el conocimiento de los expertos están muy lejos de poder ser reemplazados por la tecnología, es innegable que esta sí aporta valor al proceso.

Volviendo al ejemplo del uso del GPS, si bien es posible orientarse usando “métodos tradicionales”, como la posición del sol o las estrellas, el GPS aporta confiabilidad, eficiencia y seguridad en el proceso de exploración de campo. Esperamos que en un futuro no muy lejano, podamos disponer de elementos tecnológicos, que tal vez no garanticen de forma alguna el descubrimiento de nuevos yacimientos (como un GPS), pero que provean a los expertos con orientaciones e indicaciones cada vez

El verdadero desafío será explorar en áreas cubiertas. Por lo tanto, la pasión por el trabajo de campo para los actuales y futuros geólogos de exploración debiera estar sustentada en sólidos conocimientos de ciencia y análisis de datos.

más inteligentes y acertadas de dónde se podrían encontrar estos nuevos hallazgos.

- *“Finding mines is a high-risk business.* Exacto. ¿Pero cómo se evalúan estos riesgos asociados? La ciencia de datos (Data science) provee de variadas herramientas teóricas y prácticas que permiten hacer esto posible. Pero esto, requiere de un cambio de visión tanto en los expertos como en los tomadores de decisión.

El avance de la computación y el desarrollo de algoritmos para la simulación de escenarios han permitido, en distintas industrias, la reducción de incertidumbre. En minería esto ya es una realidad y en exploración no será diferente. Para ello, el uso de datos históricos y simulación de escenarios, apoyado por técnicas de aprendizaje estadístico y machine learning, permitirán cuantificar la incertidumbre asociada a estos procesos.

Finalmente, una estrategia a nivel país para dar disponibilidad a datos públicos asociados a nuestro territorio se hace crucial. Ejemplos de este tipo de políticas públicas se encuentran en países como Australia (Geoscience Australia), Canadá (Natural Resources Canada) o EE.UU (USGS), los cuales disponibilizan al público gran cantidad de información asociada a topografías, mapeos geológicos, datos geofísicos, mapas geotermales, imáge-



nes satelitales multiespectrales, entre otras. Esto sería de gran ayuda, al centralizar y canalizar la información para ayudar al proceso de exploración minera desde una visión global.

Agradecemos el apoyo de los geólogos Carlos Marquardt, Pamela Pérez, Ximena Pérez, Martín Marquardt y Jan Gröpper, por compartir sus puntos de vista sobre el tema a través de entrevistas que fueron una fuente invaluable de información.

# Referencias

Lowell, Dave (2014). *Intrepid Explorer – The Autobiography of the World’s Best Mine Finder*. Tucson, Arizona: The University of Arizona Press.

Chi-Durán, R., Comte, D., Díaz, M., & Silva, J. F. (2017). Automatic detection of P- and S-wave arrival times: new strategies based on the modified fractal method and basic matching pursuit. *Journal of Seismology*, 21(5), 1171–1184. <https://doi.org/10.1007/s10950-017-9658-0>



McCuaig, T. C., & Hronsky, J. (2017). The mineral systems concept: the key to exploration targeting. *Transactions of the Institutions of Mining and Metallurgy, Section B: Applied Earth Science*. <https://doi.org/10.1080/03717453.2017.1306274>

Comte, D., Carrizo, D., Roecker, S., Ortega-Culaciati, F., & Peyrat, S. (2016). Three-dimensional elastic wave speeds in the northern Chile subduction zone: Variations in hydration in the supraslab mantle. *Geophysical Journal International*, 207(2), 1080–1105. <https://doi.org/10.1093/gji/ggw318>

Gandhi, S. M., & Sarkar, B. C. (2016). *Essentials of Mineral Exploration and Evaluation*. *Essentials of Mineral Exploration and Evaluation*. <https://doi.org/10.1016/B978-0-12-805329-4.00017-X>

Lisitsin, V. (2015). Spatial data analysis of mineral deposit point patterns: Applications to exploration targeting. *Ore Geology Reviews*. <https://doi.org/10.1016/j.oregeorev.2015.05.019>

Sadeghi, B., Madani, N., & Carranza, E. J. M. (2015). Combination of geostatistical simulation and fractal modeling for mineral resource classification. *Journal of Geochemical Exploration*. <https://doi.org/10.1016/j.gexplo.2014.11.007>

Rodriguez-Galiano, V. F., Chica-Olmo, M., & Chica-Rivas, M. (2014). Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, Southern Spain. *International Journal of Geographical Information Science*. <https://doi.org/10.1080/13658816.2014.885527>

Abedi, M., Norouzi, G. H., & Bahroudi, A. (2012). Support vector machine for multi-classification of mineral prospectivity areas. *Computers and Geosciences*. <https://doi.org/10.1016/j.cageo.2011.12.014>

# *Big Data* en operaciones mineras

## **Luis Felipe Orellana**

AMTC y Departamento de Ingeniería de Minas,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## **Nelson Morales**

AMTC y Departamento de Ingeniería de Minas,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## **Alejandro Ehrenfeld**

AMTC y Departamento de Ingeniería de Minas,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

56



## 1. Introducción

La explotación de los recursos minerales ha transitado a lo largo de la historia desde la recolección más rudimentaria en tiempos pasados hasta hoy, donde métodos de explotación masivos y/o selectivos alojan operaciones mineras cada vez más complejas por sus condiciones geográficas, geológicas y ambientales, pero también más sofisticadas debido a las diversas tecnologías que coexisten. Por ejemplo, en este último caso, la introducción de maquinaria y su crecimiento en tamaño y productividad a mediados del siglo pasado, y más recientemente los procesos de automatización parcial y completa de los sistemas de carguío y transporte de mineral, han significado cambios profundos en la filosofía de operación de distintas operaciones en Chile y el mundo.

Al respecto, los avances en tecnologías de sensorización, captación y capacidad de transmisión de datos, han hecho visible y disponible una mayor cantidad y variedad de información para los y las operadores de las explotaciones mineras. Así, hoy se puede acceder en línea a la localización espacial de un equipo de transporte, es posible disponer de su estado mecánico o electrónico, revisar videos de su desplazamiento y acceder a información agregada sobre la congestión en las diferentes rutas. Por otro lado, la perforación, la tronadura y el carguío no solo aportan datos operacionales vitales sobre los equipos que las involucran, sino que también se están convirtiendo en fuentes continuas de variables geo-minero-metalúrgicas. Finalmente, otros avances tecnológicos nos permiten identificar y monitorear variables también críticas para la operación, como la estabilidad de taludes, la locación de las personas y su información biométrica, e imágenes y videos



desde plataformas audiovisuales montadas en drones, y otros, disminuyendo así la exposición de las personas a los peligros inherentes de las operaciones mineras.

Pero ¿qué podemos hacer frente a toda esta masividad y diversidad de datos?

Los análisis masivos de datos o el *big data* aparece como un cambio revolucionario en la forma de hacer minería. El *big data* aplicado a la minería permitirá viabilizar esta masividad y diversidad de datos en la operación diaria mediante la capacidad de cómputo y la síntesis de información para así proveer a los y las operarios mineros con las herramientas e información correctas. De esta manera, el *big data* no solo proporcionará información para saber qué está pasando, sino que también, para que las personas puedan tomar las mejores decisiones en tiempo real, con una visión integrada a los otros procesos del negocio minero, proyectando así el impacto de sus decisiones en el futuro de la operación. La información, sumada a la capacidad analítica, permitirá a los y las operarios de una mina, la mejora continua de los procesos, detectando cuellos de botella, previniendo inconvenientes, mejorando la integración aguas abajo, todo en un ambiente más seguro de la operación.

Para la actividad minera no es nuevo enfrentar y absorber procesos radicales de cambios tecnológico. Sin embargo, la nueva revolución tecnológica que enfrenta la minería producto del *big data* representa un desafío mayor respecto de los grandes cambios tecnológicos

anteriores, como consecuencia de los importantes efectos en productividad esperados. Así, en su dimensión más amplia, el *big data* permitirá profundizar una nueva ola de eficiencias en la extracción del mineral mediante la facilitación de mayor y mejor información en todos los niveles de decisión. De igual manera, será posible reconocer patrones de comportamiento y relaciones previamente desconocidos en los sistemas mineros. Más aún, la utilización del *big data* impactará necesariamente en la morfología de las organizaciones. Esto es bien comprendido por las operaciones y compañías mineras, las cuales por ejemplo ya han creado o están construyendo grupos de científicos de datos al interior de sus organizaciones.

No obstante los impulsos tecnológicos, son las tradiciones, lenguajes locales, los usos o reglas empíricas las que moldean cada operación y las maneras en las nuevas tecnologías son implementadas, estableciendo así convivencias a ratos caóticas, no obstante, naturales. En efecto, no es casualidad encontrar referencias del tipo "la minería no es una ciencia, es un arte". Más aún, esta convivencia entre la irrupción tecnológica y la cultura organizacional propia de cada operación es consecuencia inherente de la unicidad de los yacimientos, lo que se traduce en singularidades del proceso minero, constituyendo enraizadas prácticas operacionales que no se pueden perder de vista.

En este capítulo, abordaremos el uso e implementación de *big data* en la operación minera. Fundamentalmente, exploraremos el conjunto subprocesos o subactividades - desde el arran-

que hasta la descarga de mineral en la planta de procesamiento de minerales - que agrupados, permiten el cumplimiento de los planes de producción mineros. No analizaremos, en consecuencia, los impactos del *big data* durante el diseño del proyecto de minero, la construcción de la infraestructura o la preparación minera, entre otros.

Sin embargo, si bien el foco de este capítulo está centrado en las operaciones tal como lo hemos definido en el párrafo anterior, no hay



que desconocer que existen aplicaciones interesantes en otros ámbitos de la minería que están asociados directamente o conviven a la operación minera. En efecto, tanto durante el diseño, la preparación mina o a la construcción de infraestructura se pueden destacar importantes actividades que están experimentando cam-

bios relevantes producto de la implementación de *big data*. Podemos destacar, por ejemplo.

- La creciente digitalización de la información durante la construcción de infraestructura y preparación minera, lo cual se refleja en múltiples aspectos, en particular la eliminación de la duplicidad de datos al eliminar reingresos, la homogeneización y disponibilidad más eficiente de las bases de datos, así como también su propio enriquecimiento mediante, por ejemplo, geolocalización o inclusión de imágenes del estado de frentes o eventos.
- La utilización de *Building Information Modelling*<sup>1</sup> (BIM) en los proyectos mineros, lo cual permite no solo el mejor seguimiento y gestión de los proyectos de construcción, sino que también una mejor gestión de los activos una vez se encuentran en operación (Eastman et al., 2008; Zhou et al., 2017). El uso de BIM se constituye como reservorio de la información abriendo las puertas a su utilización como fuente de conocimiento para proyectos futuros.
- La construcción de estructuras de costos o indicadores clave (KPI) robustos a través de bases de datos sistemáticamente habitadas y estructuradas, permiten la corrección y proyección de estimaciones y términos de referencia futuros tanto para la toma de decisiones en línea como información como para su uso como insumo en procesos de planificación de largo plazo.

<sup>1</sup> ISO 19650-1:2018 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles - <https://www.iso.org>

Para la actividad minera no es nuevo enfrentar y absorber procesos radicales de cambios tecnológicos. Sin embargo, la nueva revolución tecnológica que enfrenta la minería producto del *big data* representa un desafío mayor respecto de los grandes cambios tecnológicos anteriores, como consecuencia de los importantes efectos en productividad esperados.



## 2. ¿Qué se hace hoy?

La determinación del método de extracción, a cielo abierto o minería subterránea, está basada, entre otros parámetros de referencia, por las características espaciales y geométricas del yacimiento, la riqueza del cuerpo mineralizado y el tipo de mineral, las condiciones geomecánicas y geológicas, consideraciones ambientales y licencias sociales para operar, la tecnología vigente, el marco regulatorio, la seguridad de la operación en su sentido más amplio, variables económicas y vaivenes del mercado, entre otros (Brady & Brown, 2006; Deb & Sarkar, 2017; Wyllie & Mah, 2017). Luego, será el método de extracción el que determinará los aspectos claves de la operación minera.

La operación minera es el proceso de la explotación de minas en minería a cielo abierto o subterránea que agrupa todos los subprocesos o subactividades orientadas al cumpli-

miento del plan de producción minero. Entre las actividades de la operación minera contamos la perforación y tronadura o excavación de rocas en su sentido más general, los sistemas de manejo de minerales (carguío y transporte, equipos de reducción de tamaño, otros), actividades auxiliares o servicios de suministro, y la planificación minera de corto plazo. En el caso de minería subterránea es necesario también agregar las subactividades asociadas a la ventilación de la faena, monitoreo sísmico, así como también en ciertas ocasiones considerar chancadores primarios.

Durante la operación minera, se genera una gran cantidad de variables operativas mínimas y propias de cada sistema minero. Dependiendo de si la operación es a cielo abierto o subterránea, podemos mencionar - sin ser exhaustivos ni específicos - las leyes del material (%), consumo de explosivos (g/t) y accesorios de voladura, rendimientos de perforación (m/h), productividades (t/h), tiempos de ciclo (s), velocidades de los camiones (m/s) el consumo de diésel de los equipos de carguío y transporte (lt/h), velocidades de extracción (t/m<sup>2</sup>/h), generación de polvo, consumo energético de los ventiladores (kWh), fragmentación y distribuciones de tamaño ( $d_{50}$ ,  $d_{80}$ ), humedad, sismicidad, etc. También variables categóricas, por ejemplo, la presencia o no de barro en los puntos de extracción de minería de caving, cierre y apertura de sectores en producción, entre otros.

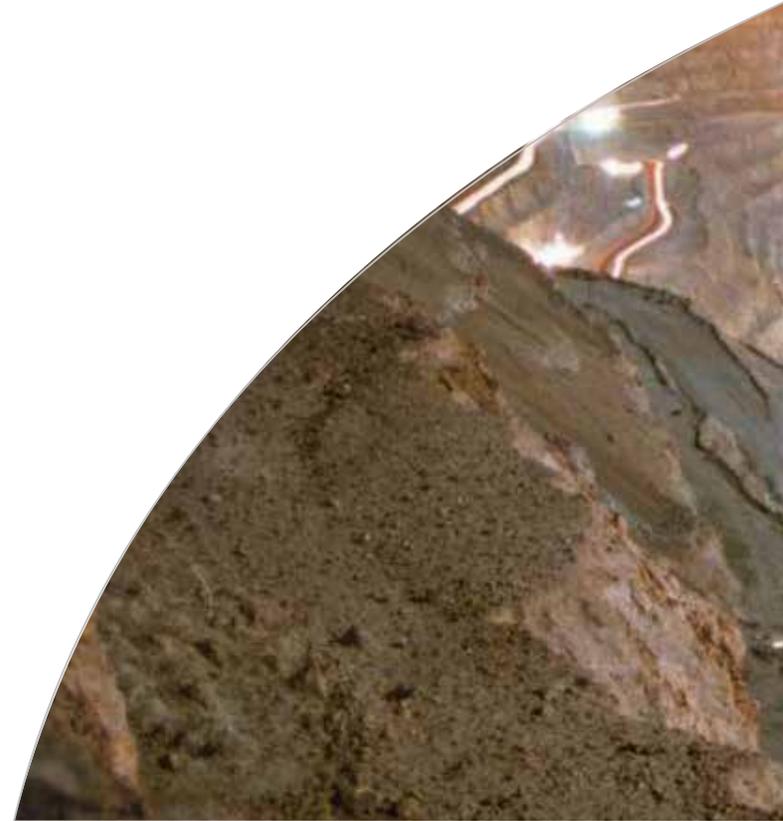
De la mano de la remotización y la automatización de los sistemas mineros, del uso de tec-



nologías como el GPS, la digitalización masiva y en general, del abaratamiento de diversos tipos de sensores disponibles en el mercado, nuevos escenarios de operación han surgido. En efecto, estas condiciones han permitido un aumento sostenido del volumen generado de información de las variables operativas de cada sistema minero, y la aparición de nuevas y múltiples fuentes de datos. Sumado a lo anterior, un incremento considerable de la velocidad de transmisión y almacenamiento de datos. Así, nuevas fuentes de información han sido incorporadas progresivamente a la operación minera, complejizando y desafiando el sistema de manejo y gobernanza de datos, pero a su vez, iluminando vacíos de información. Por ejemplo, es posible capturar datos en línea sobre el “estado de salud” de los equipos mineros, disponer de información geo-minero-metalúrgica avanzada de los frentes de avance en minería, generar protocolos de seguimiento y monitoreo de las personas, ya sea por sistemas de TAG, información biométrica o de imágenes, analizar la distribución espacio-temporal de sismicidad inducida, entre otros.

Todos estos elementos, pilares intrínsecos que favorecen la pronta implementación del *big data* en las operaciones mineras.

Como hemos indicado, todas estas variables operativas del sistema demandan y determinan la instrumentación necesaria para una operación minera estable e idealmente continua. Sin embargo, estas variables habitan en escenarios naturales (ej. incertidumbre geo-



lógica, resistencia de la roca, fallas geológicas, flujos de agua, etc.) y operacionales (ej. sismicidad inducida, levantamiento de polvo, etc.) de gran incertidumbre y variabilidad (Brady & Brown, 2006; Deb & Sarkar, 2017; Wyllie & Mah, 2017). Así, la calidad, precisión y volumen de datos generados no solo dependerá del arreglo de instrumentos y sensores disponibles, sino que también estará condicionada al método de explotación utilizado y las condiciones ambientales de operación.

En minería a cielo abierto, ya desde la perforación se puede capturar data tanto de los



equipos involucrados como del frente mismo siendo perforado. Aguas abajo, es posible medir la fragmentación y distribución de tamaño resultante de la tronadura a través de análisis de imágenes (Sereshki et al., 2016). Además, la utilización de drones y videocámaras tienen el potencial de generar abundantes volúmenes de fotografías y horas de video continuo de la operación, que incluyen aplicaciones en el ámbito de la seguridad, la estimación de volúmenes de material, la reconciliación y la caracterización de leyes. El uso de radares para el análisis de estabilidad de taludes permite un monitoreo en línea fo-

calizado de los desplazamientos y la generación de alarmas frente a inestabilidades estructurales (Kumar & Rathee, 2017). Las palas, camiones, equipos auxiliares, ya sea autónomos o tripulados, son monitoreados en línea, con la posibilidad de acceder a información detallada de su funcionamiento y estado mecánico y electrónico, pero también mediante arreglos de sistemas de geolocalización y/o sistema de tags, vía distintos software que complementan la data del sistema de despacho, generando sendas bases de datos que permiten identificar zonas lentas, perfiles de velocidad de la flota, cuellos de botellas, en-

La minería subterránea ve la incorporación de aplicaciones de *big data* con cierto rezago respecto de las operaciones a cielo abierto. Este fenómeno es natural ya que la minería subterránea se desarrolla en ambientes de operación incluso menos favorables que la minería a cielo abierto.



tre otros análisis en línea . Finalmente, el levantamiento de data geometalúrgica en los frentes permite pensar en el tracking a nivel de palada o camión y la consecuente optimización tanto de la operación como del posterior procesamiento del mineral.

La minería subterránea ve la incorporación de aplicaciones de *big data* con cierto rezago respecto de las operaciones a cielo abierto. Este fenómeno es natural ya que la minería subterránea se desarrolla en ambientes de operación incluso menos favorables que la minería a cielo abierto. En efecto, si bien la minería a cielo abierto puede localizarse en lugares geográficamente aislados (desierto, alta montaña, etc.), la minería subterránea impone adicionalmente restricciones a la transmisión

de datos y generación de información producto del difícil acceso a la ubicación de los subprocesos de la operación minera que se busca monitorear. Por ejemplo, en minería subterránea masiva tipo *block y panel caving*, la complejidad del proceso de *caving* o hundimiento condiciona que ciertas variables o eventos críticos no se puedan medir o monitorear con la certeza deseada ej. agua - barro, altura de columna de explotación, colgaduras, entre



otros (Castro et al., 2018; Orellana et al., 2017; K. Sánchez et al., 2019; V. Sánchez et al., 2019). En efecto, para diversos parámetros de operación en minería de *caving*, aún es necesario muestreos para la medición, o simplemente no existe la instrumentación adecuada, lo cual impide la medición ciertas variables de la operación. Esto último resulta en volúmenes de datos más bien restringidos, no sólo

en calidad de información, sino que también en cantidad. Esta restricción de información tiene el potencial de impactar directamente en el cumplimiento del plan de producción minero, pues al no identificar las fuentes de información en el sistema, los y las tomadores de decisión se ven obligados u obligadas a interrumpir la capacidad productiva del ciclo minero ej. cierre de puntos de extracción.

Pese a lo anterior, existen diversos sensores e instrumentos que se utilizan en minería subterránea. Por ejemplo, para monitorear la sismicidad de los distintos sectores en explotación se recurre a arreglos geófonos y acelerómetros los cuales acoplados con sistemas de adquisición de datos permiten su monitoreo continuo. Tecnología tipo 3D scanner tipo láser integrados con sistemas de procesamiento de datos pueden ser utilizados para el monitoreo de los daños y sobre excavaciones generadas por tronadura en túneles, galerías, y cavernas (Monsalve et al., 2018). Esta misma tecnología es utilizada para el monitoreo de carpetas de rodados y así generar programas de mantención ad-hoc evitando sobrecostos por reparación.

La progresiva automatización del sistema de manejo de minerales es clave para entender la evolución experimentada por minas subterráneas y su relación con el manejo y procesamiento de grandes volúmenes de datos. Martillos picadores y LHD semiautónomos son claros ejemplos de sistemas remotos de telecomando, que han permitido aumentos en productividad, así como también sustantivas

mejoras en seguridad al disminuir la exposición de los y las operadoras a la faena. En ambos casos, los equipos generan información en línea que es relevante para la operación. Actualmente, los Centros Integrados de Operaciones (CIO) han jugado un importante rol en este aspecto según ya se ha introducido en el Capítulo "Big data en Minería".

La operación minera y sus subprocesos o subactividades son sin duda, un conjunto crítico de la industria minera. En resumen, los aspectos más relevantes del uso actual del *big data* en la operación minera se pueden agrupar en los siguientes:

Captura y uso en línea del estado y funcionamiento de equipos. Producto del monitoreo continuo y en vivo de las variables operaciones de los equipos, la industria está migrando a modelos de falla y mantenimiento predictivos, en donde la decisión de realizar mantenimiento a un equipo se basa en "síntomas", facilitando que la mantención se ejecute previo a la ocurrencia de la falla. De esta forma se aumenta la predictibilidad del sistema, el cual tiene una operación más estable, evitando así los tiempos perdidos adicionales que puede generar, por ejemplo, un equipo parado en ruta u obstaculizando un frente.

Utilización de nuevas y diversas fuentes de información (imágenes, videos, biometría, etc.) para robustecer datos existentes y optimizar el funcionamiento de la operación. El uso de nuevas fuentes complementarias de información permite a la operación minera

contrastar la coherencia de los datos - eliminando los datos corruptos - que, a su vez, facilitan la construcción de análisis más robustos para la toma de decisiones en tiempo real. Así, con una adecuada gobernanza y en un escenario de decisiones complejas en ambientes operacionales adversos, es posible optimizar el funcionamiento de la faena mediante la entrega de información relevante y oportuna de los procesos, aumentando la seguridad del personal minero, y agregando valor al negocio minero.

Enriquecimiento de la base de datos geo-minero-metalúrgica de la operación para alimentar el despacho de flotas, la planificación a corto plazo y una mejor integración con los procesos extractivos. En este caso, se pueden considerar todas las tecnologías que permiten la captura de información en el frente de extracción, la cual en conjunto con herramientas de seguimiento y los softwares de despacho de flotas, permiten una mejor supervisión de cada equipo y sus actividades. Así, por ejemplo, cada baldada (y sus propiedades) del sistema de manejo de materiales puede ser monitoreada, con los consecuentes beneficios para i) la toma de decisiones en tiempo real: ¿dónde enviar esa baldada? ¿cómo ajustar la planta en función de las propiedades del mineral en las próximas horas o turnos?; ii) la planificación de corto y mediano plazo, al corregir la secuencia de perforación, y iii) la planificación de largo plazo, proveyendo modelos enriquecidos que permitan modelar la incertidumbre y hacerla parte del proceso de planificación en sus distintas etapas.

### 3. Desafíos y oportunidades

En el contexto actual de las operaciones mineras, podemos clasificar los desafíos del *big data* en tres aspectos: condiciones ambientales adversas de la operación, desarrollos tecnológicos, y la dimensión humana. Revisemos cada uno de ellos.

Diversos aspectos pueden mencionarse en relación con los desafíos del *big data* y las condiciones ambientales adversas y los desarrollos tecnológicos necesarios para enfrentarlas. Por ejemplo, más allá de las características propias de cada método de

explotación, en ambos casos existen condiciones adversas para la tecnología digital, ya sea por presencia de polvo, humedad, u otras condiciones específicas. Esto convierte al diseño de *hardware* y su mantención en un proceso clave que los desarrollos tecnológicos, tanto a nivel de equipos como tratamiento y transferencia de datos, sean específicos para la operación minera. En efecto, esto último es en particular relevante ya que tanto en minería a cielo abierto como minería subterránea el cociente *volumen de información/ancho de banda* es grande: en el caso de la minería a cielo abierto, a consecuencia de la gran cantidad de equipos distribuidos sobre áreas signifi-





ficativas, y en el caso de minería subterránea, realizado por las condiciones ambientales adversas y dificultades inherentes a la transmisión y disponibilidad de infraestructura. Más aún, el caso de la minería es todavía más complejo si se considera que las condiciones medioambientales son permanentemente dinámicas producto de la evolución de los yacimientos, los cuales son cada vez más profundos, con características geomecánicas más complejas, o donde incluso puede haber transición de un método de extracción a otro, o coexistencia de los mismos, con las complejidades para el diseño, operación y planificación minera que estos escenarios implican (Flores & Catalan, 2019; King et al., 2017; Morales et al., 2019; Parra et al., 2018)

Paradójicamente, al mismo tiempo que las condiciones de una operación pueden cambiar drásticamente, los tiempos de puesta en marcha de un proyecto minero pueden estar fácilmente en el rango de una década y luego la vida de un proyecto puede durar varios decenios. Esto hace que la tecnología deba cambiar constantemente, implicando que el diseño del proyecto original no sea el más apto para la adopción de nuevos elementos. Todas estas dificultades aparecen como particularmente importantes si se miran en el contexto más amplio, en donde la industria minera es relativamente pequeña si se compara con otros como la agricultura y la construcción con los cuales comparte proveedores.

Sin embargo, el desafío importante del *big data* en minería no es sólo técnico en su dimensión más ingenieril o tecnológica, sino que también reacciona a la dimensión humana de la operación. En efecto, existen al menos dos elementos que conviene destacar de esta dimensión humana: la gobernanza de los datos, y los cambios, rotación e interacción de las personas.

Respecto de la gobernanza, se debe mencionar como desafío importante una larga historia de grandes barreras contractuales, impidiendo el acceso libre de la información por parte del dueño a datos y variables operacionales alojados tanto en equipos mineros y obras tercerizadas respectivamente. Un segundo elemento se refiere a la dificultad de generar instancias y procedimientos comunes incluso en faenas mineras que perte-

El desafío importante del *big data* en minería no es solo técnico en su dimensión más ingenieril o tecnológica, sino que también reacciona a la dimensión humana de la operación. En efecto, existen al menos dos elementos que conviene destacar de esta dimensión humana: la gobernanza de los datos, y los cambios, rotación e interacción de las personas.

necen a un mismo controlador y que requiere pasar de una cultura en donde existen dueños de ciertos datos, a una en donde la operación es la propietaria y distintos usuarios pueden acceder a ella en forma integrada. En particular, no existe así igual manejo de contenidos de datos, texto, multimedia y uso de software impidiendo la normalización de información y adecuada organización. Esta situación pone en desventaja a la industria y también genera pérdidas en el valor del negocio.

Desde el punto de vista de la interacción de las personas, es necesario generar los puentes y lenguajes comunes entre los y las es-

pecialistas mineros y mineras y los científicos de datos. En este sentido, es importante que los y las ingenieros de mina comprendan las potencialidades y limitaciones del *big data*, manejando un lenguaje común que les permita interactuar con la misma facilidad con que lo hacen con geólogos o proveedores de equipos. Recíprocamente, el conocimiento específico de la industria minera debe permear hacia los científicos de datos que se especialicen en esta área, a fin de asegurar, por ejemplo, un tratamiento apropiado de los datos y selección correcta de las metodologías para su análisis.

Otro de los desafíos interesantes del *big data* en las operaciones mineras corresponde a la consolidación de los datos originados de distintas fuentes y/o subprocesos, en un sistema robusto y cohesivo. Esto requiere, como se indicaba anteriormente, de una gobernanza clara y eficiente que permita a las operaciones aumentar eficientemente el valor del negocio, pero también de la definición de estándares que permitan bases de datos comunes e integradas.





En efecto, una visión más estratégica que puede facilitar la implantación de *big data* supone identificar las fuentes de valor de cada uno de esos productos en la cadena de operación, donde se originan esos valores y cómo afectan a los procesos que le siguen.

#### 4. Visión y comentarios finales

La industria minera no está ajena a esta nueva ola de fuerzas disruptivas inducidas por la revolución de los datos. El *big data* llegó para quedarse y convertirse en una componente fundamental del negocio minero. En efecto su uso cambiará no solo la forma en que se operan las minas, sino que también la manera en que se organizan las compañías controladoras.

Los desafíos son vastos e importantes. Las particularidades de una operación minera, sobre todo en minería subterránea, hacen que las componentes técnicas sean particularmente desafiantes. Se debe ir hacia modelos en donde la necesidad de la operación gatille la aplicación de *big data* y en donde la gestión del cambio, tanto a nivel individual como cultural de las compañías, aparece como clave para el éxito de estas metodologías.

Los futuros operarios y operarias mineros dispondrán de una capacidad analítica sin precedentes, serán eficientes en el uso de la información y se adaptarán a escenarios de cambios dinámicos potenciados por la disponibilidad de información y algoritmos que lo liberarán de labores más rutinarias para llevarlo al ámbito de la optimización de la operación minera.

Algunos aspectos importantes en este sentido tienen que ver con la deslocalización de ciertas decisiones, las cuales por diversos motivos son tomadas, ya sea con información parcial (simplemente porque no está disponible), o no se dispone de la capacidad para anticipar sus efectos en otras partes de la operación, o en el mediano o largo plazo.

#### 5. Agradecimientos

Quisiéramos agradecer la colaboración de Nancy Pérez, Directora Senior del Área de Transformación en Ingeniería y Construcción de la Vicepresidencia de Proyectos, de CODELCO y a Cristian López, director del Área de Transformación Digital de División El Teniente, de CODELCO, a quienes entrevistamos y contribuyeron de manera muy importante para dar su visión en y experiencia en *big data* en minería. También quisiéramos agradecer a Roberto Miranda, Jefe de Proyectos del Laboratorio de Geomecánica de AMTC, quien contribuyó con ideas y textos iniciales para este capítulo.

# Referencias

Brady, B. H. G., & Brown, E. T. (2006). *Rock Mechanics for underground mining: Third edition*. Rock Mechanics for underground mining: Third edition.

Castro, R., Garcés, D., Brzovic, A., & Armijo, F. (2018). Quantifying Wet Muck Entry Risk for Long-term Planning in Block Caving. *Rock Mechanics and Rock Engineering*, 51(9), 2965–2978.

Deb, M., & Sarkar, S. C. (2017). Sustainable Development of Mineral Resources. In *Springer Geology*.

Eastman, C., Teicholz, P., Sacks, R., & Liston, K. (2008). *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. Wiley Publishing.

Flores, G., & Catalan, A. (2019). A transition from a large open pit into a novel “macroblock variant” block caving geometry at Chuquicamata mine, Codelco Chile. *Journal of Rock Mechanics and Geotechnical Engineering*, 11(3), 549–561.

King, B., Goycoolea, M., & Newman, A. (2017). Optimizing the open pit-to-underground mining transition. *European Journal of Operational Research*, 257(1), 297–309.

Kumar, A., & Rathee, R. (2017). Monitoring and evaluating of slope stability for setting out of critical limit at slope stability radar. *International Journal of Geo-Engineering*, 8(1).

Marinelli, M., Lambropoulos, S., & Petroustas, K. (2014). Earthmoving trucks condition level prediction using neural networks. *Journal of Quality in Maintenance Engineering*, 20(2), 182–192.

McGaughey, J. (2019). Artificial intelligence and big data analytics in mining geomechanics. In J. W. (Ed.), *Proceedings of the Ninth International Conference on Deep and High Stress Mining* (pp. 45–54). The Southern Africa Institute of Mining and Metallurgy. Recuperado de: [https://papers.acg.uwa.edu.au/p/1952\\_04\\_McGaughey/](https://papers.acg.uwa.edu.au/p/1952_04_McGaughey/)

Monsalve, J. J., Baggett, J., Bishop, R., & Ripipi, N. (2018). Application of laser scanning for rock mass characterization and discrete fracture network generation in an underground limestone mine. *Proceedings of the 37th International Conference on Ground Control in Mining, ICG-CM 2018*, 29(1), 183–192.

Morales, N., Seguel, S., Cáceres, A., Jélvez, E., & Alarcón, M. (2019). Incorporation of geome-tallurgical attributes and geological uncertainty into long-term open-pit mine planning. *Minerals*, 9(2).

Orellana, L. F., Castro, R., Hekmat, A., & Arancibia, E. (2017). Productivity of a Continuous Mining System for Block Caving Mines. *Rock Mechanics and Rock Engineering*, 50(3), 657–663.



Parra, A., Morales, N., Vallejos, J., & Nguyen, P. M. V. (2018). Open pit mine planning considering geomechanical fundamentals. *International Journal of Mining, Reclamation and Environment*, 32(4), 221–238.

Peng, S., & Vayenas, N. (2014). Maintainability Analysis of Underground Mining Equipment Using Genetic Algorithms: Case Studies with an LHD Vehicle. *Journal of Mining*, 2014(5), 1–10.

Pérez-Barnuevo, L., Lévesque, S., & Bazin, C. (2018). Automated recognition of drill core textures: A geometallurgical tool for mineral processing prediction. *Minerals Engineering*, 118(September 2017), 87–96.

Ribeiro e Sousa, L., Miranda, T., Leal e Sousa, R., & Tinoco, J. (2017). The Use of Data Mining Techniques in Rockburst Risk Assessment. *Engineering*, 3(4), 552–558.

Sánchez, K., Palma, S., & Castro, R. L. (2019). Numerical Modelling of Water Flow Through Granular Material for Isolated and Simultaneous Extractions in Block Caving. *Rock Mechanics and Rock Engineering*, 52(1), 133–147.

Sánchez, V., Castro, R. L., & Palma, S. (2019). Gravity flow characterization of fine granular material for Block Caving. *International Journal of Rock Mechanics and Mining Sciences*, 114(August 2017), 24–32.

Sereshki, F., Hoseini, S. M., & Ataei, M. (2016). Blast fragmentation analysis using image processing. *International Journal of Mining and Geo-Engineering*, 50(2), 211–218.

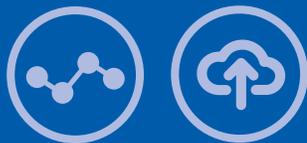
Vagenas, N., & Nuziale, T. (2001). Genetic algorithms for reliability assessment of mining equipment. *Journal of Quality in Maintenance Engineering*, 7(4), 302–311.

Wyllie, D. C., & Mah, C. W. (2017). *Rock slope engineering: Civil and mining, 4th edition. Rock Slope Engineering: Fourth Edition.*

Zhou, W., Qin, H., Qiu, J., Fan, H., Lai, J., Wang, K., & Wang, L. (2017). Building information modelling review with potential applications in tunnel engineering of China. *Royal Society Open Science*, 4(8).

# Big Data en el procesamiento de minerales y metalurgia extractiva

74



## Humberto Estay

AMTC, Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile

## Pia Lois

Departamento de Ingeniería de Minas,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile, y Julius Kruttschnitt Mineral  
Research Centre de la Universidad de Queensland

## Gonzalo Montes-Atenas

Minerals and Metals Characterization and Separation  
Research Group (M2CS), Departamento de Ingeniería de Minas,  
Facultad de Ciencias Físicas y Matemáticas  
Universidad de Chile





## 1. ¿Qué se hace hoy?

El control y la mantención de las plantas de procesamiento de minerales y metalurgia extractiva (conminución, flotación e hidrometalurgia) son, sin duda, unos de los puntos más críticos de la industria minera. Por ejemplo, solo la etapa de conminución puede representar hasta un 50% del consumo energético de una operación minera y, por lo tanto, gran parte del costo operacional (Ballantyne & Powell, 2014).

El objetivo principal del control del proceso es maximizar su eficiencia a un mínimo costo (Lynch, 2015; Spencer, 2014). En general, esto depende directamente de variables relevantes de cada proceso, como la determinación del tamaño objetivo de la molienda, el cual liberaría el mineral de interés; el consumo de ácido en lixiviación de cobre; el consumo de reactivos en flotación



Si bien, los datos que se generan a partir de la medición de instrumentos en línea y muestreos pueden acumular una gran cantidad de información, los procesos minero-metalúrgicos están lejos de que el Data Science permita la generación del volumen de datos de otras industrias, particularmente las de tecnología digital, tales como Google, Facebook o Amazon.



y también del tonelaje a procesar. Para el monitoreo de una planta, parámetros tales como la distribución del tamaño de partícula, el flujo, la ley del elemento de interés o la densidad, son parámetros importantes de cuantificar y determinan la instrumentación necesaria en la planta. Generalmente, estas medidas no se enfocan en tener una gran precisión, sino en obtener un gran volumen de datos, con una recurrencia que asegura una operación estable y continua.

Actualmente, se emplean medidores de flujo de tipo magnéticos, densímetros nucleares, analizadores de rayos X, controladores adaptativos, sensores de nivel en estanques,

manómetros en cañerías, entre otros, para el monitoreo y control de cada flujo, desde la conminución hasta la generación de relaves (Hodouin, 2011; Lynch, 2015). Estos sensores generan un volumen de datos de varios Gigabytes de información por minuto. Sin embargo, la complejidad de los procesos minero-metalúrgicos provoca que la medición de ciertas variables críticas no se pueda realizar con los instrumentos que existen actualmente, determinando la necesidad de medir a través de muestreos o, en algunos casos, la imposibilidad de medir ciertas variables. Algunas variables que no pueden ser medidas en línea sirven para determinar la eficiencia y productividad de una planta de



procesos, tales como la recuperación en flotación o lixiviación, el consumo de reactivos (ácido en el caso de lixiviación, espumantes y colectores en el caso de flotación) y la concentración de elementos en soluciones de lixiviación. Por otra parte, las características de la alimentación deben ser estudiadas mediante muestreos de planta, tales como la dureza, el grado de liberación, la ley de mineral y la mineralogía. En ambos casos, estos análisis son determinados por análisis de laboratorio, luego de muestreos que se hacen en períodos de horas o turnos. Generalmente, estos resultados sirven para cuantificar eficiencias de proceso o cuadrar balances metalúrgicos para la producción. Si bien, los datos que se generan a partir de la medición de instrumentos en línea y muestreos pueden acumular una gran cantidad de información, los procesos minero-metalúrgicos están lejos de que el *Data Science* permita la generación del volumen de datos de otras industrias, particularmente las de tecnología digital, tales como Google, Facebook o Amazon.

En las operaciones mineras existen diferencias importantes en el uso de los datos, determinado fundamentalmente por el estándar de instrumentación que tiene cada faena. La capacidad de generación de información y datos de medición y control en cada operación depende de los instrumentos instalados.

Asimismo, el uso de la información generada por medición continua (instrumentos) o discontinua (muestreo) tiene diferentes objetivos. La primera es en general utilizada para un control de proceso que asegure continuidad operacional, y no necesariamente para lograr condiciones óptimas de rendimiento.

En tanto, la medición discontinua es empleada para reportar productividad y tomar medidas de mejoras de proceso, lamentablemente de forma reactiva. Estas limitaciones no permiten hoy aplicar metodologías de predicción o control que optimicen el proceso de forma global.

Por ello es relevante avanzar en sistemas de medición aptos para los diferentes procesos o requerimientos y en sistemas de procesamiento de datos y predicción de resultados más robustos.

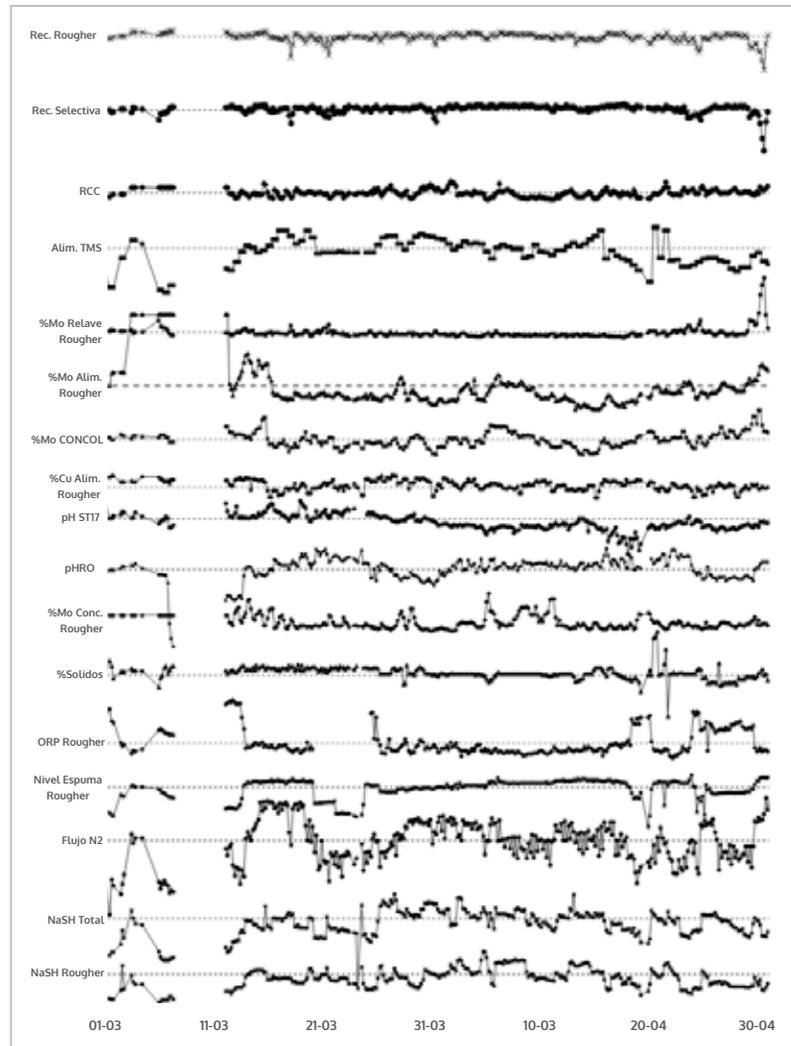
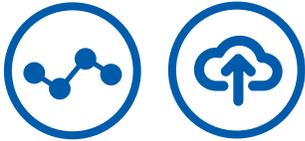
En la actualidad, la mayoría de las plantas posee un sistema de control (Process Intelligence - PI) que almacena la información y que ayuda a la automatización de la planta, la cual tiene por objetivo obtener información recurrente para tomar acciones efectivas que aseguren una operación estable (Sbárbaro & Villar, 2010). Sus funciones son: Asegurar la interacción sistema-operador, adquirir y procesar la información de la planta para mantener su funcionamiento, guardar la información de los equipos conectados y distribuirla a los computadores de los usuarios. Este sistema puede también estar asociado a un sistema de control experto para cada equipo, los cuales pueden ser integrados

en una misma plataforma, tales como Outotec ACT o ECS de FLSmith, las cuales utilizan algoritmos para establecer relaciones entre las variables provenientes de los distintos equipos y manejar la planta. OSIsoft PI System es uno de los más utilizados en la industria, el cual permite la colección automática y en tiempo real de datos de la planta, además de su análisis. Por otro lado, permite la interacción con el operador, reportando key performance Indicators (KPI values) y promoviendo la automatización de la planta (Marte, 2016). La Figura 1 muestra un ejemplo de datos almacenados por un PI System en un proceso de flotación, donde se muestran un subconjunto de las múltiples variables que se pueden extraer y analizar derivadas de este sistema de manejo de datos en plantas de procesamiento de minerales.

En las operaciones mineras existen diferencias importantes en el uso de los datos, determinadas fundamentalmente por el estándar de instrumentación que tiene cada faena.

**Figura 1.**

Ejemplo gráfico de las múltiples variables que se miden en un PI System (Ladrón de Guevara, 2016)



La excesiva limpieza de datos puede llevar fácilmente a estudiar escenarios sesgados basados en juicios erróneos.



Para cualquiera de los casos de información (datos continuos o discretos), la condición clásica es que los datos obtenidos son los disponibles y que es prácticamente imposible generar mejores de los ya generados en una ventana espacio-temporal (Loukides, 2012). Una típica rutina de análisis de datos en una planta de procesamiento de minerales o de metalurgia extractiva se muestra a continuación:

1. Una primera etapa de procedimientos simplificados que busquen limpiar los datos removiendo valores implausibles (*data conditioning*).

Esto no se debe confundir con la remoción de datos improbables, los cuales pueden, estar indicando anomalías reales dentro de una planta, por lo que la excesiva limpieza de datos puede llevar fácilmente

a estudiar escenarios sesgados basados en juicios erróneos. A lo anterior se añade un problema clave, relacionado con el lenguaje, el cual en muchos casos no es riguroso y la nomenclatura respecto de los efectos y las causalidades dentro de una planta es altamente variable.

2. Trazabilidad espacio-temporal y consistencia entre los resultados de diferentes mediciones.

Esto es incorporar en el análisis las diferencias temporales existentes entre las diversas mediciones, de tal forma de comprender el impacto de las variaciones de los parámetros y la causalidad que estos generan. Uno de los mayores problemas detrás de esto es que en una planta concentradora no se sabe de manera certera qué está ingresando efectivamente al circuito de flotación. Estudios geometalúrgicos (asumidos como proxies de operaciones) han probado ser un gran avance en esta área, reduciendo de manera significativa los riesgos detrás del funcionamiento de una operación (Harbort et al. 2017). Sin embargo, no se ha presentado como una solución a todos los problemas, y una buena correlación a todo evento -entre parámetros derivados de mediciones intrínsecas del mineral al ser procesado- no ha permitido generar modelos generalizados que predigan con exactitud la eficiencia del proceso. Interesantemente, la precisión es un aspecto que se ha logrado en faenas mineras, pero constituye, en términos generales, solo una segunda derivada del problema.



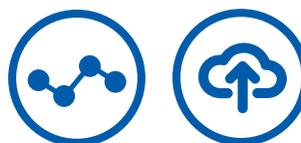


### 3. Análisis estadístico de los datos.

De acuerdo a *Data Science*, la definición de estadística es más general que el simple registro y análisis de datos, buscando una aproximación más bien holística, donde se considere el ciclo de vida de los datos. Esto es, de dónde vienen, cómo se usan y a dónde van. En el caso de la flotación es fácil indicar de dónde vienen (más allá de los desafíos que involucre la determinación de su representatividad y reproducibilidad). El cómo se utilizan habitualmente en la actualidad, debido a la falta de modelos predictivos, está relacionado más bien con una parametrización de los datos en términos de modelos multivariable y su relación con variables de eficiencia (recuperación y ley) (Gharai and Venugopal, 2015). El hacia dónde van esos datos es un desafío poco abordado, pues los resultados metalúrgicos, más allá de generar indicaciones de la producción,

no se transfieren entre unidades operativas presentes en la secuencia del negocio. La toma de decisiones local en base a información marcada (equivalente a lo que puede ser un *trending-topic*) es realizada de manera periódica por operadores y metalurgistas. El problema detrás de esto es que no es claro el momento en que estas decisiones deben ser implementadas.

En una faena minera, los datos generados internamente desde el *PI System* no incorporan tiempos de decisión. Para ejemplificar esto, una campaña de muestreo metalúrgico o *plant surveying*, actividad desarrollada desde los años 60s (Kelsall, 1961), requiere de una planta en estado estacionario, pues los modelos más utilizados solo abordan balances de masa en dicho estado. Al consultar sobre si la planta está en estado estacionario, solo se verifican algunos flujos como que los equipos no estén rebalsando de forma anómala, etc. Pero no existe un estándar que indique la estabilidad de la planta, información que varía de sitio en sitio y que en muchos casos su identificación es de carácter más bien cualitativa. Grandes esfuerzos se llevan a cabo por algunas faenas mineras al implementar unidades de integración de información o la puesta en marcha de cursos de capacitación que permitan a las diferentes áreas del negocio hablar un solo idioma y comprender la relevancia de la información de manera más transversal.



En el contexto de la lixiviación en pilas, en donde los tiempos de proceso son meses o años, se determinan áreas de pilas de decenas de hectáreas y con alturas que varían entre 3 a 70 metros, el tema de la información disponible se vuelve crítico. Estos mega-reactores, por su volumen, requieren una gran cantidad de sensores que permitan controlar el proceso y a su vez asegurar la buena transmisión de datos. Este alto requerimiento se conjuga con la falta de instrumentos propios de lixiviación en pilas, que permitan medir variables relevantes, como la humedad interna y la permeabilidad de la pila en sus diferentes zonas. En los últimos años se han propuesto diferentes métodos para cuantificar estas variables, como la tomografía de resistividad eléctrica (Hydrogeosense, 2014), la medición de la humedad a través de lectura infrarroja de imágenes usando drones (Ruíz del Solar et al., 2014), el monitoreo de condiciones de humectación mediante cámaras

Grandes esfuerzos se llevan a cabo por algunas faenas mineras al implementar unidades de integración de información o la puesta en marcha de cursos de capacitación que permitan a las diferentes áreas del negocio hablar un solo idioma y comprender la relevancia de la información de manera más transversal.



termográficas (Aplix, 2010) y la transmisión remota sin cables (Emerson, 2011). Tanto el control operacional como los modelos de producción propios de cada faena se basan en balances de masa estacionarios con ecuaciones empíricas determinadas del proceso en particular. Por lo tanto, la información generada es actualmente procesada con modelos simples, limitados y de base empírica, los cuales no contienen características de optimización o predicción.

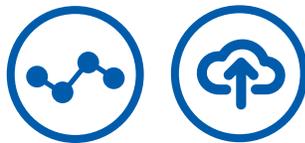
Sin embargo, existen variados modelos de lixiviación en pilas en ámbitos de consultoría e investigación (Marsden y Botz, 2017) que tienen como principal objetivo entender los fenómenos del proceso o generar estimaciones de producción a mediano y largo plazo. Hoy en día, no existen modelos de control operacional o predicción de producción de corto plazo en lixiviación en pilas.

En tanto, para conminución y flotación existen, actualmente, sistemas altamente predictivos y que permiten generar estrategias de control ante situaciones periódicas en la planta. Sin embargo, cambios en la roca y/o operacionales implicarían alteraciones en los objetivos del proceso, siendo la determinación de los límites de trabajo de cada equipo y sus fluctuaciones, de acuerdo al tipo de roca, densidades o carga circulante, más importante que el sistema de control en sí mismo (Hodouin, 2011). Debido a esto, la interpretación de los datos y el manejo del sistema de control por parte de los metalurgistas son de suma importancia para mantener el funcionamiento de una

El principal objetivo de estas metodologías es poder generar respuestas predictivas en la planta, a partir de la caracterización geológica y, por tanto, poder alimentar un modelo de planificación minera con información más certera.

planta. Los metalurgistas deben entender el proceso que está siendo controlado, la instrumentación de la planta y las medidas relevantes que pueden ser aplicadas para controlar el sistema. El entendimiento de los datos generados y almacenados en los sistemas de control es importante para tomar decisiones, ya que, por ejemplo, incrementar la velocidad de alimentación del molino podría aumentar la utilización de energía. Sin embargo, otros parámetros, como la densidad o la carga circulante pueden verse afectados, alterando el producto aguas abajo.

Por otra parte, la cantidad y uso de información es uno de los mayores desafíos que tiene el desarrollo de la geo-metalurgia o geo-minero-metalurgia. Estos conceptos buscan herramientas que puedan conciliar la

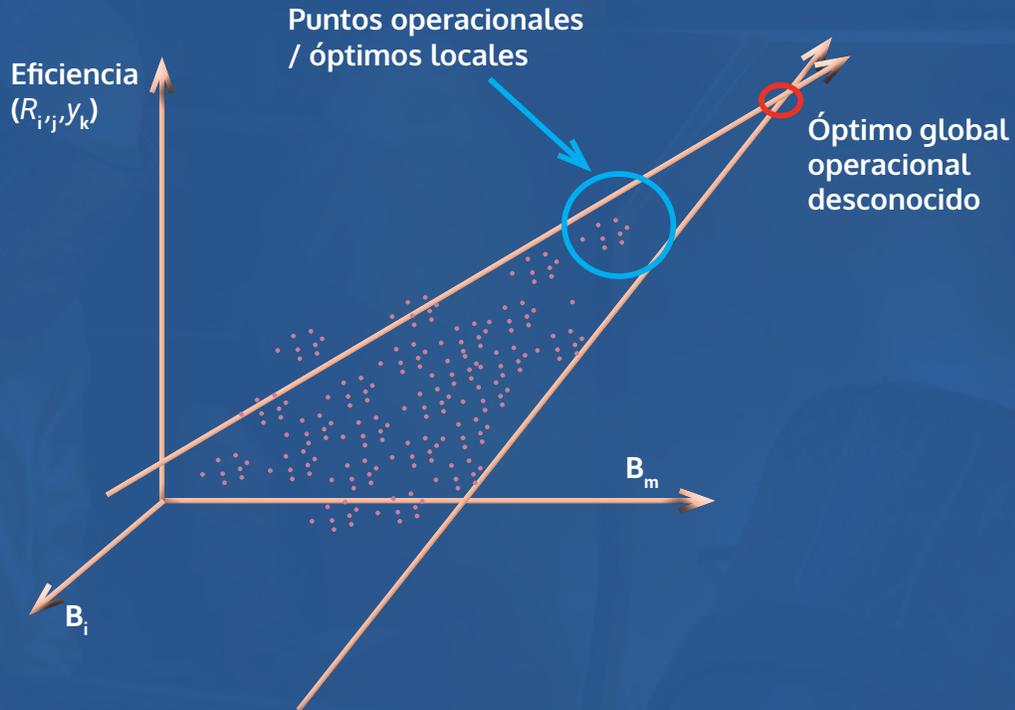


información geológica y su respuesta a nivel metalúrgico, basándose en que la geología debería tener mayor información de la materia prima usada en los procesos. El principal objetivo de estas metodologías es poder generar respuestas predictivas en la planta, a partir de la caracterización geológica y, por tanto, poder alimentar un modelo de planificación minera con información más certera. Sin embargo, para que se puedan generar herramientas predictivas, se debería contar con gran cantidad de información de pruebas metalúrgicas que se puedan asociar a la caracterización geológica. Por razones de costo, hasta hoy, esto se hace impracticable, provocando que la información geológica sea en general mucho mayor a la información metalúrgica. Para plantas operativas, se podrían implementar metodologías a partir de respuestas históricas de los minerales en la planta.

Finalmente, haciendo el paralelo con las industrias de tecnología, como Amazon y otras, las cuales utilizan intensivamente el uso de *big data*, guardando las búsquedas y correlacionándolas con aquellas de otros usuarios para crear recomendaciones sorprendentemente apropiadas, los cuales se establecen como *data products*.

**Figura 2.**

Esquema simplificado y de la descripción de los óptimos locales obtenidos en base a la experiencia de los operadores vs óptimo operacional "desconocido".



En efecto, las personas dejan un rastro de datos. Lo anterior se basa en que toda decisión se genera en base a un conjunto de datos de conocida o supuesta relevancia. Esta estrategia se puede aplicar entre turnos y/o días para una planta de procesamiento de minerales, sin embargo, estos sistemas son complejos y las variables medidas no son necesariamente aquellas que gobiernan el proceso. Aún si así fuese, el punto operacional de una correcta eficiencia no podría ser considerado un óptimo global, sino que se trataría de un óptimo local (no global o definitivo) que definiría una sugerencia de cómo hacer funcionar la planta de flotación en un determinado punto operacional (Ver Figura 2).

En resumen, los aspectos más relevantes del uso actual del *big data* en procesamiento de minerales y metalurgia extractiva son los siguientes:

- Las plantas actuales tienen diversas características para la medición, el registro y el control de variables, que dependen fundamentalmente del estándar y tamaño de la planta.
- Existen variables críticas en los procesos minerales, como la caracterización mineralógica, la distribución granulométrica, la determinación de ley en rocas y la determinación de elementos en soluciones, las cuales no pueden ser medidas de forma continua,

ya que no existen instrumentos que puedan realizar esto bajo las condiciones de complejidad de una planta real.

- La generación de datos continuos de algunas variables, como los flujos, la temperatura, la corriente o presión y los datos discretos, como los comentados en el punto anterior, para un mismo período de tiempo, determinan complicaciones inherentes a la aplicación de modelos predictivos.
- Por lo tanto, uno de los principales desafíos que existen en esta área para poder implementar el uso y los beneficios del *big data*, es el desarrollo de nuevos instrumentos o sensores para variables críticas del proceso.

## 2. ¿Qué está cambiando?

Tener equipamiento con sensores, obviamente, no es algo nuevo. Estos sistemas han evolucionado por más de treinta años, desde un sistema con mediciones básicas y controles analógicos e individuales de cada equipo, hacia un control en tiempo real que permite monitorear el proceso. Además, incluyen herramientas para analizar la información y caracterizar las señales para detectar fallos, y por qué no, anteponerse a ellos si la información es interpretada apropiadamente (Figura 3).

El avance en la implementación de los sistemas de control y la interpretación de los datos, junto con una mayor seguridad

Figura 3.

Avance de los sistemas de control en las últimas décadas.

Figura extraída de Sbárbaro and Villar (2010).



de la información, son pasos claves para la transformación digital en minería, y en particular, en el procesamiento de minerales. Además, la utilización de esta información permitirá a futuro tomar decisiones mejor informadas y mejorar la productividad (Hodouin, 2011).

No obstante, el punto más importante, y que permitirá el desarrollo y transformación del estado actual del procesamiento de minerales, es la interconectividad de los equipos. Esta interconectividad ha sido ampliamente referida en la literatura como Internet of Things (IoT) (Daugherty, Banerjee, Negm, & Alter, 2015; Zhang, Ma, Yang, Lv, & Liu, 2018). El IoT ha avanzado a lo largo del tiempo, debido principalmente al cambio en

El punto más importante, y que permitirá el desarrollo y transformación del estado actual del procesamiento de minerales, es la interconectividad de los equipos.



La transmisión de datos. Este ha evolucionado desde tecnologías Fieldbus, infrarrojo hacia bluetooth y tecnologías "Wireless". Este continuo avance en las tecnologías de transmisión determinará un aumento en la velocidad de respuesta de los equipos. La velocidad de transmisión de la información, ya sea del sensor al centro de almacenamiento y procesamiento o hacia las unidades de la planta, se verá aumentada, disminuyendo los tiempos de respuesta e incrementando la automatización de la planta y su productividad.

La reducción de leyes de cabeza y una mineralogía más compleja de los minerales, en conjunto con el avance de nuevas tecnologías, han motivado la búsqueda e implementación de nuevos desarrollos en términos de instrumentación y modelos



predictivos. Un ejemplo de ello es la incorporación de imágenes de alta resolución para caracterizar el material que entra a la planta. Ciertamente, uno de los principales problemas en el procesamiento de minerales es el desconocimiento de cómo el material se comportará en la concentradora. Sistemas como VisioRock (Guyot, Monredon, Larosa, & Broussaud, 2004) o WipFrag (Núñez, Silva, & Cipriano, 2011) permiten controlar cámaras ubicadas en la correa transportadora, además de conectarse al sistema de control para entregar información acerca del tamaño de alimentación e informar acerca de cambios en la roca. En esta misma línea, existen equipos comerciales que se pueden instalar en el proceso (generalmente correas transportadoras) para análisis elemental - espectroscopia de difracción de rayos X (XRF) y espectroscopia de ruptura inducida por láser (LIBS). Sin embargo, la presencia de estos equipos aún no es masiva en faenas mineras, quizás por costo o simplemente por no ser consideradas en las etapas de diseño.

Por otra parte, el análisis en línea de mineralogía aún está en fases de investigación. Al respecto, hay avances interesantes empleando análisis hiper-espectral (Ehrenfeld et al., 2017), pero aún falta desarrollo para su implementación en condiciones reales, donde el mineral está expuesto a movimiento, polvo, humedad, entre otros. Sin duda, el entendimiento de cómo la geología controla las propiedades de la roca es aún un tema de mucha investigación. En el caso de la cominución, uno de los principales motivos de que esta interpretación sea difícil, se debe a que no existe la utilización de pruebas que

La caracterización de ley de cabeza, ripios y mineralogía y concentración de especies en las soluciones, de forma continua y automática, es aún una de las grandes limitantes que existen hoy en hidrometalurgia.



desacoplen las propiedades de la roca del modo de quiebre. Siendo cada uno de ellos específico para un determinado equipo.

Por otra parte, la caracterización de ley de cabeza, ripios y mineralogía y concentración de especies en las soluciones, de forma continua y automática, es aún una de las grandes limitantes que existen hoy en hidrometalurgia, ya que el muestreo que se realiza no permite contar con información en línea que ayude a generar acciones de control predictivo o incluso reactivo en corto tiempo.

Particularmente, para lixiviación en pilas, se deben incorporar dos variables críticas para la productividad y seguridad de estas operaciones: permeabilidad y humedad en

los lechos. Por una parte, la permeabilidad permite definir el flujo de riego del proceso – y por tanto la velocidad de éste – y, por otra parte, la humedad permite establecer la estabilidad geotécnica de una pila y la eficiencia de humectación del lecho.

En tanto, la instrumentación en línea para el análisis de elementos en solución aún no se desarrolla con el nivel de confiabilidad necesario. En este sentido, los analizadores de electrodos de ion selectivo presentan un atractivo interesante de implementación, ya que a través de una medida potenciométrica se puede determinar el contenido de algún elemento en solución. Sin embargo, en la actualidad, estos electrodos pierden precisión en presencia de otros elementos existentes





en la solución. Este hecho limita de forma importante su aplicación en hidrometalurgia, donde las características de las soluciones presentan altos contenidos de diferentes iones.

Particularmente, para lixiviación en pilas, existen desarrollos para medir perfiles de humedad en el lecho, pero aún sin poder implementarse en línea y conectados a un sistema de control y registro. Los desarrollos en tomografía de resistividad eléctrica (Hydrogeosense, 2014) y medición de humedad a través de lectura infrarroja de imágenes usando drones (Ruíz del Solar et al., 2014) podrían ser empleados a futuro para controlar el riego en las pilas. Por otra parte, cabe destacar los desarrollos tecnológicos en la transmisión de variables remotas sin cables (Emerson, 2011). Dada la magnitud de las pilas de lixiviación, cualquier instalación de instrumentos cableados hace impracticable la posibilidad de instrumentar y controlar estas operaciones.

Respecto a los modelos predictivos, los desarrollos más interesantes en hidrometalurgia pertenecen a consultoras o empresas de ingeniería (Becker et al., 2012; Figueroa et al., 2015; Menacho, 2017), enfocados primordialmente a la estimación de producción a mediano y largo plazo. Las necesidades operacionales han motivado la complejización de estos modelos, para que puedan dar respuestas de apoyo a decisiones operacionales de mediano plazo, como el cambio de parámetros de operación de riego, carga de mineral, estimación de inventarios de

elementos en el pila o evaluación de impactos frente a cambios planificados en tasas de tratamiento. Paralelamente, hay grupos de investigación que han desarrollado modelos fenomenológicos para bio-lixiviación de minerales sulfurados en pilas (Petersen and Dixon, 2007; McBride et al., 2018). Estos modelos han sido escasamente aplicados en operaciones industriales, debido a su complejidad, pero mayormente por no incorporar condiciones operacionales propias, como la dinámica y variabilidad asociada a un plan de riego.

Así como existen modelos empíricos y fenomenológicos para lixiviación en pilas, para procesos de conminución y flotación existen experiencias donde los datos que se generan a partir de la instrumentación

y sistemas de control se transforman en información relevante para mejorar el proceso, mediante técnicas de Machine Learning, las cuales permiten establecer correlaciones y crear modelos entre grandes bases de datos. El machine learning se aplica habitualmente en cuatro pasos: desarrollo de ingeniería, entrenamiento del modelo, validación del modelo y selección y aplicación del modelo a nuevos datos (Harris, 2018). En estos casos, se requiere de un número abundante de datos de entrenamiento (training set). Algunos estudios demuestran su utilidad para anteponerse a cambios en el proceso.

Por ejemplo, Mudge, Roach, and Tickner (2017) modelaron el tonelaje de uno de los molinos de *Olympic Dam* usando variables operaciones del *PI System*, tales



como  $Fe/SiO_2$ , el porcentaje de llenado del molino, la densidad, la potencia del molino, entre otros. Utilizando un algoritmo de redes neuronales lograron una predicción bastante acertada del tonelaje de algunos meses de producción. Por su parte, P. Bardinas, Aldrich, and F. A. Napier (2018), realizaron un estudio similar para predecir el tamaño del producto de un circuito de molienda, usando información del proceso y aplicando algoritmos de Machine Learning. Estos trabajos son claros ejemplos de cómo toda esta cantidad de datos puede ser utilizada, generando nuevos estándares de trabajo en un sistema integrado computador-usuario.

En esta línea de avances, se pueden destacar esfuerzos que están haciendo grandes compañías mineras como Codelco, BHP

y Antofagasta Minerals, para incorporar procedimientos, metodologías y tecnología que permita estandarizar la obtención y uso de los datos, para desarrollar predicciones en diferentes ámbitos. En este sentido Codelco ha sido pionero en la implementación del control operacional de su planta de División Ministro Hales (DMH), ubicada cerca de Calama, desde sus oficinas en Santiago. En esta misma línea, BHP se encuentra actualmente implementando una iniciativa similar con su planta concentradora y de lixiviación en Escondida. Estas iniciativas no sólo están desafiando el uso de metodologías de *big data*, sino que además están separando los esfuerzos de diagnóstico y definición operacional, de las contingencias propias de mantener una planta operativa, evitando posibles sesgos y tomando definiciones con una visión global del problema, tanto a nivel de resultados, como a nivel de disciplinas. Otras acciones recientes, también por parte de Codelco, es el desarrollo de un contrato marco de implementación de *big data*, analítica e inteligencia artificial a nivel corporativo, para implementar diferentes desafíos divisionales.

Codelco ha sido pionero en la implementación del control operacional de su planta de División Ministro Hales (DMH), ubicada cerca de Calama, desde sus oficinas en Santiago.



En la misma línea, BHP y Antofagasta Minerals han definido propuestas para avanzar en la implementación de técnicas de *big data* y softwares para desarrollo de modelos de monitoreo de calidad de aire y mantención predictiva, respectivamente.

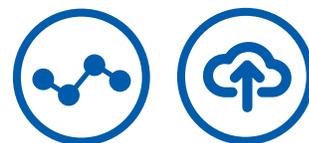
En resumen, los aspectos más relevantes de los avances que se están implementando para el uso del *bg data* en las áreas de procesamiento de minerales y metalurgia extractiva son los siguientes:

- Existen avances en el desarrollo de nuevas tecnologías de medición de variables en línea, particularmente para la caracterización de minerales, basadas en el análisis de rayos X, imágenes, hiper-espectral o láser. Asimismo, hay avances en caracterización de lechos de pilas de lixiviación.
- El avance en el desarrollo de modelos empíricos, fenomenológicos o basados en Machine Learning para los diferentes procesos u operaciones unitarias está permitiendo una mejor comprensión y toma de decisiones operacionales, aunque aún falta desarrollo para que puedan ser modelos predictivos o de respuesta en tiempo real.

### 3. Visión: Ideas y propuestas

La generación y uso de *big data* es aún limitada, pero en auge en procesamiento de minerales y metalurgia extractiva. Nuevos conceptos y estrategias que intentan incorporar diferentes aspectos del negocio, tales como la "Geo-Minero-Metalurgia", originalmente presentado a nivel mundial por

el Departamento de Ingeniería de Minas de la Universidad de Chile, busca incorporar las restricciones necesarias para evaluar el valor del negocio minero de forma más precisa en el tiempo. Pero esta integración es aún un desafío para muchas empresas mineras, ya que requiere de un manejo global, capaz de generar una visión a largo plazo, más allá de las metas generales de cada unidad del negocio minero, y de crear valor a partir de los datos. Por otro lado, este desafío tiene subyacente otros aspectos cruciales para el éxito de una operación, tampoco bien definido aún, como lo son: ¿qué variable muestrear? Y, ¿cómo muestrearlas? Incluso, en muchas operaciones existen variables que, luego de muchos años, se siguen midiendo, por un tema más bien histórico que por una necesidad real, las cuales muchas veces están asociadas a modelos empíricos que no sostienen una explicación fenomenológica de los procesos. Estas interrogantes podrían resolverse al generar una cuantificación del efecto de la geología sobre los procesos de una manera consistente y que permita explicar sus variaciones. Como se ha descrito anteriormente, el avance tecnológico reciente en la caracterización de minerales y la integración del *big data* podrían generar el conocimiento necesario para el real desarrollo de la geo-metalurgia.





El avance tecnológico reciente en la caracterización de minerales y la integración del *big data* podrían generar el conocimiento necesario para el real desarrollo de la geo-metalurgia.

95

De acuerdo a lo anteriormente descrito, existen variables que generan una cantidad de datos - algunos relevantes y otros no tanto - pero que son usados para el registro o control automático de equipos y accesorios menores (bombas, etc.), lo que garantiza continuidad operacional, sin apuntar a la maximización de eficiencia de los procesos. Esto se debe, en parte, a la falta de instrumentos en línea o de medición continua de variables críticas, como mineralogía y particularmente, en la caracterización de soluciones acuosas. Por otra parte, la ausencia de modelos que puedan usar estos datos y generen respuestas en corto plazo de manera confiable, limitan el uso del *big data* que se pueda generar.

Una forma clara de disminuir costos operacionales es, sin duda, implementar nuevas tecnologías para la utilización del *big data*, conectado por sensores, muestreos y técnicas analíticas. Las velocidades de conexión entre equipos pueden generar estructuras más conectadas, favoreciendo la operación de estos equipos como una sola unidad (Daugherty et al., 2015). Cada equipo podría tomar decisiones para modificar las variables sin la interacción de humanos, haciendo los procesos más rápidos.

Uno de los puntos que puede verse más beneficiado es el manejo de riesgos y amenazas, generando una mantención proactiva y evitando el número de errores fatales que bajan la producción. Estos algoritmos, capaces de aprender patrones, serían capaces de detectar cuándo un equipo necesita

reparación pronta, disminuyendo así los costos. De acuerdo a Daugherty et al. (2015), las reparaciones de mantención utilizando modelos predictivos de Machine Learning reducirían en un 30% los costos totales de reparaciones, sobre todo aquellos asociados a cuando los equipos dejan de funcionar. En esta línea, la visión de avanzar desde la detección y diagnóstico, hasta el pronóstico de fallas o rendimientos de procesos es uno de los principales desafíos de la industria de procesos en la era 4.0 (Reis & Gins, 2017).

Sin embargo, la enorme cantidad de datos puede, efectivamente, ser un problema en el caso del procesamiento de minerales y la metalurgia extractiva. En muchos casos, la utilidad de los datos es aún una interrogante, por lo que el manejo de estos no es aún parte del problema.

Se requiere una mejor identificación de qué medir y más y mejores mediciones, en conjunto con un establecimiento de requerimientos temporales de medición (cada cuánto tiempo se necesita medir). En este contexto, Machine Learning nuevamente se presenta como una opción válida, a pesar de que el cuerpo de datos de entrenamiento sería masivo. Sin embargo, y lamentablemente, dichos datos de entrenamiento, en principio, solo obedecerían a óptimos locales, derivados de las destrezas operacionales de operadores y metalurgistas.

La construcción de gráficos de dispersión para realizar estudios de conexión entre parámetros y la revisión de datos originales es imprescindible como primer paso hacia la implementación de *big data*.

En conclusión, hay desafíos concretos que deben ser resueltos para que se implemente el uso de *big data*, con el objetivo de mejorar la eficiencia operacional:

1. Desarrollar sensores confiables para la medición de mineralogía y la caracterización de elementos en línea (minerales, pulpas y soluciones).
2. Implementar y masificar elementos de *Internet of Things (IoT)* o Industria 4.0, como la transmisión de datos *wireless*.
3. Estandarizar datos e interoperabilidad en los modelos desarrollados, es decir, estos deben ser capaces de interactuar con la información generada por los instrumentos de la planta y responder en tiempo limitado, para poder ser integrados en el sistema de control de una planta.
4. Implementar la medición continua de variables críticas de los procesos que actualmente no existen, como la medición de la dureza de minerales, la granulometría, la humedad y la permeabilidad en pilas, el consumo de ácido de minerales en pilas de lixiviación, los requerimientos de colectores y espumantes de acuerdo al mineral en flotación, entre otros. Esta medición continua servirá de base para poder conectarse al sistema de control central de una planta.
5. Mejorar los modelos de los diferentes procesos involucrados (o desarrollar nuevos) en una faena minera (chancado, molienda,



flotación y lixiviación en pilas), con el objetivo de aumentar la confiabilidad y la predictibilidad de los resultados para condiciones operacionales o minerales diferentes a las que haya tratado una planta, incluyendo el manejo de datos con alta variabilidad, como los que se presentan en los procesos reales. Estos modelos pueden ser de base fenomenológica, semi-empírica o estocástica.

- 6. Mejorar las metodologías de geo-metalurgia, para permitir la conexión fluida de la información entre la geología y la planta de proceso y su impacto en los modelos de planificación minera.

Por último, esta revolución digital en minería traerá consigo un cambio obligado en la fuerza de trabajo, donde especialidades asociadas al análisis de datos, al desarrollo de software y al diseño digital, serán de suma importancia para probar todos los nuevos desarrollos, interpretar la gran cantidad de datos asociada y trabajar en un sistema con una infraestructura para el progreso del IoT. Por ello, la formación de capital humano que cuente con habilidades en estas áreas será fundamental para poder superar los desafíos que se impondrán a futuro.

En este sentido, en Chile tenemos una oportunidad única para ser líderes mundiales en el desarrollo de estas nuevas tecnologías y el capital humano asociado a estas, ya que se cuenta con las plantas de procesos mineros más grandes del mundo, las que a su vez presentan las mayores complejidades, debido a la explotación de minerales de menor ley y caracterización mineralógica variable, debido a la profundización de los yacimientos.

# Referencias

Hydrogeosense, 2014. Monitoring HGS. Obtenido de <http://hydrogeosense.com/services/monitoring/>

Aplik, 2010. SISTEMA DE VISIÓN ARTIFICIAL PARA EL MONITOREO CONTINUO DE PILAS DE LIXIVIACIÓN. Obtenido de <http://www.aplik.cl/apliweb/wp-content/uploads/tecnologiaMonitoreoPilas.pdf>

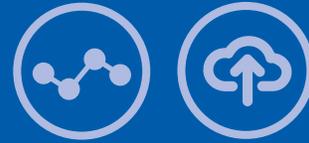
Emerson, 2011. Managing Cyanide Streams with Wireless Analytical Measurements. Obtenido de <http://www.emerson.com/documents/automation/application-data-managing-cyanide-streams-wireless-analytical-measurements-en-70776.pdf>

Ballantyne, G. R., & Powell, M. S. (2014). Benchmarking comminution energy consumption for the processing of copper and gold ores. *Minerals Engineering*, 65, 109-114. doi:10.1016/j.mineng.2014.05.017

P. Bardinas, J., Aldrich, C., & F. A. Napier, L. (2018). Predicting the Operating States of Grinding Circuits by Use of Recurrence Texture Analysis of Time Series Data. *Processes*, 6(2), 17. doi:10.3390/pr6020017

J. Becker, A. Figueroa, P. Silva, F. Arriagada, *Dynamic mass balance simulations in multi-lift leaching heaps*, in: *Hydroprocess 2012*, Gecamin, Santiago, 2012.

Daugherty, P., Banerjee, P., Negm, W., & Alter, A. (2015). Driving Unconventional Growth through the Industrial Internet of Things. In. [www.accenture.com](http://www.accenture.com): Accenture.



Ehrenfeld, A., Egaña, Á., Guerrero, P., Liberman, S., Hanna, V., Voisin, L., & Adams, M. (2017). Geometallurgical Variables Characterization Using Hyperspectral Images and Machine Learning Technics. In 38th Application of Computers and Operations Research in the Mineral Industry (pp. 61–66)

Figueroa, P. Silva, F. Arriagada, R. Peralta, Copper Production Model for San Francisco Dump Leaching Operation, in: Hydroprocess 2015, Gecamin, Santiago, 2015.

Guyot, O., Monredon, T., Larosa, D., & Broussaud, A. (2004). VisioRock, an integrated vision technology for advanced control of comminution circuits. Minerals Engineering, 17(11), 1227-1235. doi:10.1016/j.mineng.2004.05.017

Harbort, Gregory & Jones, Kelly & Morgan, Dylan & Sola, Christine. (2017). Integrating geometallurgy with copper concentrator design and operation. In the book: We are Metallurgists, not Magicians, p. 37-54.

Harris S. (2018) Data Science from Scratch: Intermediate guide for learning Data science and master it like a pro (Comprehensive guide). 89 p.

Hodouin, D. (2011). Methods for automatic control, observation, and optimization in mineral processing plants. Journal of Process Control, 21(2), 211-225. doi:10.1016/j.jprocont.2010.10.016

Kelsall D.F. (1961), Application of of probability in the Assessment of flotation Systems, Trans IMM, 70, 191-204.

[Ladrón de Guevara, R. \(2016\)](#) Análisis estadístico y experimental de parámetros involucrados en la flotación selectiva de molibdeno en minera los pelambres, Memoria de Ingeniero Civil de Minas, Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, 142 p.

[Loukides, M. \(2012\)](#) What is data Science? The future belongs to the companies and people that turn data into products. 27 p.

[Lynch, A. J. \(2015\)](#). Comminution handbook (Vol. Spectrum series / Australasian Institute of Mining and Metallurgy). Carlton, Victoria: Australasian Institute of Mining and Metallurgy.

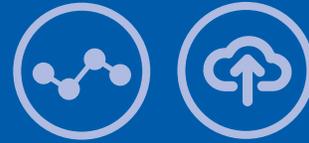
[J.O. Marsden, M.M. Botz](#), Heap leach modeling: A review of approaches to metal production forecasting, *Miner. Metall. Process.* 34 (2017) 53–64. doi:10.19150/mmp.7505.

[Marte, L. \(2016\)](#). Operational Excellence with the PI System at Barrick Gold. Paper presented at the OSIsoft Users Conference, 2016, San Francisco.

[D. McBride, J. Gebhardt, N. Croft, M. Cross](#), Heap leaching: Modelling and forecasting using CFD technology, *Minerals*. 8 (2018). doi:10.3390/min8010009.

[J. Menacho](#), Scale Up of Hydrodynamic and Metallurgical Response in Heap Leaching Operations, in: *Hydroprocess 2017*, Gecamin, Santiago, 2017.

[Mousumi Gharai & R. Venugopal \(2015\)](#): Modeling of flotation process – an overview of different approaches, *Mineral Processing and Extractive Metallurgy Review*, 57 p.



Mudge, C., Roach, G., & Tickner, J. (2017). *Big Data* improving the productivity of minerals processing (Presentation). (ep173774). from CSIRO

Núñez, F., Silva, D., & Cipriano, A. (2011). Characterization and Modeling of Semi-Autogenous Mill Performance Under Ore Size Distribution Disturbances. *IFAC Proceedings Volumes*, 44(1), 9941-9946. doi:<https://doi.org/10.3182/20110828-6-IT-1002.02199>

J. Petersen, D. Dixon, Modeling and optimization of Heap Bioleach Processes, in: D.E. Rawlings, D.W. Johnson (Eds.), *Biomining*, Springer, NY, 2007: pp. 153–175.

Ruíz del Solar, J., Vallejos, P., Asenjo, R., Daud, O., Correa, M. Método de detección de zonas de saturación, mediante plataforma aérea no tripulada, equipada con cámara térmica infrarroja, para la gestión y el control de riego en pilas de lixiviación. WO2015157874A1 Application, 2014.

Reis, M., Gins, G., 2017. Industrial Process Monitoring in the *Big Data*/Industry 4.0 Era: From Detection, to Diagnosis, to Prognosis. *Processes* 5, 35. doi:10.3390/pr5030035

Runge, K C, Tabosa, E, Crosbie, R and McMaster, J K, 2012. Effect of flotation feed density on the operation of a flotation cell, in *Proceedings 11th AusIMM Mill Operators' Conference*, pp 171-178 (The Australasian Institute of Mining and Metallurgy: Melbourne).

Sbárbaro, D., & Villar, R. d. (2010). *Advanced control and supervision of mineral processing plants*: London

Spencer, S. (2014). *Monitoring of Mineral Processing to Improve Productivity* (Presentation). from CSIRO

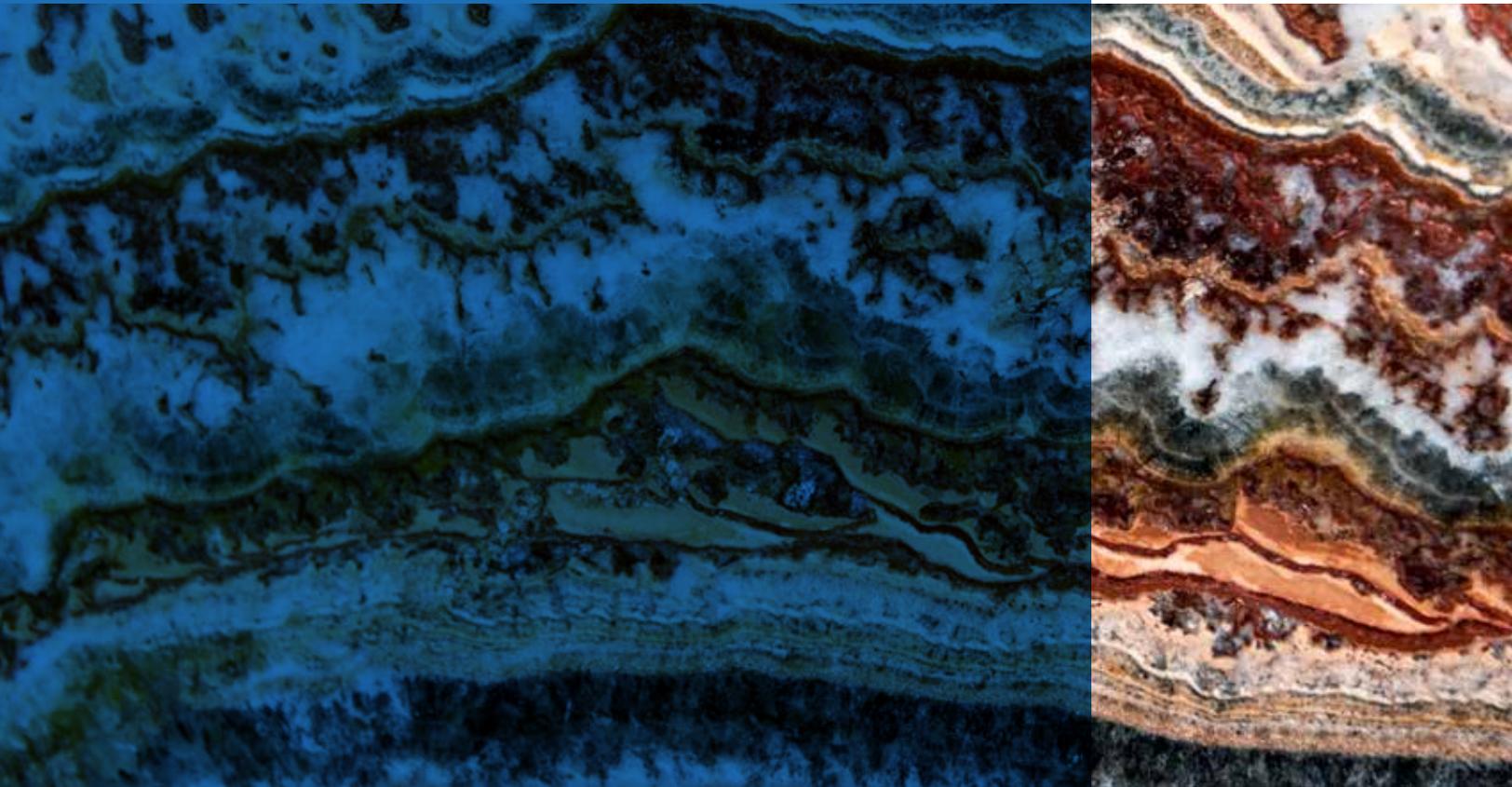
# Casos aplicados: *Big Data* en geometalurgia

**Julián M. Ortiz**

The Robert M. Buchan  
Department of Mining Queen's  
University, Canada

**Willy Kracht**

Departamento de Ingeniería de Minas  
Universidad de Chile  
Advanced Mining Technology Center  
Universidad de Chile



## 1. Introducción

La geometalurgia combina información de la geología y de los procesos mineros y metalúrgicos para crear modelos espaciales predictivos para procesamiento de minerales y metalurgia que pueden ser usados para optimizar el plan minero, el diseño y los procesos, considerando restricciones claves del proyecto, relacionadas con el medio ambiente, así como con el uso eficiente y disponibilidad de agua y energía (Ortiz et al, 2015). Para esto, es necesario incorporar el contexto espacial de la información disponible y su naturaleza geológica, para la construcción de modelos predictivos que permitan estimar y predecir el desempeño de distintos procesos a lo largo de la cadena de valor minera (Boisvert et al, 2013). El principal desafío de este enfoque es integrar el conocimiento experto de las geociencias con los modelos estadísticos utilizados. *Machine learning* y *deep learning* pueden ayudar con algunos de estos problemas, mediante herramientas de *knowledge discovery* desde la abundante información disponible en proyectos mineros (desde la exploración a la producción y la operación), así como mediante predicción de comportamientos de mezclas de materiales al someterlos a distintos procesos, mezclas de materiales de orígenes inciertos en la mina (producción de distintas fases o sectores, mezclada en un *stockpile* y/o remanejo de materiales), procesamiento de minerales y procesos metalúrgicos complejos, y retroalimentación de datos de operación. Incorporar la variabilidad y por lo tanto entender la incertidumbre asociada a los valores estimados o predichos, resulta fundamental para poder realizar un análisis de riesgo apropiado.



Un esquema general del proceso se muestra en la Figura 1.

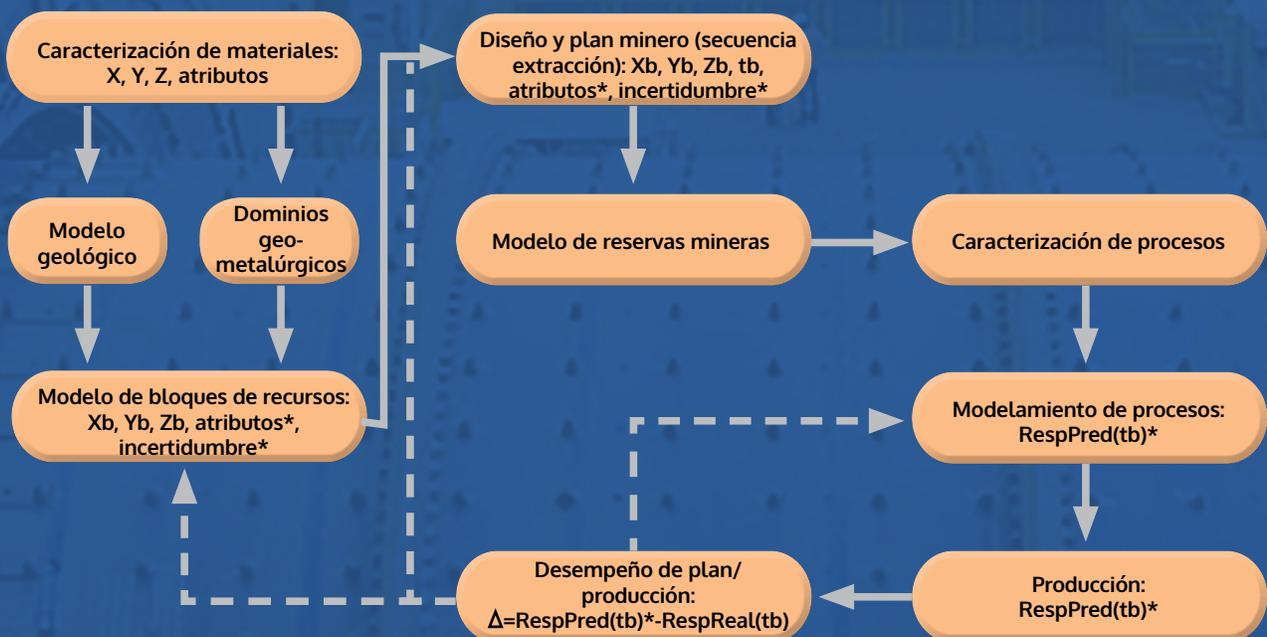
Las características específicas de los materiales pueden muestrearse en ciertas posiciones del espacio (X, Y, Z), entregando valores reales de la propiedad analizada (asumiendo los errores de muestreo pueden despreciarse)<sup>1</sup>. Se puede entonces combinar el conocimiento experto con la información de estos distintos ensayos que caracterizan el material, los que luego son usados para construir el modelo *in situ* (modelo geológico, dominios geometalúrgicos y modelo de recursos geológicos).

Este modelo *in situ* entrega una aproximación de los valores reales de los atributos, habitualmente en una grilla regular conocida como modelo de bloques donde los valores son estimados y su incertidumbre puede evaluarse mediante simulaciones estocásticas (cada bloque con coordenadas  $X_b$ ,  $Y_b$ ,  $Z_b$ ). Este modelo de bloques alimenta el proceso de planificación minera, donde se realizan optimizaciones para determinar el tiempo ( $t_b$ ) en el que cada bloque es extraído y enviado a distintos flujos de material, dependiendo de las propiedades estimadas en el bloque (estéril, stockpiles, procesos).

<sup>1</sup> En rigor, los valores reales nunca son conocidos sino estimados (mediciones). Si los errores de muestreo se desprecian, entonces los valores medidos tienden a los valores reales.

**Figura 1.**

Flujo de trabajo general para el modelamiento geometalúrgico.





La inherente variabilidad en la geología y la mineralogía de un depósito mineral impone un cierto grado de incertidumbre en la respuesta esperada de cada proceso.

Las distintas respuestas (estabilidad del lastre desde el punto de vista de su potencial drenaje ácido, recuperación de elementos o metales de interés, consumo de agua y energía durante el procesamiento, características de los relaves, etc.) dependerán de las propiedades reales de los bloques y de los procesos involucrados. Por lo tanto, es necesario disponer también de modelos de los procesos. Finalmente, el desempeño predicho ( $RespPred(tb)^*$ ) y real ( $RespReal(tb)^*$ ) pueden compararse y retroalimentar los modelos para actualizarlos/mejorarlos.

La inherente variabilidad en la geología y la mineralogía de un depósito mineral impone un cierto grado de incertidumbre en la respuesta esperada de cada proceso. Usando un enfoque geometalúrgico, se puede reducir esta incertidumbre, lo que permite operar bajo un riesgo técnico y operacional reducido.

Dado que tanto los aspectos geológicos como los mineros y metalúrgicos tienen impactos en las fases de exploración y explotación minera, el modelamiento geometalúrgico se puede aplicar en todas las etapas de un proyecto minero. Esto es, puede aplicarse en las etapas tempranas de ingeniería de perfil, durante el estudio de factibilidad, así como también en las etapas de desarrollo y operación y durante el cierre de mina.

La construcción de un modelo predictivo geometalúrgico requiere completar una secuencia de tareas, desde la adquisición de datos, modelamiento espacial, hasta el modelamiento de procesos (Boisvert et al, 2013; Dominy et al, 2018). Estas tareas pueden ser fácilmente interpretadas como etapas en un flujo de trabajo. Un ejemplo típico se muestra a continuación:

1. Adquisición de datos: análisis geoquímicos, mapeos geológicos, análisis químicos de elementos principales e impurezas, información hiperspectral.
2. Ensayos metalúrgicos: una batería de ensayos de laboratorio pueden ser realizados sobre muestras representativas de los principales dominios (o unidades) geológicos del depósito, de forma de conocer la respuesta a cierto proceso.
3. Caracterizar las unidades o dominios geometalúrgicos mediante *clustering* de muestras con similar desempeño en el proceso.
4. Construir un modelo espacial de las unidades geometalúrgicas y de los atributos de interés usando técnicas geoestadísticas convencionales.
5. Inferir la respuesta al proceso basado en las características locales de cada bloque de material.

Este flujo de trabajo, que en principio aparece lineal, no puede realmente aplicarse de esta forma secuencial. Requiere múltiples iteraciones para mejorarlo y llegar a un modelo predictivo adecuado, incluyendo pasos como selección de variables, agregación de variables, *clustering* o clasificación, escalamiento, modelamiento de correlaciones entre múltiples variables, modelamiento de procesos, etc.

En las próximas secciones, se muestran algunas aplicaciones donde se enfatizan las herramientas estadísticas requeridas y la integración del conocimiento experto geocientífico en los procesos.

## 2. Casos aplicados

**Aplicación 1: modelamiento de tipos de alteración como un proxy para el comportamiento del mineral en flotación.**



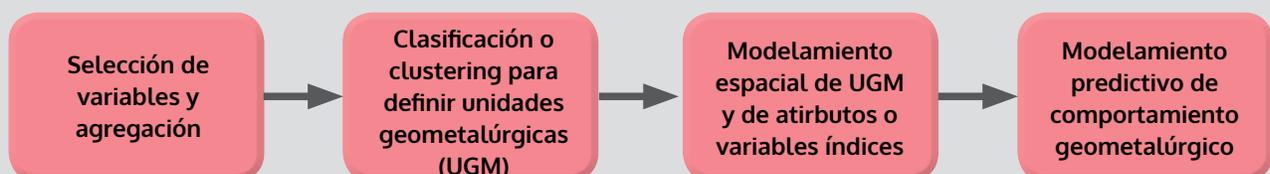
Existe un vínculo entre los tipos de alteración y el desempeño en flotación en depósitos porfíricos de cobre (Sillitoe, 2010; Bulatovic et al. 1999). Los tipos de alteración son identificados por el equipo de geología durante la etapa de mapeo de los testigos de sondajes o de los *cuttings* de aire reverso. Este proceso se presta a errores de interpretación de los tipos de alteración, debido a que ciertas características de la alteración pueden haberse visto afectadas por la súper-imposición de otros procesos, o por oxidación superficial en el sondaje. Por otro lado, los tipos de alteración puros son referenciales para el geólogo de mapeo, por lo que la mayor parte de las muestras presenta algún nivel de mezcla de tipos de alteración, y se etiqueta según sus características predominantes.

En esta aplicación se propone que las características específicas de cada tipo de al-

teración pueden ser modeladas mediante la geoquímica de las muestras (Miranda, 2015). Los tipos de alteración mapeados pueden ser usados como valores referenciales para entrenar un modelo predictivo. De esta manera, se puede buscar aquellos elementos que permitan una máxima discriminación entre los distintos tipos de alteración, para construir un modelo estadístico predictivo, usando herramientas de *machine learning*. Estos tipos de alteración pueden entonces utilizarse para poblar el modelo de bloques y predecir el desempeño en flotación, una vez que los atributos geoquímicos han sido estimados o simulados en el modelo. Los pasos para construir este modelo se muestran en la Figura 2. En esta aplicación, se presentan algunos detalles de los dos primeros pasos de este flujo de trabajo (selección de variables y modelo de clasificación para definir los tipos de alteración).

**Figura 2**

Flujo de trabajo para la predicción de unidades de alteración como proxy de desempeño de flotación.



El análisis se desarrolla con una base de datos de muestras de sondaje con información de geoquímica de muestras (disolución en agua regia) – 39 variables, análisis de leyes de interés económico – 3 variables, y el mapeo del tipo de alteración – 1 variable agrupada en 5 categorías principales. A esto, se agregan 11 variables sintéticas creadas para destacar ciertas características de los tipos de alteración, tal como se explica más adelante. La base de datos completa consta de más de 13600 muestras.

Previo a la construcción del modelo, la información disponible se divide en tres grupos para poder generar un modelo que evite el *overfitting* o sobreajuste a los datos.

1. **Entrenamiento:** Para la confección de los modelos de clasificación se seleccionó un conjunto de sondajes aleatoriamente con el fin de completar al menos 80 muestras de cada tipo de alteración. Este conjunto de datos representa un 4% de las muestras totales disponibles.
2. **Ajuste:** Compuesta por un 17% de los datos tomados al azar de la base de datos. Este conjunto de datos es utilizado para estudiar el error que produce el modelo, buscando que no exista sobreajuste.
3. **Validación:** El 79% de información restante permite entender cómo el desempeño del modelo generaliza a datos nuevos. Con este conjunto se determina el porcentaje de acierto final de cada modelo predictivo analizado.

Algunas de las técnicas estadísticas requeridas para construir este flujo de trabajo, se describen en más detalle a continuación:

- **Selección de variables y agregación,** basado en la correlación con la respuesta. El primer paso en el proceso es conocer qué variables tienen las correlaciones más fuertes con la respuesta, en este caso, el tipo de alteración. Esto se hace mediante diversos métodos de *data mining*.

En esta aplicación, se utilizó una selección *forward* modificada (Chen et al. 2004), donde la combinación de variables más relevantes para predecir la respuesta es seleccionada de manera iterativa. Además de este procedimiento, se puede realizar agregación de variables, usando la comprensión de la geología en este paso (Babak and Deutsch, 2009; Mariethoz et al, 2009). Por ejemplo, variables sintéticas que destacan ciertas características de cada tipo de alteración pueden ser agregadas al análisis:  $K^*Al$  sirve para destacar la presencia de sericita-muscovita en la alteración fílica;  $Al/Mg$  destaca las arcillas ricas en aluminio presentes en la alteración argílica, sobre las arcillas ricas en magnesio que son dominantes en otros tipos de alteración;  $K/(Ca+Na)$  destaca el intercambio de cationes de K sobre los de Ca y Na en la alteración potásica.

- **Clasificación para definir las unidades geometalúrgicas.** Utilizando las variables geoquímicas seleccionadas y las variables sintéticas agregadas, se realiza una clasificación para maximizar el grado de coincidencia con la información disponible en el mapeo geológico, entendiendo que esta información no es totalmente consistente, por lo que el procedimiento permite identificar zonas donde el mapeo puede requerir una revisión. Distintos métodos de clasificación pueden utilizarse para etiquetar los tipos de alteración, basados en la información geoquímica.

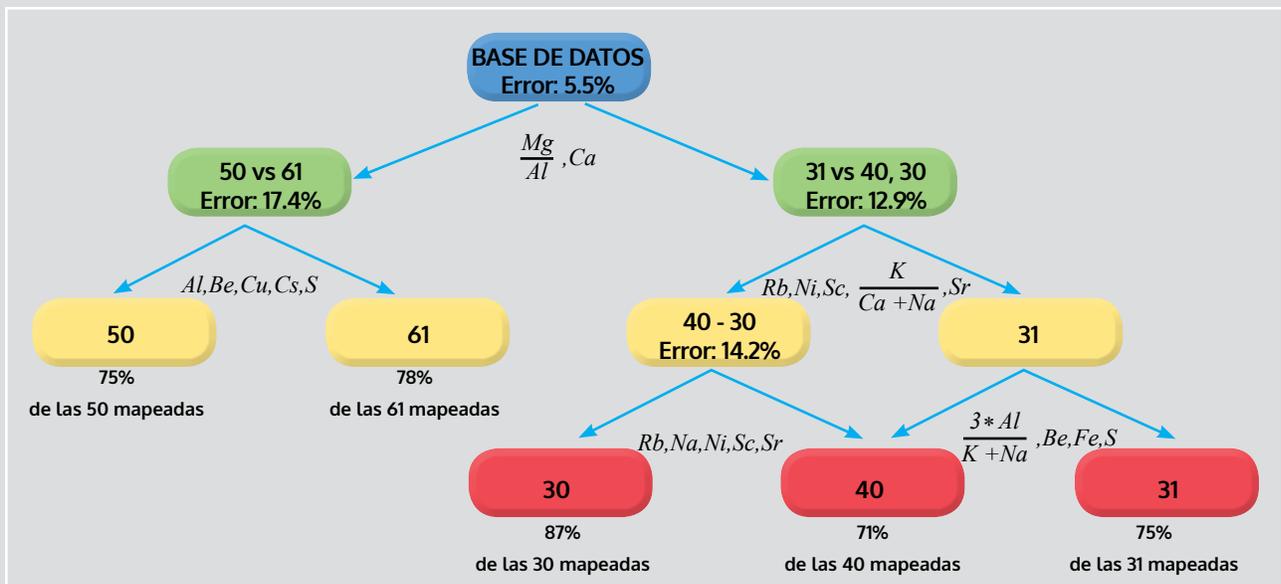
En esta aplicación se utilizan criterios basados en máxima discriminación utilizando una sola variable, *k-means clustering*, regresión logística y redes neuronales artificiales. Mayores detalles de estas y otras técnicas pueden encontrarse en cualquier

libro de *machine learning o data science* (Grus, 2015). Estas técnicas permiten construir un árbol de clasificación para determinar la etiqueta de cada muestra (Figura 3). El mejor método puede seleccionarse usando la curva ROC (Fawcett, 2004), que permite comparar el desempeño de los distintos métodos. Además, este análisis puede complementarse usando funciones de membresía para describir la incertidumbre asociada al etiquetado de los tipos de alteración (Pimentel and de Souza, 2013; Cardenas, 2015). Otros métodos existen



**Figura 3:**

Árbol de clasificación final modelo predictivo de tipos de alteración basado en geoquímica.





para realizar este procedimiento de clasificación (Sepúlveda and Dowd, 2018). El resultado del modelo predictivo puede verse en detalle en la matriz de confusión que muestra el porcentaje de acierto en la clasificación al comparar la predicción con los tipos de alteración mapeados en la base de datos de validación (Tabla 1).

Diferencias entre los tipos de alteración obtenidos en la predicción y los mapeados inicialmente son un indicador de inconsistencia durante el proceso de mapeo, o bien, una señal de variabilidad local. Por lo tanto, estas áreas pueden ser consideradas para una campaña de remapeos, para mejorar el modelo final de

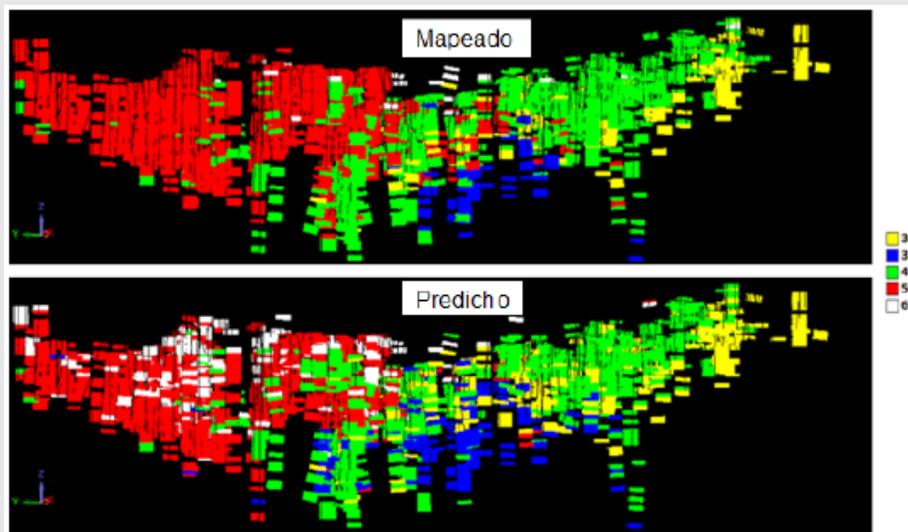
**Tabla 1**

Matriz de confusión del modelo predictivo final de tipos de alteración, basado en geoquímica de muestras.

	Alt. Predicha	30	31	40	50	61
Alt. Mapeada	N° de Datos	578	358	1801	3471	902
30	261	87%	3%	10%	0%	0%
31	115	9%	75%	13%	4%	0%
40	2169	15%	7%	71%	5%	2%
50	4474	0%	3%	5%	75%	18%
61	91	0%	0%	8%	14%	78%
<b>Acierto Promedio:</b>		<b>77.1%</b>				

**Figura 4**

Sección representativa con el mapeo de tipos de alteración original (arriba) y tipo de alteración definido por el modelo predictivo (abajo).



alteraciones y su capacidad predictiva relacionada con el proceso de flotación (Figura 4).

Además, el uso de funciones de membresía permite determinar porcentajes de pertenencia que pueden ser un indicador de mezclas que se alimentan al proceso de flotación, a ser considerados durante el control de procesos en la planta de flotación.

Esta aplicación muestra cómo se puede construir sin mayor dificultad y alcanzando un porcentaje de acierto alto, un modelo predictivo de los tipos de alteración a partir de la información geoquímica disponible en cada muestra. El modelo predictivo generado per-

mite determinar de manera cuantitativa los tipos de alteración a partir de un entrenamiento con una fracción muy pequeña de la base de datos total, donde un mapeo cualitativo ha sido definido por un equipo de geólogos de mapeo. Las muestras con su tipo de alteración definido mediante la aplicación de este modelo predictivo, pueden ser usadas en etapas posteriores para definir los dominios geometalúrgicos y asignar a cada bloque del modelo un tipo de alteración, que permita anticipar su comportamiento en la planta, en particular en relación al proceso de flotación, donde la alteración afecta la recuperación de los elementos de interés.



### **Aplicación 2: modelamiento de zonas minerales utilizando un campo de direcciones variables basado en la interpretación geológica**

Distintos atributos geológicos capturados en el proceso de mapeo, pueden relacionarse con el comportamiento del mineral en distintos procesos. Tal como se presentó en la sección anterior, los tipos de alteración definen la recuperación que se puede alcanzar en la flotación, y éstos pueden ser determinados de manera cualitativa a partir de la información geoquímica, que está disponible de manera abundante en muchos proyectos mineros. En el caso de las zonas minerales, éstas muchas veces definen el destino de cada bloque a un tipo de proceso. En minería del cobre, la distinción entre bloques oxidados y sulfurados, define el proceso – lixiviación

o flotación – por lo que su determinación es importante para poder construir un modelo geometalúrgico integrado.

En esta sección presentamos una aplicación en la que se integra información geológica asociada a la disposición geométrica de los distintos dominios, al modelamiento geoestadístico de estas unidades. En particular, se presentan los pasos para construir un campo de direcciones variables a partir de los contactos entre zonas minerales, para condicionar el modelamiento estocástico de estos dominios. Esto corresponde al tercer paso en el flujo de trabajo presentado en la Figura 2, en este caso en relación a la definición de unidades de zonas minerales.

Los detalles de los pasos, se describen a continuación:

**Modelamiento espacial de unidades geometalúrgicas.** Las muestras individuales poseen las concentraciones de los elementos de interés y las categorías definidas en la etapa de mapeo, las que típicamente incluyen las litologías, zonas minerales y tipos de alteración (ver aplicación anterior), aunque otros atributos tales como texturas (Lobos et al, 2016) y asociaciones estructurales también pueden ser considerados. Las categorías asignadas a las muestras pueden ser utilizadas para construir un modelo de la distribución espacial de estas unidades, las que pueden asociarse al comportamiento geometalúrgico de la roca.

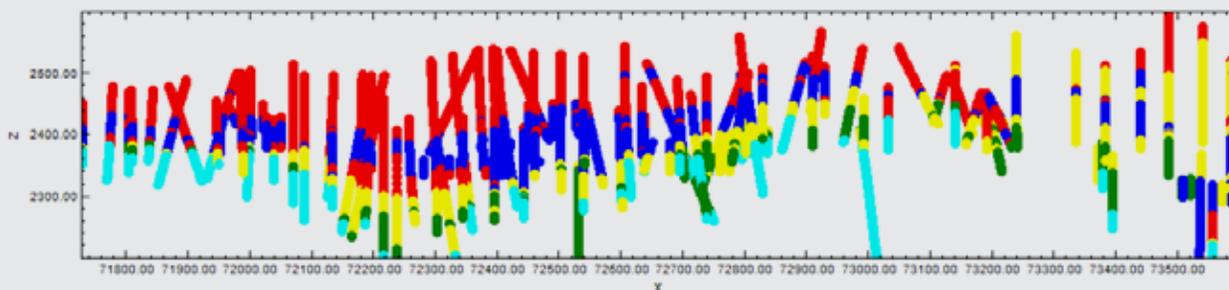
Este modelamiento puede realizarse con técnicas convencionales de modelamiento 3D o bien con métodos geoestadísticos, tales como simulación de indicadores, truncación Gaussiana, simulación plurigaussiana o simulación con estadísticas de múltiples puntos (Chiles and Delfiner, 2012; Mariethoz and Caers, 2014). En esta aplicación, se presenta

un procedimiento para incorporar información geológica secundaria mediante un campo de direcciones variables en el modelamiento de estos dominios, lo que permite mejorar la caracterización de la continuidad espacial de las variables al tomar en cuenta procesos geológicos que erosionan, compactan o deforman los volúmenes mineralizados (Gutierrez, 2014).

En esta aplicación, se utiliza una base de datos de más de 46000 muestras de un pórfido cuprífero, donde las zonas minerales han sido mapeadas en todas las muestras en una de cinco categorías y se quiere poblar el modelo de bloques con esta variable. Las cinco categorías consideradas son: Lixiviado (0), Óxidos (1), Sulfuros (2), Mixtos (3), Hipógeno (4). Una sección típica de la disposición de estas zonas en el yacimiento se muestra en la Figura 5 (Gutierrez, 2014), donde se aprecia que las unidades no son horizontales, sino que ondulan posiblemente de acuerdo a un paleo nivel freático.

**Figura 5**

Sección representativa de las zonas minerales. LIX (0): lixiviado, se muestra en rojo, OX (1): óxidos, en azul, SULF (2): sulfuros, en amarillo, MIX (3): mixtos, en verde y HYPO (4): hipógeno, en celeste.

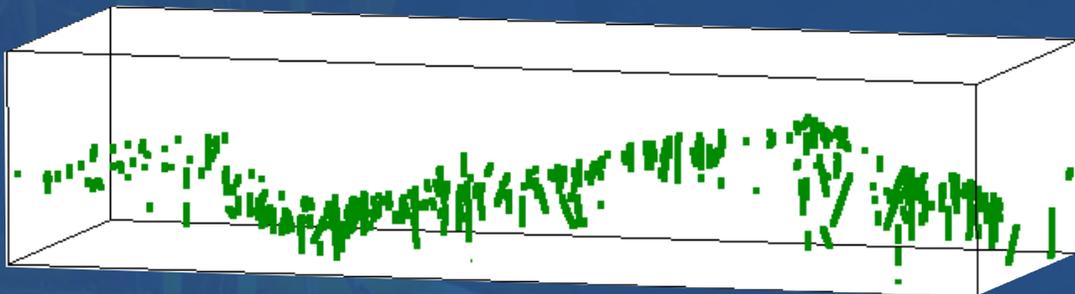


El modelamiento del campo de direcciones variable integrado a la metodología convencional de simulación categórica, se basa en direcciones locales estimadas sobre las muestras de la unidad de Mixtos (Figura 6). En la Figura 7 se muestran las direcciones condicionantes y el campo final utilizado para ser integrado en el paso de simulación estocástica.

Finalmente, se adapta el método de simulación secuencial de indicadores en su versión de variables categóricas (Deutsch, 2006), para poder incorporar el campo de direcciones variables (Boisvert and Deutsch, 2011) y obtener un modelo que se aproxima de mejor manera a la realidad, tal como se muestra en la validación realizada con las zonas minerales mapeadas en los pozos de tronadura, los que se utilizan para verificar el gra-

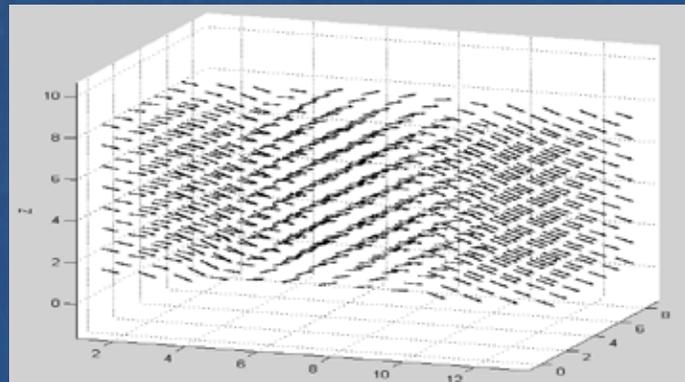
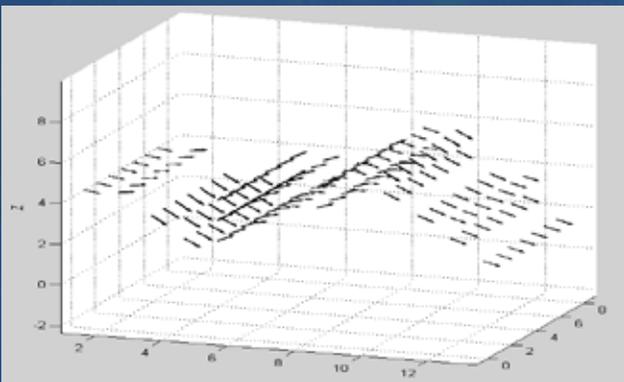
**Figura 6**

Despliegue de muestras pertenecientes a unidad de Mixtos usada para sustentar el modelo del campo de direcciones variables.



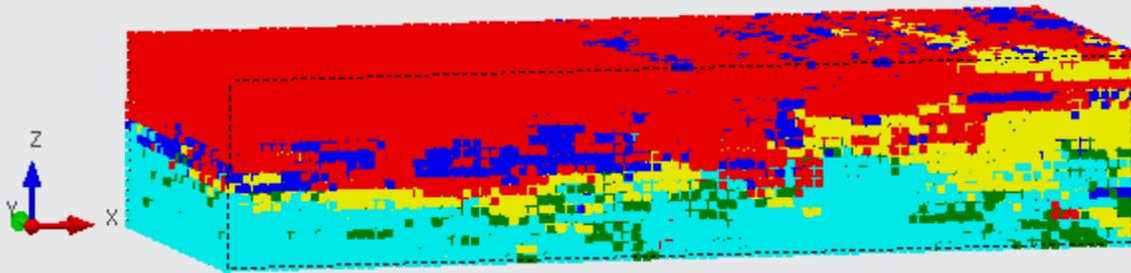
**Figura 7**

Datos condicionantes y campo de direcciones variables final estimado.



**Figura 8**

Una realización mediante simulación secuencial de indicadores utilizando el campo de direcciones variable.



**Tabla 2**

Porcentaje de acierto por categoría de método convencional y método que integra campo de direcciones variable.

Caso	LIX	OX	SULF	MIX	HYPO	TOTAL
SISIM	96.7%	78.2%	71.7%	14.3%	18.6%	81.2%
SISIM_LVA	97.1%	87.1%	78.7%	16.6%	26.6%	85.4%

do de acierto del modelo. Una realización del modelo final, se muestra en la Figura 8. El porcentaje de acierto al comparar con los mapeos en pozos de tronadura se compara con la técnica convencional de simulación en la Tabla 2.

**Modelamiento espacial de atributos y variables índice.** Si se dispone de ensayos metalúrgicos, que permiten relacionar la composición de la roca con alguna respuesta metalúrgica como la recuperación en flotación o la asignación a un proceso, entonces la construcción del modelo de las variables asociadas a la composición de la roca, además del modelo de los dominios geometalúrgico,

es necesario para predecir la respuesta (recuperación en flotación o proceso asignado al bloque).

Esto se desarrolla en dos etapas: primero, los atributos resultantes de ensayos geometalúrgicos o bien variables índice, es decir, variables que se consideran relevantes para predecir la respuesta (por ejemplo, las concentraciones de distintos elementos disponibles en la información geoquímica), deben ser modeladas en el espacio. Esto puede realizarse con técnicas geoestadísticas como kriging o simulación (Chiles and Delfiner, 2012). Técnicas avanzadas de modelamiento multivariable existen y están bien documentadas

en la literatura (Sepúlveda et al, 2017; Tolo-sana-Delgado et al, 2018; Bolgkoranou and Ortiz, 2019). Información secundaria puede integrarse también para condicionar las anisotropías locales de estas variables (Garrido et al, 2016; Navarro et al, 2016). En segundo lugar, algún tipo de escalamiento se requiere para ir desde la información de los ensayos de laboratorio a un soporte que esté asociado a la tasa de procesamiento. Este es probablemente uno de los pasos más complejos, dado que estos últimos no están disponibles y se requieren algunas corridas de prueba en la planta de procesamiento, lo que resulta costoso y difícil de hacer. Algunos enfoques basados en el uso de variables aditivas para predecir variables no aditivas (Garrido et al, 2019) o en reglas de escalamiento mediante funciones potencia (Deutsch, 2015) han sido desarrolladas y podrían integrarse al flujo de trabajo.

**Predicción espacial de comportamiento metalúrgico.** Un modelo que relaciona las variables índices con la respuesta geometalúrgica de interés puede construirse mediante técnicas estadísticas multivariantes. Regresión multilineal, redes neuronales artificiales, métodos basados en kernel, etc., están entre un gran conjunto de posibilidades. En cada bloque, donde se han modelado los atributos relevantes vinculados a la respuesta, se puede aplicar el modelo predictivo, de manera de obtener una predicción de la respuesta metalúrgica en cada bloque, lo que puede ser utilizado en la optimización del plan minero (Lopez et al, 2016).

### Aplicación 3: predicción del efecto de la interacción agua-roca en procesamiento de minerales

Las rocas son formadas por diversos procesos geológicos: ígneos, metamórficos, sedimentarios e hidrotermales, entre otros, y bajo condiciones termodinámicas y fisicoquímicas específicas. En depósitos hidrotermales, por ejemplo, se encuentran distintas condiciones fisicoquímicas, donde asociaciones cristalinas



específicas cristalizan en equilibrio con fluidos hidrotermales. Asociaciones minerales hidrotermales y minerales en general, al estar en contacto con agua, tienden al equilibrio, generando condiciones hidrotermales de *buffer*, en particular pH, Eh y la composición geoquímica. Este comportamiento no se restringe a

depósitos minerales, también ocurre cuando los minerales están siendo procesados.

En procesamiento de minerales, las rocas son molidas en presencia de agua. Esto incrementa las reacciones agua-roca que afectan las condiciones fisicoquímicas en el proceso (Bruckard et al., 2011). Esto es importante dado que el desempeño de la flotación depende de la composición mineral, el funcionamiento de los elementos químicos, pH, Eh

mediante el tamaño de partículas ( $P_{80}$ ), poniendo poca atención al efecto que las condiciones fisicoquímicas determinadas en el proceso de molienda tienen en la flotación.

El efecto de la composición mineral y de la interacción agua-roca en las condiciones fisicoquímicas de la pulpa, así como el desempeño de los procesos de conminución y flotación, puede ser modelado con un enfoque geometalúrgico, combinando el modelamien-



y la composición mineral del agua (Hu et al., 2009; Peng et al., 2012; Owusu et al., 2014).

Actualmente, no existe un modelo exhaustivo que integre conminución y flotación. Los modelos disponibles usualmente apuntan a los procesos por separado, vinculados

to geoquímico para generar un modelo exhaustivo que vincule la conminución con la flotación, no solo a través de la distribución de tamaños de partículas, sino que también considerando las condiciones fisicoquímicas y químicas de la pulpa que son transferidas de un proceso al siguiente (Lois, 2015).

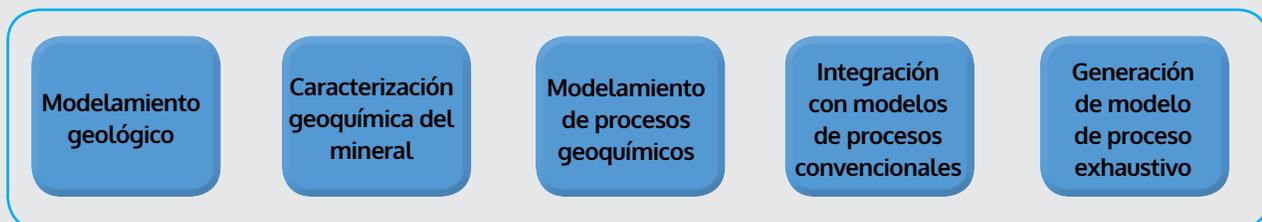
Usando la geoquímica para predecir la naturaleza de las interacciones agua-roca en procesamiento de minerales permite una mejor comprensión del efecto que las distintas asociaciones minerales, típicas en sistemas de pórfidos cupríferos, tienen en las condiciones químicas y fisicoquímicas de la pulpa, y en el desempeño metalúrgico combinado de la conminución y flotación. Esto permite la generación de un modelo llamado geometalúrgico de procesamiento de minerales, que depende, no solo de la forma de operación de los procesos, sino que también de las características de los minerales alimentados. Los pasos para construir este tipo de modelo, se muestran esquemáticamente en la Figura 9.

#### Aplicación 4: implementación de modelos geometalúrgicos de consumo de energía para la integración de ERNC

La integración de energías renovables no convencionales (ERNC) en procesos mineros y metalúrgicos impone un desafío producto de su intermitencia, la que, combinada con la variabilidad en el consumo de energía por parte de los distintos procesos, hace que no necesariamente se pueda conciliar la oferta de energía por parte de fuentes de ERNC con la demanda energética por parte de las distintas operaciones unitarias. Desde el punto de vista de la demanda energética, la geometalurgia ofrece una alternativa de mode-

**Figura 9**

Etapas para construir un modelo geometalúrgico de proceso basado en la interacción agua-roca en procesamiento de minerales.



lamiento que permite llenar el espacio que hay entre los modelos de largo plazo y estado estacionario, utilizados para dimensionar el tamaño de los equipos e instalaciones y aquellos modelos dinámicos, de corto plazo, utilizados en horizontes temporales de corto alcance, con objetivos como por ejemplo el control de procesos.

En el caso de la predicción de variabilidad de demanda de energía en procesos, se requiere contar con una caracterización geometalúrgica que permita inferir consumos específicos en los diferentes procesos. Estos modelos son necesarios si se opta por incorporar estrategias de gestión de demanda de energía (*demand side management*, DSM) que permitan hacer un mejor uso de las ERNC (Pamparana, et al., 2017) y su variabilidad, sin tener que recurrir a instalaciones sobredimensionadas de almacenamiento de energía.

La integración de energías renovables no convencionales (ERNC) en procesos mineros y metalúrgicos impone un desafío producto de su intermitencia.





Ortiz et al. (2019) muestran un modelo integrado para dimensionar un sistema fotovoltaico (PV) asistido por baterías (*Battery Energy Storage System*, BESS) que sirva para proveer energía solar a la operación de un molino SAG, donde la variabilidad en términos de moliendabilidad del mineral se modela con herramientas geoestadísticas y la variabilidad en radiación solar es simulada mediante cadenas de Markov. El sistema completo, construido sobre data sintética, es optimizado considerando un plan minero fijo (secuencia de extracción) con el objetivo de entender el potencial que presenta la incorporación de DSM. El tamaño y costos de la

infraestructura necesaria se determinó considerando 50 simulaciones de secuencias de moliendabilidad (estimada) que son luego comparadas contra moliendabilidad real-sintética para determinar el consumo de energía real esperado. La estimación de variabilidad en el consumo de energía permite un adecuado dimensionamiento del sistema PV-BESS, mostrando no solo potencial para reducir la huella de carbono del proceso sino también reducir costos. La incorporación de *stockpiles* diferenciados (duro - blando) para implementar DSM muestra ventajas adicionales, en términos de reducción del tamaño de las baterías en el componente BESS del sistema.



### 3. Visión: Ideas y propuestas

En las secciones anteriores, se presentaron algunos de los problemas y desafíos de implementar un flujo de trabajo para modelamiento geometalúrgico.

Existen aún muchas brechas y desafíos pendientes que requieren mayor investigación para poder disponer de las metodologías necesarias para completar un modelo geometalúrgico completo. Los flujos de trabajo aquí presentados constituyen herramientas para simplificar el enfoque de modelamiento y comprender los distintos

La estimación de variabilidad en el consumo de energía permite un adecuado dimensionamiento del sistema PV-BESS, mostrando no solo potencial para reducir la huella de carbono del proceso sino también reducir costos.



componentes interrelacionados en el modelo completo. Una vez que cada etapa del modelo ha sido desarrollada, análisis de sensibilidad pueden mejorar la comprensión de los procesos fenomenológicos a los que el mineral y el material estéril son sometidos en el procesamiento.

A continuación, identificamos los siguientes temas pendientes que requieren investigación y desarrollo adicionales:

- **Adquisición de datos y control y aseguramiento de calidad de datos geometalúrgicos.** Existe una necesidad importante de desarrollar procedimientos para adquirir datos de manera sistemática y de bajo costo, en relación a composición

elemental, proporciones minerales, asociaciones minerales, tamaño de fragmentos, parámetros geotécnicos, entre otros. Restricciones de tiempo y la falta de comprensión de la relevancia de este tipo de información para modelar los procesos han prevenido que estos pasos se integren en los hábitos de trabajo en distintas operaciones. Datos más ricos, tales como información hiperespectral, mineralogía cuantitativa, modelos texturales cuantitativos, pueden ser extremadamente importantes para entender de mejor manera el comportamiento de la roca a distintos procesos físicos y químicos. Además, protocolos de control y aseguramiento de calidad no son aplicados de manera regular a este tipo de datos.

Existe una necesidad importante de desarrollar procedimientos para adquirir datos de manera sistemática y de bajo costo, en relación a composición elemental, proporciones minerales, asociaciones minerales, tamaño de fragmentos, parámetros geotécnicos, entre otros.

- **Número de datos insuficiente para construir modelos espaciales y relaciones predictivas**, debido a su costo. Hay una necesidad de desarrollar ensayos y mediciones índice (proxy), para reducir el costo de adquisición de datos geometalúrgicos. La sistematización de mediciones de bajo costo que puedan correlacionarse con ciertas respuestas geometalúrgicas relevantes es una necesidad para aumentar la cantidad de información y cobertura espacial necesaria para la construcción de los modelos
- **Herramientas insuficientes para descubrir relaciones por medio de aproximaciones estadísticas**. Dado que los métodos predictivos estadísticos presentan limitaciones en su capacidad de ajustarse a datos, en particular cuando las relaciones son complejas o cuando las respuestas se ven afectadas por mezclas poblacionales, es necesario aplicar con cuidado métodos de tipo caja negra y apuntar esfuerzos al desarrollo de modelos fenomenológicos, mediante la integración de la geología en el entendimiento del procesamiento de minerales y procesos metalúrgicos.
- **Modelos metalúrgicos pobres**. Hay una falta de comprensión de muchos de los procesos físicos y químicos que ocurren en cada una de las etapas del procesamiento de minerales y procesos metalúrgicos.

Se requieren estudios experimentales, así como análisis teóricos para mejorar el conocimiento en el área, los que luego deben ser testeados y validados en terreno, para comprender los problemas de escalamiento y su efecto en las variables operacionales de los procesos. La integración de variabilidad e incertidumbre es, posiblemente una de las principales limitaciones de los modelos metalúrgicos hoy disponibles. Muchos de los "simuladores metalúrgicos" son simples predictores con parámetros de entrada variables, pero que no reportan la variabilidad en la salida, debido a los múltiples parámetros no considerados en el modelo y que afectan su respuesta.

#### 4. Conclusiones

La geometalurgia requiere integrar el conocimiento geológico en el modelamiento del procesamiento de minerales y procesos metalúrgicos, así como la comprensión de los fenómenos físicos y químicos que ocurren en dichos procesos. Existe un conjunto típico de análisis y etapas de trabajo que se utilizan para la construcción de un modelo geometalúrgico completo.

El uso de flujos de trabajo, donde variables de entrada y de salida pueden vincularse para modelar cada etapa y luego integrarse en el modelo completo de una serie de procesos, es útil, dado que permite mejorar el entendimiento de los procesos específicos, da nociones acerca de qué variables y datos son los más relevantes para mejorar el poder predictivo de los modelos, y reducir el problema a una unidad de tamaño manejable.



Vinculando estos pequeños flujos de trabajo en el flujo de trabajo integrado para modelar el proceso completo del mineral puede manejarse, incluso integrando modelos que utilizan distintos niveles de sofisticación en cada etapa. Por ejemplo, se puede combinar un modelo estocástico de la distribución espacial de las propiedades de la roca bastante sofisticado, con un modelo predictivo simplificado de la respuesta a la flotación. Esto puede ser suficiente para entregar nociones del comportamiento de los materiales en el proceso. Estos flujos de trabajo pueden extenderse a otras aplicaciones como fracturamiento de las rocas, modelamiento medioambiental de drenaje ácido, consumo de agua y energía, solo por nombrar algunos.

Los modelos predictivos utilizados en cada etapa pueden construirse a partir de la abundante variedad de enfoques estadísticos disponibles y que hoy disponen además de

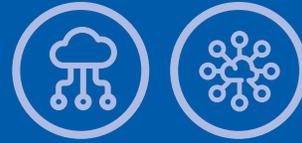
una gran cantidad de código abierto para su rápida implementación. Métodos estadísticos tradicionales, así como los métodos más avanzados disponibles con técnicas de *machine learning* o *deep learning*, pueden ser usados para construir rápidamente una variedad de enfoques de predicción, y pueden ser comparados con facilidad contra datos separados para su validación y testeo.

El enfoque presentado en este capítulo es principalmente una integración de pasos, y como tal, la mayoría de las herramientas utilizadas ya existen. Sin embargo, es necesario contar con experiencia y conocimiento de los procesos para combinarlos e interpretar los resultados, considerando el contexto geológico, minero y metalúrgico de las operaciones modeladas. Esto último constituye probablemente el desafío más difícil de superar.



# Referencias

- Babak, O., Deutsch, C.V. (2009). Collocated Cokriging Based on Merged Secondary Attributes. *Mathematical Geosciences*, 41(8), 921-926.
- Boisvert, J.B.; Deutsch, C.V. (2011). Programs for kriging and sequential Gaussian simulation with locally varying anisotropy using non-Euclidean distances. *Computers & Geosciences* 37(4), 495-510.
- Boisvert, J. B., Rossi, M. E., Ehrig, K. and Deutsch, C. V. (2013). Geometallurgical modeling at Olympic Dam Mine, South Australia. *Mathematical Geosciences*, 45(8), 901-925.
- Bolgkoranou, M., Ortiz, J.M. (2019). Multivariate geostatistical simulation of compositional data using Principal Component Analysis – Application to a Nickel laterite deposit, APCOM 2019.
- Bruckard, W. J., Sparrow, G. J. and Woodcock, J.T. (2011). A review of the effects of the grinding environment on the flotation of copper sulphides. *International Journal of Mineral Processing* 100(1), 1-13.
- Bulatovic, S., Wyslouzil, D. and Kant, C. (1999). Effect of clay slimes on copper, molybdenum flotation from porphyry ores. In proceedings of the Copper 99 – Cobre 99 International Conference, vol. 2, no. 1-3, 95-112.
- Cardenas, E. (2015). Caracterización geoquímica y mineralógica de alteraciones hidrotermales en pórfido cuprífero Escondida, Tesis Magíster en Ciencias, mención Geología, Universidad de Chile, 205 p.
- Chen, C-L., Kaber, D. B. and Dempsey, P. G. (2004). Using feedforward neural networks and forward selection of input variables for an ergonomics data classification problem. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 14(1), 31-49.



Chiles, J. P. and Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty*. Wiley, 2nd edition, 734 p.

Deutsch, C.V. (2006). A sequential indicator simulation program for categorical variables with point and block data: *BlockSIS*, *Computers & Geosciences* 32, 1669-1681.

Deutsch, J.L. (2015). *Multivariate Spatial Modeling of Metallurgical Rock Properties*, PhD Thesis, University of Alberta, Edmonton, 209 p.

Dominy, S.C., O'Connor, L., Parbhakar-Fox, A., Glass, H.J., Purevgerel, S. (2018). *Geometallurgy—A Route to More Resilient Mine Operations*, *Minerals* 8(12), 560, 33 p.

Fawcett, T. (2004). *Roc graphs: notes and practical considerations for researchers*. Intelligent Enterprise Technologies Laboratory, Tech. Rep.

Garrido, M., Navarro, F., Ortiz, J., Moreira, J., (2016). Resource assessment with unfolding methodology: a case study, abstract accepted,

MININ 2016 – 6th International Conference on Innovations in Mine Operations, August 21-23, 2016, Santiago, Chile.

Garrido, M., Ortiz, J. M., Villaseca, F., Kracht, W., Townley, B., Miranda, R. (2019). Change of support using non-additive variables with Gibbs Sampler: Application to metallurgical recovery of sulphide ores. *Computers and Geosciences*, 122: 68-76.

Grus, J. (2015). *Data Science from Scratch – First Principles with Python*, O'Reilly, 311 p.

Gutierrez, R. (2014). *Simulación geoestadísticas de variables categóricas con campos de anisotropías variables*, Tesis de Magíster en Minería, Universidad de Chile, 114 p.

Hu, Y., Sun, W. and Wang, D. (2009). *Electrochemistry of flotation of sulphide minerals*. Springer.

Lobos, R., Silva, J.F., Ortiz, J. M., Díaz, G., Egaña, A., (2016). Analysis and Classification of Natural Rock Textures based on New Transform-based Features, *Mathematical Geosciences*, 48(7), 835-870.

Lois, P. (2015). Interacción agua/roca durante molienda e impactos fisicoquímicos y geoquímicos en la flotación: caso pórfido Escondida, Tesis de Magíster en Ciencias, mención Geología, Universidad de Chile, 228 p

Lopez, A. Barberan, A., Alarcon, M., Vargas, E., Ortiz, J., Morales, N., Emery, X., Egana, A., McFarlane, A., Friedrich, C. (2016). Progress towards Data-driven Mine Planning via a Virtual Geometallurgical Laboratory, Geomet 2016, The Third AusIMM International Geometallurgy Conference, 15-16 June, 2016, Perth, Australia.

Mariethoz, G. and Caers, J. (2014). Multiple-point Geostatistics: Stochastic Modeling with Training Images. Wiley, 376 p.

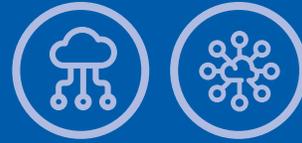
Mariethoz, G., Renard, P., Froidevaux, R. (2009). Integrating collocated auxiliary parameters in geostatistical simulations using joint probability distributions and probability aggregation. Water Resources Research 45(8), W08421, 13 p.

Miranda, R. (2015). Análisis multivariable de alteraciones, Tesis Magíster en Minería, Universidad de Chile, 204 p.

Navarro, F., Garrido, M., Diaz, G., Ortiz, J., Egaña, A., (2016). New human-computer interfaces to capture expert geological knowledge, abstract accepted, MININ 2016 – 6th International Conference on Innovations in Mine Operations, August 21-23, 2016, Santiago, Chile.

Ortiz J.M., Kracht, W., Townley, B., Lois, P., Cárdenas, E., Miranda, R., Álvarez, M., (2015). Workflows in Geometallurgical Prediction: Challenges and Outlook, in Proceedings of the 17th Annual Conference of the International Association for Mathematical Geosciences IAMG 2015, Schaeben, H., Tolosana-Delgado, R., van den Boogaart, K. G., van den Boogaart, R. (editors), 6p.

Ortiz, J.M., Kracht, W., Pamparana, G., and Haas, J., (2019). Optimization of a SAG mill energy system: integrating rock hardness, solar irradiation,



climate change and demand side management. Submitted to *Mathematical Geosciences* (under review)

Owusu, C., Brito e Abreu, S., Skinner, W., Addai-Mensah, J. and Zanin, M. (2014). The influence of pyrite content on the flotation of chalcopyrite/pyrite mixtures. *Minerals Engineering* 55, 87-95.

Pamparana, G., Kracht, W., Haas, J., Díaz-Ferrán, G., Palma-Behnke, R., and Román, R., (2017). Integrating photovoltaic solar energy and a battery energy storage system to operate a semi-autogenous grinding mill. *Journal of Cleaner Production* 165, 273–280

Peng, Y., Wang, B. and Gerson, A. (2012). The effect of electrochemical potential on the activation of pyrite by copper and lead ions during grinding. *International Journal of Mineral Processing* 102, 141-149.

Pimentel, B. and de Souza, R. (2013). A multivariate fuzzy c-means method. *Applied Soft Computing*, 13(4), 1592-1607.

Sepúlveda, E. Dowd, P.A. (2018). Fuzzy Clustering with Spatial Correction and Its Application to Geometallurgical Domaining. *Mathematical Geosciences* 50(8): 895-928.

Sepúlveda, E. Dowd, P.A., Xu, C., Addo, E. (2017). Multivariate Modelling of Geometallurgical Variables by Projection Pursuit. *Mathematical Geosciences* 49(1), 121-143.

Sillitoe, R. (2010). Porphyry Copper Systems. *Economic Geology*, 105(1), 3-41.

Tolosana-Delgado, R., Mueller, U., van den Boogaart, K.G. (2018). Geostatistics for compositional data: an overview. *Mathematical Geosciences*, online first, 42 p.



# Beauchef Minería

Es una iniciativa de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile (FCFM). Su misión es ser un centro de pensamiento que cumpla funciones orientadoras y de apoyo a las políticas públicas en el ámbito minero. De esta forma, busca contribuir al aumento de la competitividad de la industria minera nacional, cooperar con el fortalecimiento y articulación de su ecosistema de innovación y posicionar las actividades de la Facultad de Ciencias Físicas y Matemáticas en el ámbito nacional e internacional.

Como parte importante de su quehacer, Beauchef Minería desarrolla estudios sobre temas de alto interés para la industria minera, los que dan cuenta de las actividades de I+D+i en minería de la Facultad.

Beauchef Minería cuenta con el apoyo de dos Centros y dos Departamentos de la FCFM: Departamento de Ingeniería de Minas, Departamento de Ingeniería Industrial, el Advanced Mining Technology Center y el Centro de Modelamiento Matemático. Con el objetivo de generar investigación multidisciplinaria y fomentar una relación interfacultad, estas áreas colaboran entregando sus conocimientos y experiencia para aportar al debate en temas relevantes para la minería.

---

## **Departamento de Ingeniería de Minas**

Para el Departamento de Ingeniería de Minas, Beauchef Minería representa una plataforma que permite complementar nuestras competencias técnicas en minería y metalurgia con la participación en la discusión pública asociada al desarrollo de la actividad minera. Esta iniciativa ha permitido canalizar, de manera coherente y ordenada, el aporte de la Facultad de Ciencias Físicas y Matemáticas en estos temas.

---

## **Advanced Mining Technology Center (AMTC)**

El Advanced Mining Technology Center busca generar investigación multidisciplinaria de clase mundial, transferir nuevas tecnologías y formar capital humano avanzado. Asimismo, AMTC apunta ser un colaborador relevante en la definición de políticas públicas para nuestro país, por esta razón, ser parte de Beauchef Minería nos permitirá avanzar en la vinculación de la labor académica en el marco de los desafíos de la industria minera y las necesidades de nuestra sociedad.

---

## **Centro de Modelamiento Matemático (CMM)**

El Centro de Modelamiento Matemático ha sido un actor activo en la introducción de herramientas matemáticas y estadísticas sofisticadas para la modelación de procesos claves en la minería chilena y su transformación a la era digital. La iniciativa Beauchef Minería nos permite visibilizar y acelerar la adopción de nuestras tecnologías en la minería.

---

## **Departamento de Ingeniería Industrial**

El Departamento de Ingeniería Industrial de la Universidad de Chile es pionero en la formación de Ingenieros Civiles Industriales y nuestros programas de postgrado son reconocidos a nivel nacional e internacional. El departamento destaca por su investigación aplicada y formación en economía y gestión, particularmente en el área de la minería.

---



Av. Beauchef N° 850, Santiago, Chile.

Teléfono: +562 29771000

contacto@beauchefmineria.cl

<http://www.beauchefmineria.cl/>

ISBN: 978-956-19-1186-4