

La ciencia de datos como herramienta para la innovación pública

María Paz Hermosilla Cornejo

¿Cómo podemos identificar conversaciones con mayor riesgo en un chat de ayuda a jóvenes, para priorizarlas cuando hay alta demanda de apoyo? ¿Cómo podemos prevenir el envenenamiento por plomo en bebés? ¿Cómo podemos reducir los tiempos de respuesta de Bomberos a incendios y accidentes? Todos estos problemas públicos tienen en común lo siguiente: la ciencia de datos puede ayudar a resolverlos.

La ciencia de datos es el conjunto de metodologías para inferir ideas procesables a partir de los datos (Iguar y Seguí, 2017). Se busca crear información útil para la toma de decisiones de una organización. La ciencia de datos es una herramienta clave de la innovación pública, entendida como “el proceso de crear nuevas ideas y convertirlas en valor para la sociedad” (Bason, 2010:34).

El GobLab, laboratorio de innovación pública de la Universidad Adolfo Ibáñez en Chile, tiene la misión de contribuir a la innovación en políticas públicas para beneficiar a la sociedad. Trabaja con organismos públicos, organizaciones de la sociedad civil e investigadores para lograr políticas públicas más eficaces, eficientes y equitativas mediante la ciencia de datos. Una de las formas de lograr su misión es a través de la formación de capacidades, lo que ha sido parte de su quehacer desde su creación en 2017. El primer programa creado por el GobLab fue un diplomado en ciencia de datos para políticas públicas, con el objetivo de enseñar análisis de datos y machine learning en el software R a profesionales de las ciencias sociales que no supieran programar. Sin embargo, no es suficiente formar a los analistas de datos si sus jefaturas no están sensibilizados respecto de qué problemas de gestión pública se pueden abordar con datos, y entienden cómo pasar de un problema a una formulación de proyecto viable que pueda abordarlo. Como lo plantea la OCDE (2015:255), en la fase de toma de decisiones del ciclo de valor de los datos son claves las jefaturas, gerentes o directivos/as.

Conscientes de esa necesidad, a finales de 2018 el GobLab se asoció con el Center for Data Science and Public Policy (DSAPP) de la Universidad de Chicago para crear un curriculum para tomadores/as de decisión de políticas públicas. Con el apoyo del Laboratorio de Gobierno y el Servicio Civil de Chile, se desarrolló el curso “Ciencia de Datos para Directivos Públicos”. Este curso tradujo al español una metodología de formulación de proyectos de ciencia de datos creada por Rayid Ghani, que en ese entonces lideraba el centro de la Universidad de Chicago, denominada en inglés: *data science project scoping*. Esta metodología había sido usada con éxito desde el 2013 por los integrantes del DSAPP con más de 100 gobiernos en diferentes partes del mundo, a través del *Data Science for Social Good Fellowship* (Ver: <https://www.dssgfellowship.org/projects/>).

En el marco del proyecto se mejoró una ficha de proyecto existente y una matriz de madurez de datos, y se creó un programa académico de 24 horas para enseñar la metodología, basado en la iteración de proyectos presentados por los mismos estudiantes. Este curriculum fue probado en 2019 en Chile con más de 70 tomadores de decisión en Santiago y en la región de Valparaíso.

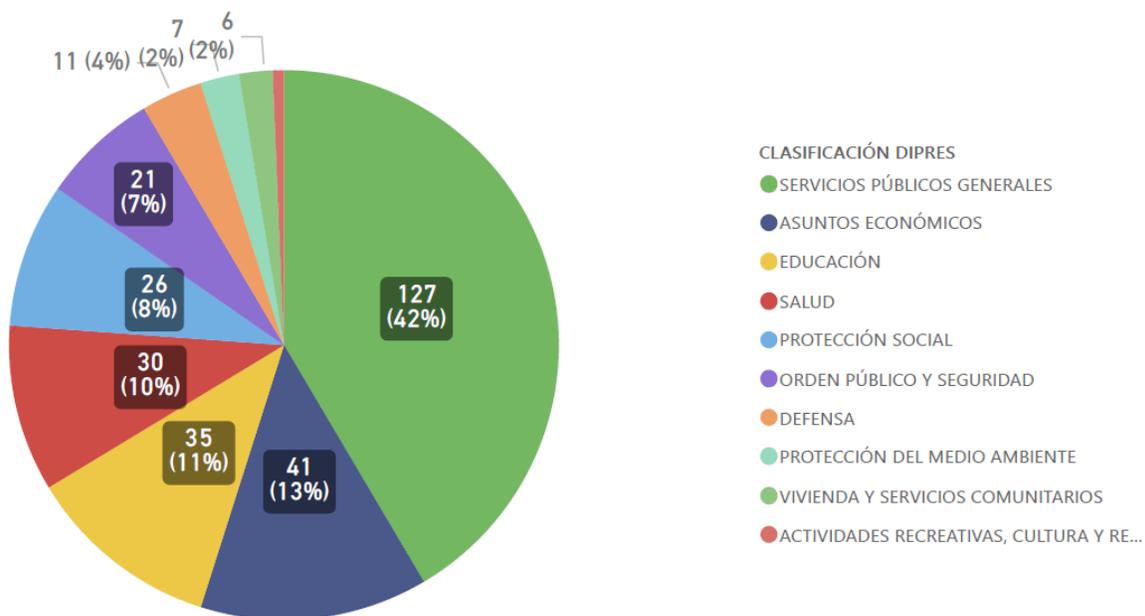
Posteriormente durante la pandemia fue adaptado al formato online y se realizó para otros 133 funcionarios en esa modalidad entre 2020 y 2021. También se creó material pedagógico adicional, incluyendo un glosario y cuatro ejemplos de proyectos. En total, 205 funcionarios públicos han pasado por el programa de formación. El 62% de los participantes tienen cargos de jefaturas y solamente el 35% son mujeres.

Desde el año 2020, también se utiliza la metodología en el diplomado de Ciencia de Datos para Políticas Públicas de la Universidad Adolfo Ibáñez. En 2021 la metodología será parte de un manual de formulación y ejecución de proyectos de inteligencia artificial responsable para tomadores de decisión de política pública que publicará el BID como parte de su iniciativa fAlr LAC (Ver: <https://fairlac.iadb.org/>)

La ciencia de datos y sus aportes a las políticas públicas

En estos 3 años de implementación del curriculum, 306 funcionarios han presentado proyectos para participar en las oportunidades de formación existentes, que en varias ocasiones han sido gratuitas. Realizamos un análisis de las temáticas utilizando la clasificación del gasto público que realiza la Dirección de Presupuestos en Chile (Dipres, 2020:151). En la figura que se muestra a continuación, se vislumbra que la ciencia de datos es una herramienta que puede aportar a diversas áreas de gestión pública.

Figura 1. Temáticas de los proyectos postulados



Fuente: elaboración propia.

Varios de los proyectos que se han formulado se han desarrollado posteriormente por los organismos públicos, en alianza con estudiantes de posgrado de la universidad. Por ejemplo, con el Servicio Nacional de Pesca, se está trabajando en modelos que permitan detectar buques de pesca fraudulentos. Con la Superintendencia del Medio Ambiente, se desarrolló un proyecto de Inteligencia artificial para el cuidado del medio ambiente y la salud de la población consistente en un modelo predictivo para episodios críticos de contaminación en el aire para una zona industrial de Chile.

El Servicio Electoral desarrolló un modelo de clasificación de infracciones en el proceso de financiamiento y propaganda electoral, y la Superintendencia de Educación realizó un proyecto respecto de la identificación y caracterización de establecimientos educacionales en relación a denuncias por clima de convivencia escolar.

Aprendizaje por la acción

La metodología de aprendizaje se caracteriza por estar basada en la iteración de un proyecto que cada participante trabaja durante el curso.

Para eso, debe seleccionar una problemática que cree pueda ser abordada con ciencia de datos y contestar las siguientes preguntas antes de la primera clase:

1. Nombre del proyecto
2. Justificación del proyecto: Describa el problema que enfrenta su organización. Describa cuáles son los actores que son afectados por el problema. ¿Por qué es un problema importante para la organización? ¿Existen riesgos asociados al proyecto?
3. Estado actual: ¿Cómo está resolviendo su organización este problema actualmente?
4. Objetivos: ¿Qué objetivos se alcanzarían si se resolviera el problema? ¿Cuál es el producto esperado?
5. Datos: ¿Qué datos necesita para solucionar el problema? ¿Tiene su organización acceso a esos datos? ¿A qué fuentes tiene acceso? Descríbalos en detalle: nombres de fuentes de datos, contenido y cuántos años de datos históricos están disponibles.

Durante el curso, los docentes retroalimentan las evaluaciones en tres oportunidades, permitiendo a las y los estudiantes ir mejorando su formulación.

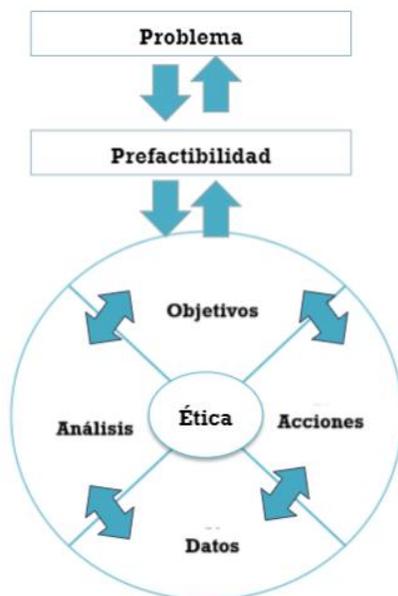
Pasos de la metodología de formulación de proyectos de ciencia de datos

La metodología que se enseña tiene 7 componentes, que no tienen una estructura lineal, sino iterativa: los productos de una fase se actualizan y mejoran con los hallazgos de las otras fases. Esto requiere de flexibilidad por parte de los funcionarios públicos, ya que la conceptualización del proyecto debe ir mejorando constantemente.

La metodología se acompaña con la ficha de proyecto, que tiene preguntas específicas que se deben contestar en cada fase.

En la figura que se muestra a continuación, se muestra un esquema general de la metodología de formulación de proyectos de ciencia de datos:

Figura 2. Esquema general de la formulación de un proyecto de ciencia de datos



Fuente: modificación de Ghani (2016).

Definición del problema

En esta etapa, se requiere describir en detalle la situación que se busca mejorar con el proyecto de ciencia de datos. Se debe contemplar una descripción de las causas del problema, qué políticas públicas existen para abordarlo y las brechas en la situación actual. Esta debe contemplar una explicación de los usuarios o entidades que se ven afectados, de qué manera están afectados y medidas que permitan cuantificar dicha afectación. Por ejemplo, el tiempo medio de espera para una cirugía, el número de estudiantes que abandonan los estudios, costo debido a la evasión fiscal, entre otros. Además, se debe explicar por qué resolver este problema es una prioridad para la agencia pública, y buscar antecedentes sobre casos de uso similares que podrían ser de utilidad para el proyecto, sea en el mismo país o en el extranjero.

Análisis de prefactibilidad

La fase de análisis de pre-factibilidad es clave, porque permite descartar proyectos en una etapa temprana. Esto es relevante para un eficiente uso del tiempo de las y los directivos públicos, que tiene un alto costo de oportunidad debido a las múltiples necesidades de la población usuaria de los servicios públicos. La gran mayoría de las preguntas en esta etapa son estratégicas, y buscan que se prioricen problemas que están dentro de las competencias del organismo público. Se busca visibilizar desde una primera etapa alianzas que sean necesarias de realizar con otros organismos públicos, y los recursos humanos y financieros necesarios para realizar el proyecto. En esta etapa es importante visualizar si los datos que necesitamos existen, y más importante aún, si se puede acceder a ellos. Finalmente, se requiere identificar de manera general los riesgos del proyecto, que pueden ser éticos, de licencia social, de implementación, entre otros.

Definición de objetivos

En esta etapa, utilizamos la metodología SMART para la definición de objetivos del proyecto de ciencia de datos. Esto significa que deben ser específicos, medibles, alcanzables, realistas, y deben tener un horizonte temporal. Aunque parezca de perogrullo, alcanzar el objetivo debería ayudar a solucionar el problema. Si el problema es que una especie de peces está amenazada por la pesca ilegal, un objetivo podría ser aumentar el cumplimiento de las cuotas de pesca por parte de los pescadores artesanales e industriales de la región en un porcentaje determinado.

Es importante recalcar que el objetivo del proyecto no es crear una solución técnica (el modelo de predicción, clasificación, o el análisis descriptivo). El producto de datos es un medio para cumplir el objetivo, como por ejemplo, aumentar la proporción de casas fiscalizadas por posible envenenamiento por plomo en donde viven mujeres embarazadas. Cumplir el objetivo ayudará a resolver el problema, en este caso, de envenenamiento por plomo en bebés.

Al definir los objetivos, es importante también visibilizar las limitaciones a dichos objetivos, ya que serán aspectos importantes a la hora de definir los análisis a realizar. Por ejemplo, si solo tengo la posibilidad de hacer 100 inspecciones al mes, entonces mi modelo debiera recomendar las 100 inspecciones más riesgosas, y ese umbral será relevante para evaluar su desempeño. Las limitaciones típicas incluyen el presupuesto, la falta de capital humano, las restricciones legales, la voluntad política y la licencia social.

Descripción de acciones

Las acciones son las actividades que las instituciones hacen/pueden hacer respecto de un problema. Son los programas o procesos existentes, que podrán ser realizadas de mejor forma cuando la institución tenga la información generada en el proyecto de ciencia de datos. Por ejemplo, la inspección de una empresa, la atención de un paciente, o la asignación de un subsidio de vivienda. El error más común de esta sección es que las y los estudiantes confunden las actividades de un proyecto de ciencia de datos (extracción de datos, exploración de datos, modelamiento) con las acciones de una institución.

Las acciones ideales deberían ayudar a alcanzar los objetivos definidos anteriormente, se requiere que exista una coherencia entre acción y objetivo.

Una cabal comprensión de la acción institucional que será mejorada con ciencia de datos es esencial para el éxito del proyecto. Esto implica conocer quién la ejecuta, de qué manera, con qué frecuencia, qué información se requiere para ejecutarla, entre otros aspectos.

Mapeo de datos

En esta etapa, se requiere definir con precisión los datos internos y externos que se requieren para realizar el proyecto, y hacer una descripción detallada de sus atributos y evaluación de su madurez. Esto implica describir su granularidad, frecuencia, determinar la existencia de identificadores únicos, la entidad o persona responsable de administrarlos, su forma de almacenamiento, entre otros aspectos.

Los proyectos típicos de ciencia de datos usan datos administrativos como la fuente primaria de datos y la mejoran con fuentes de datos disponibles públicamente (censo, otros datos abiertos). La asociación con el sector privado u organizaciones sin fines de lucro podría ser una forma de obtener datos que podrían faltar a nivel interno.

Definición de análisis

En esta fase, se requiere determinar el tipo de análisis que podría proveer la información necesaria para mejorar la acción. Los proyectos típicos de ciencia de datos incluyen una combinación de análisis: descripción, detección, predicción, entre otros. Es importante recordar que el análisis no es el objetivo del proyecto, sino un medio para ayudar a resolver el problema. En esta etapa es importante empezar a pensar cómo será el proceso de validación del análisis, de manera de saber si está ayudando a cumplir con el objetivo del proyecto.

Diagnóstico de consideraciones éticas

En esta etapa de la metodología, se analizan aspectos de proporcionalidad, licencia social, protección de datos, transparencia, equidad y rendición de cuentas. El objetivo es visibilizar los desafíos éticos específicos del proyecto, de manera de que, si decide llevar a cabo, poder realizar las gestiones necesarias para mitigarlos. Para eso se requiere justificar por qué un sistema de ciencia de datos es el medio adecuado para resolver el problema y si se han evaluado otras alternativas. Además se le pide a la tomadora de decisión de política pública que identifique los impactos negativos podría tener su proyecto.

La metodología busca que los funcionarios reflexionen respecto de la aceptación que tendrá la comunidad del uso de datos propuesto, denominado licencia social. Para eso deben ponerse en el rol de los usuarios/afectados por el problema, y también considerar la visión que podría tener la población general.

En materia de protección de datos, se debe tener claro si se está trabajando con datos personales y/o sensibles, y ser capaz de especificar cuáles datos están en esta categoría. Se debe identificar la justificación o base legal para trabajar con esos datos, y las regulaciones que podrían impactar en el proyecto, como las leyes de protección de datos, de transparencia, y regulaciones sectoriales.

En materia de transparencia, es necesario hacer un mapeo de actores que deberían estar al tanto del proyecto. Las partes interesadas suelen incluir a formuladores de políticas, trabajadores de primera línea, organizaciones de la sociedad civil, organismos públicos, personas que se verán afectadas por las acciones, etc. También es necesario considerar algún mecanismo para que las partes interesadas se comuniquen con la institución por el proyecto, y reflexionar sobre la necesidad de explicar los mecanismos de toma de decisión o análisis a implementar.

La consideración ética de la equidad abarca varios aspectos. Por un lado, se requiere identificar las inequidades de base hay en el proceso/entorno donde se inserta el proyecto. Por ejemplo, una política pública puede estar enfocada a un grupo de la población que se caracteriza por tener menor acceso a servicios, o mejor nivel de educación. También se requiere identificar si existen grupos específicos (vulnerables) para los que desees garantizar la equidad de los resultados o la protección de sus derechos. Podrían ser grupos dado su género, edad, localización, clase social, nivel educativo, urbano-rural o etnia. Finalmente, es necesario describir los sesgos que podrían tener los datos, a partir del conocimiento experto del funcionario. En materia de rendición de cuentas, se requiere definir quien elaboraría la respuesta en caso de ocurrir un requerimiento de información respecto del proyecto, quién es responsable si el sistema se equivoca, y definir los mecanismos de monitoreo, control, evaluación, incluyendo su documentación y periodicidad.

El curriculum y su aporte a las competencias para la ciencia de datos

Existen varias competencias claves para la ciencia de datos: la definición de problema, la preparación de los datos, el modelamiento y análisis, la evaluación de los resultados, la ética profesional y la comunicación (Rodolfa et al, 2019). El curriculum creado las aborda con diferente profundidad, ya que está diseñado para tomadores de decisiones y personal no técnico, por lo que no busca entregar las habilidades necesarias para realizar análisis de datos en un software específico como R o python, sino el conocimiento necesario para tomar un problema de política pública que queremos abordar con datos, en una propuesta de proyecto factible.

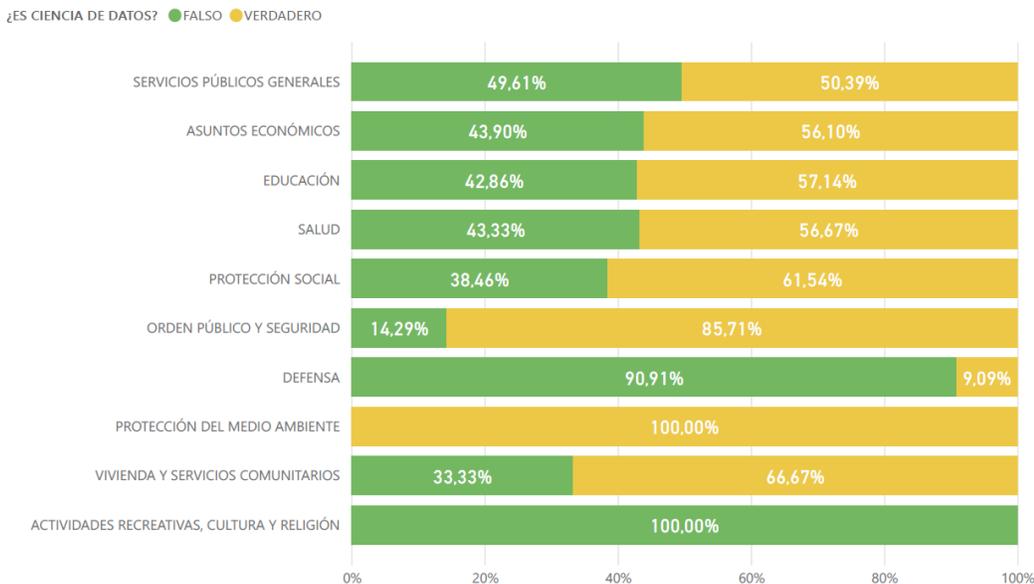
A continuación, identificamos los objetivos de aprendizaje que se abordan en el curriculum, para cada una de las competencias relevantes de la ciencia de datos, identificadas por Rodolfa et al.

Competencia clave de la ciencia de datos	Objetivo de aprendizaje abordado en el curriculum
Definición de problema	Evaluar la factibilidad de usar la ciencia de datos para abordar un problema público. Definir el alcance del problema público a resolver con el proyecto de ciencia de datos. Formular objetivos específicos y medibles para el proyecto. Describir las acciones del organismo público que serán impactadas por el proyecto de ciencia de datos.
Preparación de los datos	Conocer las características de los datos y sus principales problemas. Evaluar la madurez de una base de datos en función de los objetivos de un proyecto.
Modelamiento y análisis	Conocer los diferentes tipos de análisis que podrían ser utilizados en un proyecto de ciencia de datos.
Evaluación de los resultados	Conocer las principales métricas que se utilizan para evaluar un proyecto de ciencia de datos.
Ética profesional	Identificar los desafíos de privacidad, transparencia, equidad, responsabilidad y licencia social que tiene un proyecto de ciencia de datos.
Comunicaciones	Explicar el proyecto de ciencia de datos en un lenguaje entendible para una audiencia externa del organismo público.

Desafíos en la formulación de proyectos de ciencia de datos

Uno de los primeros hallazgos al implementar este currículum fue que el funcionariado público que postulaba a participar del curso, presentaba proyectos que no eran de ciencia de datos. De 306 proyectos postulados, 44% no correspondían a ciencia de datos, sino eran de actividades de recopilación de datos, integración o interoperabilidad, digitalización, desarrollo web, o automatización, entre otras. Existe un importante desafío previo a la ciencia de datos: tener instalado una gobernanza de datos que facilite su posterior análisis. Esta problemática se presentó en la mayoría de las áreas de la política pública.

Figura 3. Tipificación de los proyectos: ¿Es o no ciencia de datos?



Fuente: elaboración propia.

Otro desafío relevante que enfrentan los servidores públicos es elaborar una buena definición de problema, reto habitual en los proyectos de innovación, independiente de la metodología utilizada. Los errores más comunes en las descripciones de problemas son los siguientes:

a) **Describir al problema como la falta de una solución predefinida:** En muchas ocasiones se indica que el problema es que no existe un modelo predictivo o un conjunto de datos en específico. Esto es habitual porque generalmente la persona que busca innovar con ciencia de datos ya tiene una idea predefinida de lo que quiere hacer. Sin embargo, en la etapa de definición de problema no debemos cerrarnos a una solución, ya que la exploración del problema nos puede hacer darnos cuenta de que se requiere algo diferente. Para mitigar este problema, debemos preguntarnos: ¿Cuál es la consecuencia que tiene para la institución/proceso la ausencia de esa información/modelo?. Habitualmente debemos retroalimentar a los estudiantes utilizando la frase “para qué”: ¿Para qué necesitas más información? ¿Para qué quieres integrar bases de datos? ¿Qué decisiones de la organización cambiarían si tuvieras esa información/modelo?

b) **Descripciones genéricas o generales:** Un problema clásico al inicio de los proyectos es la falta de especificación o detalle cuando se describe una situación. Por ejemplo, se habla de decisiones, pero no se indica cuáles. Se habla de “decisiones mejores”, eficiencia o eficacia, sin explicar cómo sería una decisión mejor, o como se mide dicha eficiencia o eficacia.

c) **Alcance demasiado amplio o ambicioso:** Los proyectos parten muy ambiciosos queriendo mejorar muchas decisiones de un área, pero para poder ser viables requieren focalizarse en un ámbito de gestión específico. Por ejemplo, “mejorar decisiones del hospital” se convertirá luego en un proyecto de reforma del sistema de derivaciones a especialistas desde la atención primaria de salud al hospital.

d) **Falta de cuantificación:** Es habitual que las descripciones de problemas no incluyan cifras que permitan entender la cantidad de usuarios afectados, el porcentaje del presupuesto, u otros montos concretos. Para entender las deficiencias o brechas de una situación con respecto al estado deseado, se requiere cuantificar lo más posible la situación actual.

Material pedagógico desarrollado

A continuación se realiza una descripción del material pedagógico desarrollado, que está disponible para su uso bajo licencia creative commons en el sitio web: <https://goblab.uai.cl/proyecto-curso-ciencia-de-datos-para-directivos-publicos/>.

Nombre	Descripción
Ficha de proyecto	Ficha en blanco que debe ser completada durante la conceptualización y diseño del proyecto. Contiene las siguientes secciones: Definición del problema, análisis de prefactibilidad, consideraciones éticas, objetivos, acciones, mapeo de datos, análisis, conformación de equipo.
Proyectos ejemplos	4 fichas de proyectos que se utilizan durante el curriculum: <ul style="list-style-type: none"> - Detectando riesgo de suicidio en aplicaciones de chat (Fundación Todo Mejora) - DART. Inteligencia Artificial para la prevención de la ceguera (Teledx) - Commander (Cuerpo de Bomberos de Santiago) - Desarrollando modelos predictivos de riesgo para apoyar las decisiones de la línea directa de maltrato infantil (Condado de Allegheny, Pennsylvania, EE.UU.)
Bibliografía	Listado de lecturas sugeridas sobre la temática
Matriz de madurez de datos	Herramienta utilizada en clases para enseñar a evaluar la madurez de una base de datos en particular. Para diferentes atributos de una base, como acceso, almacenamiento, calidad, frecuencia y granularidad, entre otros, define los parámetros para poder clasificarla en niveles desde deficiente a avanzado.
Glosario	Definiciones de 55 conceptos relacionados con la temática del curriculum, con sus respectivas referencias.
Curriculum	Programa (syllabus) de un curso de 24 horas presencial, dividido en 3 jornadas. Contempla instrucciones para los docentes.
Videos de clases	11 videos de las clases del curso realizado en 2019 en Santiago, Chile.

Conclusión y próximos pasos

La ciencia de datos ofrece múltiples maneras de mejorar el diseño y la gestión de políticas públicas. Para eso se requieren jefaturas que comprendan el valor que el análisis de datos puede significar para su gestión. El programa desarrollado ha permitido formar a más de 200 directivos públicos en una metodología de formulación de proyectos de ciencia de datos.

Se adaptó al español una metodología existente, que ha sido mejorada con iteraciones periódicas, y cuyo material pedagógico se encuentra disponible en formato abierto bajo licencia creative commons.

Actualmente el GobLab está colaborando con la Universidad Carnegie Mellon y la Escuela de Gobierno del Instituto Tecnológico de Monterrey para continuar mejorando el material y que esté disponible tanto en inglés como en español. Se están realizando encuestas de seguimiento a los directivos públicos que han sido entrenados en la metodología, de manera de evaluar si han podido integrar los aprendizajes y monitorear el avance de los proyectos.

Se ha evidenciado la necesidad de proveer formación especializada en temáticas de gobernanza de datos, debido a que la ausencia de ésta dificulta la realización de los proyectos de ciencia de datos. También existen ciertas tipologías de proyectos que son particularmente desafiantes, como los proyectos de fiscalización, que requieren de conocimientos especializados. Para abordar esas necesidades el GobLab ha desarrollado nuevos programas de formación que actualmente dicta la Escuela de Gobierno de la Universidad Adolfo Ibáñez.

Bibliografía

- Bason, C. (2010). *Leading public sector innovation: Co-creating for a better society*. Bristol, UK; Portland, OR, USA: Bristol University Press. doi:10.2307/j.ctt9qgnsd
- Dirección de Presupuestos (2020). *Estadísticas de las Finanzas Públicas 2010-2019* https://www.dipres.gob.cl/598/articles-203350_doc_pdf.pdf
- Ghani, Rayid (2016). *Scoping Data (for Social Good) Projects* <https://www.dssgfellowship.org/2016/10/27/scoping-data-science-for-social-good-projects/>
- GobLab UAI (2020). *Desarrollando capacidades de ciencia de datos en directivos públicos, Chile*. <https://goblab.uai.cl/desarrollando-capacidades-de-ciencia-de-datos-en-directivos-publicos/>
- Igual, L., & Seguí, S. (2017). *Introduction to Data Science*. In *Introduction to Data Science* (pp. 1-4). Springer, Cham.
- OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, <https://doi.org/10.1787/9789264229358-en>.
- Rodolfa, K.T., Unánue, A.D., Gee, M., & Ghani, R. (2019). *An Experience-Centered Approach to Training Effective Data Scientists*. *Big Data*, 7, 249-261.

Reseña biográfica

María Paz Hermosilla Cornejo. Es Directora de GobLab UAI, laboratorio de innovación pública de la Escuela de Gobierno de la Universidad Adolfo Ibáñez. Magíster en Administración Pública de la Universidad de Nueva York y Periodista de la Pontificia Universidad Católica de Chile. Es Directora Académica del Diplomado en Ciencia de Datos para Políticas Públicas de la Escuela de Gobierno, y docente e investigadora en ética de datos en programas de magíster de la Facultad de Ingeniería y Ciencias y la Escuela de Negocios. Su área de especialización es la innovación pública, específicamente cómo la tecnología transforma el gobierno, innovaciones como la ciencia de datos, los datos abiertos, el crowdsourcing y los laboratorios. Fue funcionaria pública, liderando iniciativas de rediseño de los servicios de atención al ciudadano, transparencia, digitalización de trámites y datos abiertos. Lidera proyectos de investigación aplicada en ciencia de datos en diversos organismos públicos, y es parte del Comité Asesor de Datos de Interés Público del Ministerio de Ciencia, Tecnología, Innovación y Conocimiento.