

MODELOS DE MADUREZ PARA LA MEJORA DE CALIDAD DE LOS DATOS DE LOS INDICADORES DE DESARROLLO SOSTENIBLE

Ignacio Marcovecchio^{1,2}, Elsa Estevez^{1,3}, Pablo Fillottrani^{1,4}

¹ Depto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, Argentina

² United Nations University Institute on Computing and Society, Macao SAR, China

³ Instituto de Ciencias e Ingeniería de la Computación, (UNS-CONICET), Argentina

⁴ Comisión de Investigaciones Científicas Provincia de Buenos Aires, Argentina

ignaciomarcovecchio@gmail.com, {ece,prf}@cs.uns.edu.ar

Tesis de Doctorado expuesta en la ciudad de Bahía Blanca el día 24 de Octubre de 2019

RESUMEN

Disponer de datos útiles y confiables para poder tomar decisiones informadas es uno de cimientos necesarios para alcanzar los objetivos de la Agenda de Desarrollo de 2030. Sólo mediante datos confiables que describan claramente la realidad se pueden definir estrategias y asignar recursos que contribuyan a alcanzar los Objetivos de Desarrollo Sostenible. El trabajo de investigación que se presenta en este documento contribuye con el cumplimiento de Agenda 2030 a través de intervenciones que buscan fortalecer las capacidades de las entidades responsables de producir datos que describan la realidad de los distintos países. En particular, formula un modelo de madurez de la capacidad prescriptivo y multidimensional para evaluar y mejorar la capacidad de los procesos que las Oficinas Nacionales de Estadísticas utilizan para obtener e informar datos estadísticos oficiales a los indicadores de los Objetivos de Desarrollo Sostenible. Las principales contribuciones de la tesis, además del modelo mencionado, incluyen una matriz de alineación de indicadores que permite determinar la similitud entre indicadores de desarrollo, un conjunto de indicadores de gobierno digital que aportan datos y metodologías para los indicadores de desarrollo sostenible, y un conjunto de recomendaciones para utilizar modelos de madurez de la capacidad como instrumentos para estandarizar la producción de datos estadísticos.

Palabras clave: Objetivos de Desarrollo Sostenible, Modelos de Madurez de la Capacidad, Capacidad Institucional

CONTEXTO

El presente documento exhibe un resumen extendido de la tesis de doctorado realizada por el Dr. Ignacio Marcovecchio en el marco de la colaboración entre la Universidad Nacional del Sur (UNS) y el Instituto de Computación y Sociedad de la Universidad de Naciones Unidas (UNU-CS) en Macao SAR, China. El trabajo de investigación fue realizado bajo la dirección de los Profesores del Departamento de Ciencias e Ingeniería de la Computación de la UNS Dra. Elsa Estevez y Dr. Pablo Fillottrani, y con la colaboración del Dr. Mamello Thinyane, Investigador Principal de UNU-CS y Director del Laboratorio de Datos para Desarrollo Sostenible de UNU-CS [1]. En la UNS, este trabajo se enmarcó dentro de las líneas de investigación en Modelos y Aplicaciones de Interoperabilidad Semántica en Gobernabilidad Electrónica del Laboratorio de Investigación y Desarrollo en Ingeniería de Software y Sistemas de Información [2] y en Ingeniería de Software y Gobierno Digital del Instituto de Ciencias e Ingeniería de la Computación [3].

1. INTRODUCCIÓN

En septiembre de 2015 los líderes de 193 países definieron 17 objetivos para el desarrollo sostenible que el mundo se propone alcanzar para el año 2030. Estos objetivos, que se conocen como los Objetivos para el Desarrollo Sostenible (Sustainable Development Goals – SDGs), definen la agenda global de desarrollo y plantean un conjunto ambicioso de objetivos que deben mantener el balance entre los tres pilares

fundamentales del desarrollo sostenible: inclusión social, desarrollo económico y sostenibilidad del medio ambiente. Los 17 objetivos buscan alcanzar 169 metas que van a ser controladas y evaluadas a través de 232 indicadores. La Comisión de Estadísticas de Naciones Unidas [4] desarrolló un sistema global de indicadores que permite controlar el progreso de los países para cumplir con los SDGs [5].

Se prevé que el control de los indicadores de los SDGs va a demandar grandes esfuerzos para que se produzcan datos confiables y de calidad, manteniendo la premisa de que “nadie sea dejado afuera” [6]. Sin embargo, estos objetivos representan un gran desafío a la capacidad de muchos países para medir el progreso para alcanzar las metas de los SDGs [7] ya que la capacidad de los actores principales del ecosistema debe ser potenciada para que se pueda utilizar y sacar provecho de los datos. Por este motivo, resulta indispensable asegurar que todos los países cuenten con un sistema nacional de estadísticas capaz de producir y controlar datos estadísticos confiables que cumplan con los estándares y expectativas globales [6].

Disponer de datos confiables es un factor crítico para poder transformar los SDGs en herramientas útiles para la toma de decisiones y la solución de los problemas. Sin datos actualizados y confiables, el diseño y la implementación de políticas adecuadas resulta muy difícil. Por estas razones, los datos juegan un rol fundamental en el sistema de control de los SDGs.

Para supervisar el progreso, responsabilizar a los gobiernos y promover el desarrollo sostenible es indispensable contar con instituciones fuertes, capaces de satisfacer de manera rápida y eficiente las cambiantes demandas de calidad de datos e información. Este trabajo busca fortalecer la capacidad de las instituciones responsables de liderar y coordinar los Sistemas Nacionales Estadísticos (National Statistical Systems, NSS) para que puedan ser más eficientes en la obtención e integración de datos para la generación de información que permita la toma de decisiones informadas con el fin de lograr los objetivos de desarrollo sostenible.

Para poder realizar intervenciones que resulten útiles y novedosas, se intentarán contestar las siguientes preguntas de investigación:

- P1. ¿Qué prácticas estadísticas e instrumentos garantizan la calidad de los datos para los indicadores de desarrollo sostenible?
- P2. ¿Qué elementos del ecosistema de los indicadores sociales contribuyen a la producción de valor y cómo se relacionan?
- P3. ¿Qué instrumentos de soporte y mejora de la calidad de datos son adecuados para asegurar la calidad de los datos estadísticos que producen las oficinas nacionales de estadísticas?

Construyendo sobre la premisa de que cuanto más maduras son las organizaciones dentro del ecosistema nacional de datos, más confiables serán los datos que pueden producir, esta tesis propone la utilización de modelos de madurez de la capacidad (Capability Maturity Model, CMM) como herramientas para mejorar la capacidad de las entidades responsables de obtener e informar datos sobre el progreso en la implementación de la agenda de desarrollo. Más específicamente, propone la formulación de un CMM prescriptivo y multidimensional para evaluar y mejorar la capacidad de los procesos que las Oficinas Nacionales de Estadísticas (National Statistical Office, NSO) utilizan para producir datos oficiales para los indicadores de los SDGs.

El resto de este documento está organizado como se detalla a continuación. En la sección 2 se presenta la metodología que sirvió de guía para el desarrollo de este trabajo de investigación. A continuación, en la sección 3, se posiciona y enmarca el presente trabajo en la base de conocimientos a partir del relevamiento y estudio de trabajos relacionados, identificando los faltantes e identificando las similitudes y diferencias. En la sección 4 se presenta el modelo propuesto como herramienta para la mejora de la calidad de los datos de los indicadores de desarrollo sostenible, mientras que en la sección 5 se describen las actividades realizadas para la validación del modelo propuesto. Finalmente, en las últimas dos secciones se presentan una discusión de los descubrimientos realizados durante el desarrollo de este trabajo (sección 6) y las conclusiones alcanzadas (sección 7).

2. METODOLOGÍA

Con la motivación de contribuir al ecosistema mediante la introducción de nuevos artefactos y procesos para la construcción de dichos artefactos [8], este trabajo adopta una filosofía Pragmática y sigue el enfoque de la Ciencia del Diseño (Design Science, DS). En particular, adopta la vista de tres ciclos de la DS propuesta por Hevner y que entiende a la DS como tres ciclos de actividades fuertemente relacionados (*relevancia, diseño y rigor*) que diferencian a la DS de cualquier otro tipo de paradigmas [9].

Por la naturaleza del problema que se quiere resolver, que implica influir cambios en la forma de producir datos estadísticos con propósitos sociales a través del diseño y la apropiación de nuevas herramientas computacionales, la estrategia de investigación seleccionada es Investigación-Acción (*Action Research*) [10]. Asimismo, debido a que el ecosistema de datos es estudiado y evaluado de acuerdo con su evolución en el tiempo, esta investigación se enmarca en un horizonte temporal longitudinal.

Todas las decisiones de diseño de la investigación enunciadas en esta sección están fundamentadas en la visión del autor y en función de su comprensión del dominio y de los objetivos perseguidos, y están sustentadas por una extensa revisión de la bibliografía del dominio. Por ejemplo, un estudio sistemático de mapeo de la literatura en modelos de madurez que analizó más de 200 artículos resalta que la DS es el enfoque más utilizado cuando se desarrollan modelos de madurez [11]. La Figura 1 ilustra y resume el diseño de esta investigación.

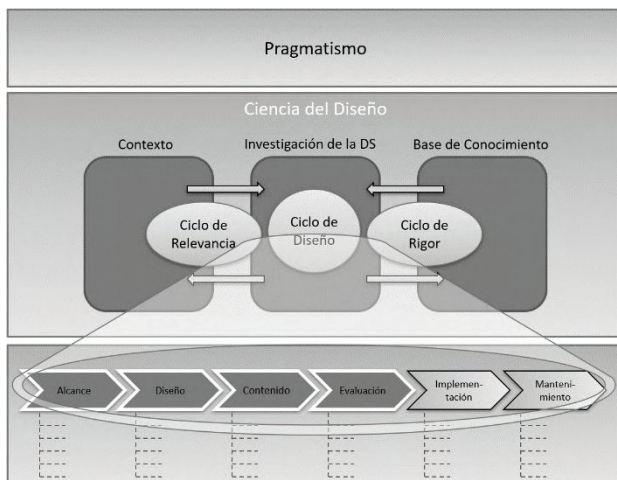


Figura 1: Diseño de la Investigación

3. TRABAJOS RELACIONADOS

A partir de una revisión extensiva de la literatura se identificaron numerosos esfuerzos que buscan mejorar el funcionamiento y los resultados generados por las entidades estadísticas nacionales. Debido a la gran cantidad y diversidad de trabajos existentes, y con el fin de facilitar su comprensión, los trabajos relevantes identificados fueron categorizados de acuerdo con el tipo de la solución. Las seis categorías definidas fueron marcos de trabajo (10), arquitecturas empresariales (2), estudios de nivel de preparación (3), procesos y estándares (2), y modelos (4). La Figura 2 resume el universo de instrumentos relevantes para la gestión de la calidad de datos de organizaciones estadísticas.

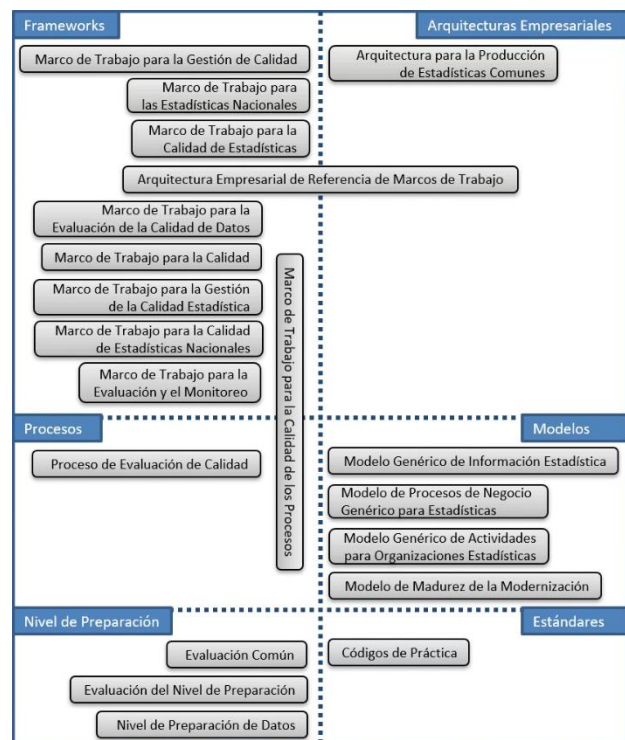


Figura 2: Trabajos Relacionados

La primer observación es que existe una demanda manifiesta de contar con datos e información confiable dentro de la comunidad estadística internacional y que existen numerosas iniciativas para mejorar la calidad y la precisión de los datos. Se observaron además dos grandes limitaciones; la primera es que gran parte de los trabajos se enfocan en evaluar y mejorar la calidad de los datos producidos, prescindiendo de cómo fueron producidos; la segunda es que ninguno de los trabajos se concentra en las demandas específicas de la Agenda 2030.

Mientras que ninguno de los trabajos relevados persigue los mismos objetivos que esta tesis, existen algunas similitudes y superposiciones. Desde el punto de vista de la utilización de modelos de madurez en el contexto de las agencias estadísticas, el trabajo que presenta más similitudes es el Modelo de Madurez de la Modernización (Modernization Maturity Model, MMM) ya que puede ser utilizado no sólo para identificar la madurez de las organizaciones estadísticas, sino que también las ayuda a modernizar la forma en la que operan [12]. Una de las diferencias fundamentales con el modelo propuesto en este trabajo radica en su enfoque: mientras que el CMM se centra en los procesos que producen datos para los indicadores de los SDGs, el MMM promueve la modernización de las formas de producir estadísticas oficiales en general. La evolución de ambos modelos también es distinta: mientras que el CMM es prescriptivo, el MMM es sólo descriptivo. Desde el punto de vista de especializarse en la producción de datos para la Agenda 2030, sólo dos de los trabajos satisfacen esta característica: la Evaluación Integrada Rápida (Rapid Integrated Assessment, RIA) que se enfoca en los planes nacionales de desarrollo [13], y la Evaluación Común de Países (Common Country Assessment, CCA) que se centra en políticas y programas nacionales [14]. Trabajos como el CCA pueden complementar al CMM que se propone en este trabajo ya que ofrecen el potencial de asegurar que el soporte otorgado por las agencias de Naciones Unidas en su conjunto, dentro de un país, sea coherente y complementario, sacando provecho de la experiencia, los recursos y el mandato de cada agencia.

Por lo tanto, luego de una extensiva revisión de la literatura – y de acuerdo con el conocimiento del autor – no existen herramientas diseñadas específicamente para mejorar los principios de los datos generados por las entidades estadísticas para el control de los SDGs y para evaluar la madurez de la capacidad de los procesos que emplean dichas entidades para producir datos estadísticos para los indicadores de los SDGs.

4. MODELO DE MADUREZ

El ciclo de diseño es el núcleo de todo proyecto de investigación de DS y es donde los artefactos se construyen y evalúan. Para guiar el desarrollo dentro del ciclo de diseño se ha seleccionado un marco de trabajo que incluye un proceso de seis fases para el desarrollo de modelos de madurez [15]. Las fases de desarrollo incluyen la definición del alcance, diseño, contenido, evaluación, implementación y mantenimiento. El orden de las fases es importante porque las decisiones que se realizan en una fase pueden tener influencia en las fases siguientes. En esta sección se describe instanciación de las fases de diseño y contenido.

4.1 Diseño

El diseño de la arquitectura del modelo representa las bases para su desarrollo y aplicación [15]. La arquitectura de los modelos de madurez se define a través de los componentes y su interacción [16] y se compone, mínimamente, por los objetos o capacidades a ser medidos (las dimensiones sobre las cuales se realizará la evaluación de la madurez) y por una medida que permita determinar su estado (los niveles de madurez) [11].

De los ejercicios de diseño se obtuvieron cuatro dimensiones que cubren todos los aspectos relevantes para la producción de datos estadísticos para los indicadores de los SDGs. Las dimensiones resultantes son:

- *Entorno* – representa el ecosistema en el cual las NSOs están inmersas y se compone de los factores institucionales y organizacionales que conforman el NSS y que tienen influencia en la efectividad y credibilidad de las NSOs para producir datos confiables y de buena calidad para los indicadores de los SDGs.
- *Recursos* – para poder desarrollar sus actividades, las NSOs deben contar con los recursos suficientes que les permitan contar con la última tecnología y personal bien preparado, y que asimismo les permita desempeñarse de manera independiente de cualquier presión económica. Los recursos no sólo deben ser suficientes, sino que deben ser administrados de manera eficiente.

- *Datos* – son tanto la materia prima como los productos finales que utilizan las NSOs para cumplir con sus responsabilidades.
- *Procesos* – si los datos son la materia prima con la que trabaja una NSO, los procesos son las herramientas que determinan cómo dichos datos se procesan para convertirse en productos estadísticos confiables y de buena calidad que permitan la medición, el control y la toma de decisiones basadas en evidencia..

De la evolución en el desarrollo del modelo se obtuvieron cuatro niveles para categorizar la madurez:

- *Básico* – la madurez de la capacidad de la organización es baja y la confianza en sus resultados es escasa debido a que los datos producidos pueden proporcionar una descripción de la realidad inexacta, parcial o incompleta.
- *Respalddado* – la madurez de la capacidad de la organización es intermedia y los resultados prometen un grado aceptable de calidad y confiabilidad, aunque podrían no ilustrar de manera precisa y completa la realidad.
- *Gestionado* – la madurez de la capacidad de la organización es alta y ofrece una descripción completa y precisa del estado nacional.
- *Experto* – la madurez de la capacidad de la organización es muy alta, lo que proporciona un reflejo preciso y altamente confiable de la realidad a nivel nacional.

El modelo en desarrollo, además de los niveles de madurez y las dimensiones de estudio, incluye los procesos a evaluar dentro de la organización, que son los que prescriben la secuencia de actividades que las NSOs realizan para producir datos estadísticos. Los procesos fueron agrupados en fases y representan el ciclo de vida de los datos. Las fases y los procesos no prescriben el modelo de procesos de negocio que las NSOs deben implementar para la producción de datos, sino que representan un conjunto de procesos genéricos que puede mapearse con los procesos específicos que cada NSO ejecuta para producir datos en el contexto de la Agenda 2030. El ciclo de vida diseñado para este modelo se compone de las siguientes cuatro fases:

- *Recolección* – comprende los procesos que se ejecutan para adquirir los datos.
- *Procesamiento* – agrupa los procesos que procesan los datos recolectados y validan los resultados hasta alcanzar un estado en el que puedan publicarse.
- *Utilización* – incluye los procesos relacionados con la difusión de los resultados además de la vinculación de los productos producidos con los consumidores.
- *Impacto* – abarca los procesos llevados a cabo para fomentar la utilización de los datos para generar impacto y crear cambios.

El modelo resultante es el resultado de la intersección de las fases con las dimensiones, segmentadas por los niveles de madurez, como se ilustra en la Figura 3.

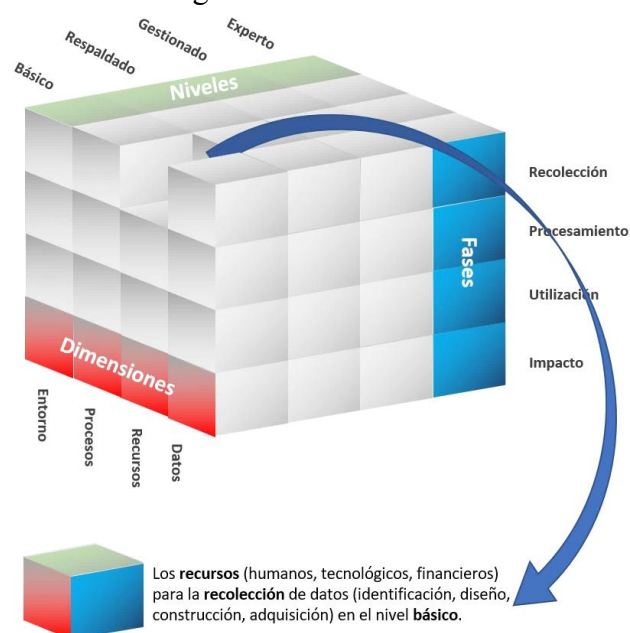


Figura 3: Arquitectura del Modelo

4.2 Contenido

En esta fase de desarrollo se identifica qué se debe medir en la evaluación de madurez y cómo se va a medir. El objetivo que se persigue en la definición del contenido es lograr componentes y subcomponentes del dominio que sean mutuamente excluyentes y colectivamente exhaustivos. Por tratarse de un dominio conocido, se realizó la identificación inicial de los componentes a través de una revisión de la literatura, la cual se mejoró a través de discusiones entre pares y entrevistas con expertos en el dominio.

De los ejercicios de definición de contenido, las cuatro dimensiones de análisis se subdividieron en 20 subdimensiones:

- *Entorno* – los factores que determinan la madurez del entorno de una NSO son su independencia profesional, un mandato claro y respaldado por el marco regulatorio, su ubicación dentro del NSS y una estructura interna que favorezcan su independencia y el acceso a los recursos necesarios, y un contexto que favorezca la colaboración dentro del ecosistema.
- *Recursos* – entre los recursos fundamentales para la producción de datos estadísticos confiables se encuentran los recursos humanos, tecnológicos, económicos y financieros, y físicos.
- *Datos* – los atributos que definen la calidad de los productos estadísticos y que determinan la madurez de una NSO para producir datos para los indicadores de los SDGs son su capacidad para gestionar la calidad de los datos; su confidencialidad; relevancia y vigencia; precisión y confiabilidad; puntualidad y oportunidad; coherencia, comparabilidad y consistencia; accesibilidad y claridad; así como su origen y nivel de desagregación.
- *Procesos* – esta dimensión evalúa la madurez de la capacidad de las técnicas y metodologías que emplea la NSO para la producción de datos estadísticos para los indicadores de los SDGs; su transparencia, imparcialidad y objetividad; la forma en que una NSO comparte e intercambia datos para, entre otras cosas, reducir la carga que se genera sobre los productores; y sus prácticas para mejorar de manera continua.

Similarmente, el contenido de las cuatro fases del ciclo de vida de la producción de datos estadísticos fue detallado a través de 14 procesos que se describen a continuación y se grafican en la Figura 4:

- *Recolección* – incluye los procesos de identificación de la necesidad de nuevos datos estadísticos, el diseño de cómo se conseguirán, la construcción de los artefactos y las herramientas necesarias para la recolección, y la adquisición de los datos necesarios.

- *Procesamiento* – contiene los procesos para el procesamiento de los datos recolectados en la fase anterior, su análisis, y la evaluación de los resultados para garantizar que sean adecuados para su propósito antes de ser difundidos.
- *Utilización* – consiste de la publicación, difusión y conexión que cubren el espectro de actividades que van desde la preparación de la difusión hasta su lanzamiento.
- *Impacto* – los procesos en esta fase incluyen actividades que promueven la utilización de datos para la toma de decisiones, la influencia en las decisiones basadas en datos y la reutilización y combinación de datos para la creación de nuevo conocimiento. Los procesos en esta fase suelen no existir entre los procesos tradicionales de las NSOs, pero son de gran importancia en el contexto de la Agenda 2030 y forman parte de la transformación y modernización que se persigue en las NSOs.

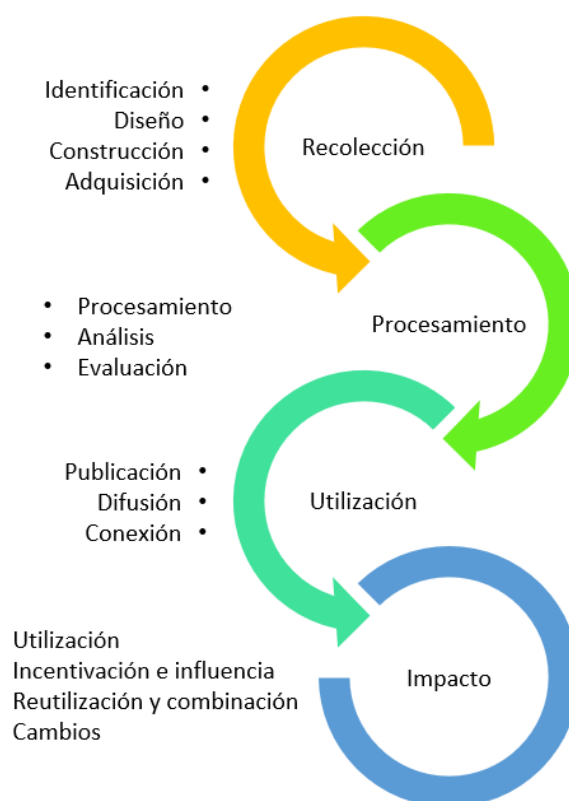


Figura 4: Fases y Procesos del Ciclo de Vida de los Datos

5. VALIDACIÓN

Para servir a sus objetivos, los modelos de madurez deben estar suficientemente validados; en caso contrario, la idoneidad y utilidad de un modelo de madurez sin ninguna aplicación ni validación resulta dudosa.

Informada por los objetivos de investigación, el diseño de la investigación y el cuerpo de conocimiento (*body of knowledge*), los ejercicios de validación realizados tuvieron como objetivo determinar si el modelo propuesto resulta adecuado para su uso (*fit-for-use*) y adecuado y útil para su propósito (*fit-for-purpose*). Mientras que el cuerpo de conocimiento informó los criterios generales para la validación, los objetivos particulares de la investigación y su diseño prescribieron los criterios específicos para la validación. La validación se realizó utilizando una combinación de múltiples métodos para evaluar los distintos criterios de validación. Para determinar la robustez del modelo se realizó una prueba de concepto (*proof of concept*) donde se evaluaron tanto su estructura como su contenido, estudiando no sólo el modelo resultante sino también cómo fue desarrollado. Los criterios para determinar la robustez del modelo fueron:

- *completitud* – se evaluó si contempla los principios que dictan la producción de estadísticas oficiales y los principios que rigen la calidad de los datos en el contexto de la Agenda 2030, y si incluye las áreas de evaluación relevantes consideradas por otros instrumentos similares;
- *correctitud* – fue evaluada a través de principios de los diseño que prescriben pautas para determinar la calidad de los modelos de madurez y de otros principios generales que evalúan la calidad de los artefactos desarrollados en el marco de la DS;
- *rigurosidad* – se analizó la instanciación del proceso de desarrollo y si fue desarrollado siguiendo estándares y buenas prácticas.

Para validar si el modelo satisface el propósito para el cual fue construido se realizó una prueba de valor (*proof of value*) donde se evaluó si cumple con los objetivos planteados y con los requerimientos preestablecidos, y si agrega valor tanto al dominio para el que fue desarrollado como a la base de conocimiento en el área.

Los criterios utilizados para estudiar la utilidad del modelo desarrollado fueron:

- *aptitud* – se evaluó si tiene la capacidad de cumplir adecuadamente la función para la que fue construido;
- *idoneidad* – se investigó si satisface los requerimientos planteados durante su concepción en las etapas de alcance y diseño;
- *originalidad* – se estudió si aporta nuevo conocimiento o nuevos instrumentos a la producción de datos para los indicadores de los SDGs y si agrega valor a la base de conocimiento en modelos de madurez y a la confiabilidad y calidad de la producción de datos estadísticos.

Los métodos de validación se resumen en la Tabla 1 e incluyeron tanto trabajo de escritorio como la participación de expertos de dominio y expertos en modelos.

Tabla 1: Criterios y Métodos de Validación

Prueba	Criterios	Métodos
Concepto	Completitud	- Comparaciones/Mapeos - Entrevistas a expertos
	Correctitud	- Principios de diseño DS - Principios de diseño CMMs
	Rigurosidad	- Trabajo de escritorio (rigor de aplicación del proceso de desarrollo)
Valor	Aptitud	- Entrevistas a expertos
	Idoneidad	- Trabajo de escritorio (validación de conformidad de requerimientos)
	Originalidad	- Literatura

Los ejercicios de validación realizados fueron satisfactorios, determinando tanto que el modelo propuesto es confiable como que puede ser utilizado para el objetivo para el que fue desarrollado. Adicionalmente, se observó que muchas de las limitaciones identificadas no son propias del modelo, sino que tienen origen en el ecosistema donde el modelo va a utilizarse. El modelo se enfoca específicamente en las NSOs y, si bien considera entre sus dimensiones el entorno en el cual se encuentran inmersas, los NSS son muy complejos y tienen un gran impacto en los resultados que se pueden lograr dentro de ellos. Una NSO madura dentro de un NSS inmaduro probablemente no pueda cumplir con los niveles de calidad y confiabilidad esperados.

Adicionalmente, muchos de los desafíos identificados por los expertos y especialistas (marco legal y regulatorio, financiamiento, interferencia política, contratación y retención de personal calificado) no son exclusivos de las NSOs sino que son desafíos comunes a muchas de las organizaciones del sector público.

Se observó también que, si bien varios de los expertos consultados coincidieron en que la introducción del modelo puede ser beneficiosa para la mejora de la calidad y la confiabilidad de los datos producidos para la Agenda 2030, algunos también resaltaron que las NSOs se encuentran bajo mucha presión de tiempo y reciben muchas demandas (muchas veces de los mismos datos por parte de distintos organismos) y que, por lo tanto, la introducción de más procesos puede no ser bien recibida o puede no cumplir con los beneficios esperados.

Otra de las observaciones fue que, por sí sólo, el modelo no puede solucionar todos los desafíos de datos que presenta la Agenda 2030 pero que, combinado con otras herramientas y otros esfuerzos, puede contribuir a lograr confianza en los datos lo que a su vez permita la toma de decisiones basada en evidencia confiable.

6. DISCUSIÓN

De la realización de este trabajo de investigación se lograron muchos descubrimientos que pueden aportar valor a otros trabajos que se realicen en el dominio. A continuación se presenta una pequeña discusión los más relevantes.

Se observa que existe una necesidad manifiesta de información confiable dentro de la comunidad estadística internacional y se han realizado numerosos esfuerzos con el fin de garantizar la calidad y precisión de los datos. Sin embargo, el proceso de la Agenda 2030 es largo y la tecnología cambia rápidamente, afectando directa e indirectamente las vidas de los seres humanos y, a su vez, los datos que se producen. Por lo tanto, la confiabilidad debe ser salvaguardada por organizaciones sólidas y maduras que sean independientes de sus empleados y de las administraciones actuales y futuras. Para este fin, los NSSs deben ser empoderados para adaptarse rápida y fácilmente a las nuevas realidades de los datos.

Todos los países, independientemente de su grado de avance y su nivel de desarrollo, pueden beneficiarse de soluciones que ayuden a desarrollar sus capacidades, como el CMM propuesto en esta tesis. Mientras que los países más desarrollados suelen disponer de mayores recursos para innovación y mejoras y, por lo tanto, suelen liderar el camino y marcar tendencias, los países en vías de desarrollo se puede beneficiar enormemente de los esfuerzos y la experiencia de aquellos que ya recorrieron ese camino. Cumplir los objetivos de la Agenda 2030 no implica una competencia entre países y sólo puede lograrse si cada país logra alcanzar sus metas y cumplir sus objetivos. Uno de los principios y creencias de los SDGs es que “los Estados Miembros de las Naciones Unidas trabajan juntos con un alto nivel de cooperación para mejorar las circunstancias de todas las personas en el mundo y los ubican en el centro del desarrollo futuro” [17, p. 1].

Las prácticas y soluciones provenientes del sector privado deben ser analizadas y adaptadas cuidadosamente ya que sus prioridades y objetivos pueden ser distintos. A manera de ejemplo, mientras que los indicadores de desarrollo prestan especial atención a la inclusión y el respeto por la privacidad de los individuos y sus comunidades, las soluciones del sector privado muchas veces se centran en otras prioridades. El modelo propuesto en este trabajo está definido para los SDG en particular, y para los indicadores sociales para el bien público en general.

Existen numerosos esfuerzos para el control de indicadores sociales y se puede (y se debería) sacar ventaja de ellos. Entre ellos se destacan las grandes inversiones que se han realizado para mejorar los datos para el control y reporte de los Objetivos de Desarrollo del Milenio (Millennium Development Goals, MDGs) [18]. De manera similar, algunos Estados Miembros de las Naciones Unidas reportan datos sobre derechos humanos desde hace más de una década en conformidad con las Revisiones Universales Periódicas (Universal Periodic Reviews, UPR) [19]. Todos estos esfuerzos, y en particular sus resultados, deberían ser estandarizados y considerados para desarrollar las sinergias que puedan facilitar.

Los datos deben incluir a todos y deben ser útiles para todos. La tendencia muestra que las empresas y los gobiernos confían cada vez más en los Grandes Datos (*Big Data*) y los datos analíticos asociados. Mientras que las empresas utilizan grandes datos para informar las decisiones y la estrategia empresarial, los gobiernos los utilizan para proporcionar una mejor prestación de servicios y participación ciudadana [20]. Enfoques complementarios como el de Datos Pequeños (*Small Data*) – en el que los datos, en lugar de ser agregados, se procesan en la misma unidad en la que se tomaron las muestras [21] – son importantes para asegurar que nadie sea excluido. El modelo propuesto en esta tesis integra ambos enfoques para promover la inclusión.

7. CONCLUSIONES

Este documento presentó el trabajo de investigación que se desarrolló con el fin de mejorar la calidad de los datos estadísticos producidos por los ecosistemas nacionales de datos para el seguimiento y control del progreso en el cumplimiento de la Agenda 2030.

Para responder la primera pregunta de investigación (P1) se identificaron, estudiaron y clasificaron los instrumentos utilizados para la gestión de la calidad de los datos oficiales para desarrollo. Para poder realizar tal estudio, primeramente se definió un marco conceptual de desarrollo sostenible que brinda los fundamentos conceptuales para todo el trabajo.

Con el fin de responder a la pregunta P2 se estudiaron y describieron los ecosistemas de producción de estadísticas oficiales para el seguimiento y control del progreso en el cumplimiento de la Agenda 2030. Se definió un modelo conceptual que describe los ecosistemas nacionales y globales de producción y reporte de datos que informan el estado de cumplimiento de las metas y los objetivos de los SDGs. Además, para enriquecer la comprensión del ecosistema global de control y reporte se realizó un mapeo a nivel de los indicadores que estudió cómo los indicadores de los instrumentos de medición de otros dominios (en este caso, Gobierno Digital) pueden proveer datos a los indicadores de los SDGs.

Para responder la tercera y última pregunta (P3) se plantearon dos objetivos de investigación que consistían en diseñar y validar un modelo para la producción de datos que permita a las NSOs avanzar gradualmente en la producción confiable de datos de calidad que sirvan para medir el cumplimiento de las metas de los SDGs. El cumplimiento de ambos objetivos se logró a través del desarrollo de tres actividades que incluyeron el diseño y desarrollo del modelo, la validación de su calidad y la utilidad, y el desarrollo un marco conceptual de modelos de madurez que consistió en un relevamiento extensivo y exhaustivo de la literatura existente sobre modelos de madurez en general y modelos de madurez de la capacidad en particular.

Las principales aportes de esta tesis son el CMM para la producción de datos para los indicadores de los SDGs y su respectiva validación. Pero además, este trabajo de investigación realizó las siguientes contribuciones:

- Una clasificación de los instrumentos relevantes para la calidad en la producción de datos estadísticos.
- Una matriz de alineación de indicadores que permite determinar la similitud entre indicadores de desarrollo.
- Un conjunto de indicadores de Gobierno Digital que aportan datos y metodologías para 10 indicadores de los SDGs.
- Un conjunto de recomendaciones para reutilizar trabajos de medición en áreas relacionadas con el desarrollo sostenible para la medición de los SDGs.
- Un conjunto de recomendaciones para utilizar modelos de madurez de la capacidad como instrumentos para estandarizar la producción de datos estadísticos.

Todas estas contribuciones fueron compartidas mediante publicaciones y otras actividades académicas y de difusión. Los resultados de esta investigación sirven además para establecer un marco de referencia para investigaciones posteriores y como plataforma para continuar con las mejoras en la calidad y confiabilidad de los datos que permiten la toma de decisiones de manera informada y sustentados por evidencias. Posibles líneas de investigación futuras incluyen el estudio de la adaptabilidad del modelo para ser utilizado en otros procesos y dominios.

8. BIBLIOGRAFÍA

- [1] United Nations University Institute on Computing and Society, “Last Mile Data Enablement and Collaboration, and Building Trust in Indicators Data.” <http://cs.unu.edu/research/sdgs/> (accessed Mar. 13, 2018).
- [2] Universidad Nacional del Sur, “Líneas de investigación – LISSI DCIC CIC.” https://lissi.cs.uns.edu.ar/?page_id=26 (accessed Mar. 18, 2018).
- [3] Universidad Nacional del Sur, “Departamento de Ciencias e Ingeniería de la Computación.” <https://cs.uns.edu.ar/home/> (accessed Mar. 18, 2018).
- [4] United Nations Statistics Division, “United Nations Statistical Commission,” 2018. <http://unstats.un.org/unsd/statcom> (accessed Jan. 31, 2017).
- [5] United Nations, “Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development,” 2017. [Online]. Available: https://unstats.un.org/sdgs/indicators/GlobalIndicatorFramework_A.RES.71.313Annex.pdf.
- [6] Independent Expert Advisory Group on a Data Revolution for Sustainable Development, “A World that Counts: Mobilising the Data Revolution for Sustainable Development,” 2014. [Online]. Available: <http://econpapers.repec.org/RePEc:ecr:col013:40319>.
- [7] Business for 2030, “Metrics & Indicators.” <http://www.businessfor2030.org/metrics-indicators/> (accessed Jan. 31, 2017).
- [8] H. A. Simon, *The Sciences of the Artificial (3rd Ed.)*. Cambridge, MA, USA: MIT Press, 1996.
- [9] A. R. Hevner, “A Three Cycle View of Design Science Research,” *Scand. J. Inf. Syst.*, vol. 19, no. 2, pp. 87–92, 2007, doi: <http://aisel.aisnet.org/sjis/vol19/iss2/4>.
- [10] E. Ferrace, *Action Research*. 2000.
- [11] R. Wendler, “The maturity of maturity model research: A systematic mapping study,” *Inf. Softw. Technol.*, vol. 54, no. 12, pp. 1317–1339, 2012, doi: 10.1016/j.infsof.2012.07.007.
- [12] “Modernisation Maturity Model (MMM) - Roadmap for Implementing Modernstats Standards - UNECE Statistics Wikis.” <http://www1.unece.org/stat/platform/pages/viewpage.action?pageId=129172266> (accessed Feb. 28, 2017).
- [13] UNDP, “Rapid Integrated Assessment (RIA),” 2017.
- [14] United Nations Development Group, “Common Country Assessment,” 2002. [Online]. Available: <http://www.sprep.org/att/IRC/eCOPIES/Countries/Vanuatu/3.pdf>.
- [15] T. De Bruin, R. Freeze, U. Kaulkarni, and M. Rosemann, “Understanding the Main Phases of Developing a Maturity Assessment Model,” in *Australasian Conference on Information Systems (ACIS)*, 2005, pp. 8–19, doi: 10.1108/14637151211225225.
- [16] D. L. Moody and G. G. Shanks, “What Makes a Good Data Model? Evaluating the Quality of Entity Relationship Models,” in *Proceedings of the 13th International Conference on the Entity-Relationship Approach*, 1994, pp. 94–111, [Online]. Available: <http://dl.acm.org/citation.cfm?id=647516.727205>.
- [17] “Sustainable Development Goals, Beliefs, and Principles.” <https://www.agora-parl.org/resources/aoe/sustainable-development-goals-beliefs-and-principles> (accessed Jun. 05, 2017).
- [18] United Nations, “The Millennium Development Goals Report,” 2015. doi: 978-92-1-101320-7.
- [19] United Nations Office of the High Commissioner for Human Rights (OHCHR), “Universal Periodic Review - A Practical Guide for Civil Society,” 2014.
- [20] M. Thinyane, “Investigating an Architectural Framework for Small Data Platforms,” in *Proceedings of the 17th European Conference on Digital Government (ECDG 2017), Lisbon, Portugal, 12-13 June, 2017*, pp. 220–227.
- [21] M. Best, “Small Data and Sustainable Development,” in *International Conference on Communication/Culture and Sustainable Development Goals: Challenges for a new generation.*, 2015, pp. 1–6.