

ACTA DE REUNIÓN N°1 AYUDA DE MEMORIA

Ciudad: Bogotá

Lugar: Sesión virtual, plataforma Teams

Tema: Reacciones frente a recomendaciones
brindadas línea de interoperabilidad y temas a
trabajar 2021

Hora: 10:00 a.m. a 12:10 a.m.

Fecha: 12/05/2021

Dependencia responsable: Secretaría Técnica
CASEN (DIRPEN)

Participantes

Miembros de la Sala Especializada de Modernización Tecnológica

León Darío Parra Bernal
Jorge Andrés Gallego Durán
Valérie Gauthier Umaña

Oficina de Sistemas – Profesional Especializado
Marly De Moya Amaris

Coordinadora Grupo Interno de Trabajo (GIT)
Planificación y Articulación Estadística (DIRPEN)
Mónica Patricia Pinzón Torres

Asesor de la Dirección General
Juan Sebastián Ordoñez

Departamento Administrativo Nacional de Estadística (DANE)

Director General
Juan Daniel Oviedo

Coordinador Gestión Organizacional- oficial de
seguridad de la información
David González

Directora Técnica de Regulación, Planeación,
Estandarización y Normalización (DIRPEN)
Julieth Alejandra Solano Villa

DIRPEN-Contratista
Mateo Cardona

Asesora Dirección General y DIRPEN
Elizabeth Moreno Barbosa

Profesional GIT Planificación y Articulación
Ruth Constanza Triana Acuña

Profesional GIT Planificación y Articulación
Estadística (DIRPEN)
María del Pilar Pulido

Orden del día

1. Solicitud autorización para grabar la reunión
2. Verificación del quórum
 - I. León Darío Parra Bernal
 - II. Valérie Gauthier Umaña
 - III. Jorge Andrés Gallego Duran
 - IV. Julieth Solano
 - V. Marly Esther De Moya Amaris
 - VI. Juan Sebastián Ordoñez
3. Propuesta de valor 2021 (Julieth Solano y expertos – 10 minutos)
4. Reacciones del DANE frente a las recomendaciones realizadas por los expertos de la Sala Especializada de Modernización Tecnológica en cuanto a la línea de interoperabilidad, presentación propuesta de líneas a trabajar 2021 y las acciones de divulgación – Webinar y notas estadísticas (Julieth Solano - Marly De Moya – 10 minutos)
5. Aportes expertos frente a temáticas en el punto anterior (Expertos de la sala – 20 minutos)
6. Propuesta de temas a abordar en la línea “Seguridad de la información” (Marly de Moya, Mateo Cardona, David González, Julieth Solano, Juan Sebastián Ordoñez – 30 minutos)
7. Aportes de expertos frente a las temáticas en el punto anterior (Expertos de la sala – 30 minutos)
8. Identificación de actores del ecosistema de datos relacionados a la línea de “seguridad de la información” invitados a participar en la sala (Integrantes de la sala – 10 minutos)
9. Conclusiones y cierre (Julieth Solano - 10 minutos)

Desarrollo

Objetivo

Presentar las consideraciones del DANE, las líneas de trabajo propuestas para el período 2021, acciones de divulgación e iniciar el desarrollo de la línea de investigación “Seguridad de la información”. Asimismo, dar a conocer el estado de avance del trabajo que viene realizando el DANE en esta materia, a fin de recibir realimentación por parte de los expertos y fortalecer estos procesos.

1. Solicitud de autorización para grabar la reunión

Se inició la grabación de la sesión con previa autorización de los participantes.

2. Verificación del quórum

Se verificó el quórum tanto de los expertos de la Sala, como del Departamento Administrativo Nacional de Estadística (DANE).

3. Propuesta de valor 2021 (Julieth Solano y expertos – 10 minutos)

Julieth Solano inició la reunión indicando que de acuerdo a las lecciones aprendidas y al balance realizado del desarrollo de las reuniones de la sala el año pasado, este año se trabajaría con una dinámica más fluida y flexible para el desarrollo de los temas y de esta manera poder generar productos y resultados interesantes de acuerdo al aporte de cada integrante de la sala.

Asimismo señaló que como propuesta de valor se busca proponer nuevos temas en analítica y seguridad y promover productos de divulgación y un documento que recoja el trabajo realizado, así como espacios tipo webinar para temas de seguridad en la información y del ecosistema de datos en cuanto a producción estadística. Por otra parte se pretende la articulación con otras salas como la de Salud, bienestar social y demografía en donde se está trabajando el tema étnico y se han identificado desafíos asociados a la recolección de información; por su parte, en la Sala de Gobierno, seguridad y justicia se está trabajando el indicador de femicidio y se presenta el desafío de acceso a la información; en la sala de Geografía, medio ambiente y ordenamiento territorial hay iniciativas el abordar el registro estadístico de inmuebles que permita conectar con otros registros como el de empresas y personas; en la sala de Economía se está trabajando en la revisión de pares del registro administrativo de componente tecnológico. Así mismo, se busca la articulación con el CAD (Comité de administración de datos) para que la sala tenga mayor visibilidad.

Frente a los temas abordados, León Darío Parra manifestó estar de acuerdo y más aún porque juega un papel importante la analítica de datos, la generación de nuevas estadísticas, así como la importancia de la realización de espacios como el webinar vinculando las otras salas.

Jorge Gallego señaló que hay una propuesta de valor interesante y le llama la atención que en otras salas hay productos más concretos y específicos a abordar, por lo tanto, es importante establecer si para la presente sala se busca trabajar en temas generales o específicos.

Por su parte Valerie Gauthier mencionó la importancia de traer a la sala invitados como Yeimy Cano y Daniel Díaz quienes complementarían los procesos a desarrollar en las temáticas señaladas.

4. Reacciones del DANE frente a las recomendaciones realizadas por los expertos de la Sala en cuanto a la línea de interoperabilidad, presentación propuesta de líneas a trabajar 2021 y las acciones de divulgación – Webinar y notas estadísticas (Julieth Solano - Marly De Moya – 10 minutos)

Marly De Moya realizó un resumen de los temas abordados en las reuniones del año pasado donde indica que de acuerdo a la línea de Big Data e interoperabilidad, se generaron instrumentos para revisión de las recomendaciones. Señala que se recibieron recomendaciones a nivel estratégico, las cuales vinculaban: Mapeo capacidades institucionales e identificación de niveles de madurez de las entidades públicas, armonización de prioridades y necesidades, homogenización y estandarización de sistemas de información, formación al ciudadano y tratamiento datos personales. Frente a estos, se espera la propuesta de modelos, lineamientos y buenas prácticas a implementar, basados en experiencias y casos de éxito internacionales. Adicionalmente, específicamente para el tratamiento de Datos personales, se esperan propuestas de mecanismos tecnológicos que garanticen la seguridad de los datos personales en los diferentes flujos de información que existen en el DANE, incluyendo operaciones de interoperabilidad y sobre servicios en la nube y lago local, entre otros.

En cuanto a las recomendaciones a nivel táctico se señaló: Articulación entre entidades, espacios de diálogo, modelo de aprendizaje colaborativo, mapeo de necesidades regionales y locales, metodología de evaluación de calidad para transferencia de datos y desarrollar pilotos. De acuerdo con esto, se espera un análisis de los proyectos pilotos realizados y sus resultados frente a experiencias y casos de éxito internacionales (tanto en los proyectos de interoperabilidad, como en los proyectos de Big Data) y recomendaciones sobre ajustes para su replicación de acuerdo a las necesidades regionales y locales.

5. Aportes expertos frente a temáticas en el punto anterior (Expertos de la sala – 20 minutos)

No hay reacciones frente al punto anterior

6. Propuesta de temas a abordar en la línea “Seguridad de la información” (Marly de Moya, Mateo Cardona, David González, Julieth Solano, Juan Sebastián Ordoñez – 30 minutos)

Mateo Cardona inició la presentación abordando la política de gobierno de RRAA y fuentes alternativas donde se busca la ejecución y aplicación de la autoridad sobre la gestión de los datos en su ciclo de vida y de los activos relacionados con estos teniendo en cuenta el principio de sistema y el principio de coherencia, desde el mes de diciembre del 2020 se busca tratar de generar una política que permita el desarrollo de estos principios, así como el aprovechamiento de RRAA.

La política anteriormente señalada busca generar un seguimiento integral al RRAA y fuentes alternativas en todo su ciclo de vida al interior de la entidad y direccionar la implementación de un marco de aseguramiento de la calidad. A la luz del proceso de revisión de pares del SEN, se está piloteando el marco de calidad armonizado del DANE para RRAA, como una primera implementación del lineamiento de gestión de calidad de la política de gobierno de RRAA y fuentes alternativas.

Por otro lado, se trabaja en la gestión de activos por medio de la identificación, clasificación y valoración de activos de información del DANE y el FONDANE, de acuerdo con su criticidad y a su nivel de confidencialidad; en el control de acceso permitiendo el acceso sólo a los usuarios autorizados por los líderes de proceso; en la seguridad de las operaciones asegurando que las operaciones de procesamiento de información sean correctas y seguras dentro de las instalaciones y la gestión de incidentes de seguridad de la información donde se busca asegurar, con un enfoque coherente y eficaz la gestión de incidentes de seguridad de la información.

Por su parte, David González señaló que, en la política general de seguridad y privacidad de la información, tiene como objetivo conservar, salvaguardar y proteger la información administrada por el DANE y el FONDANE para preservar la confidencialidad, integridad y disponibilidad de la información. Por esta razón se trabaja con una política de activos críticos de información estadística en donde cualquier RRAA que llegue al DANE se salvaguarda evitando su vulnerabilidad y la no pérdida de consistencia e integridad, lo anterior, se acompaña de guías, procedimientos, manuales; por otra parte, hay controles en seguridad física y entorno donde se busca que ninguna persona acceda al DANE tome información salvaguardada.

De acuerdo a lo expuesto anteriormente Jorge Gallego preguntó sobre cómo está la política de gestión de RRAA para uso de otros actores.

Juan Sebastián Ordoñez indicó que se empezó con revisión de pares, fuentes y usuarios para entender la necesidad puntual de algunas variables y que luego en el CAD seguramente se mirará cómo conectar al proveedor de los datos y usuario para realizar el intercambio, sin embargo, se presenta el desafío con datos sensibles por temas de privacidad de la información, entonces se debe buscar a través del CAD como garantizar la seguridad en la información.

Jorge Gallego preguntó si el CAD sería el intermediario entre generadores y usuarios, a lo cual Juan Sebastián Ordoñez confirmó que en efecto así es.

Por su parte, Julieth Solano comentó que la principal fuente de datos son los RRAA y hay diferentes niveles de acceso a estos datos y a microdatos con identificadores. Por ejemplo, al solicitar al DANE la PILA por medio de revisión de pares se busca poder hacer revisiones de calidad de datos y usos a los mismos.

Juan Sebastián Ordoñez complementó lo anterior indicando que se busca velar por la coherencia con un enfoque de calidad tanto en el input como en el output.

Marly De Moya habló sobre el Modelo de Gestión de datos de requerimientos de seguridad en donde el gobierno de datos vincula estrategias, valoración de los datos, principios ética y administración de roles y responsables con el fin de alcanzar un cambio cultural, seguido a esto en la gestión del ciclo de vida se tiene en cuenta el plan y diseño (arquitectura y modelo de datos), el uso y la mejora (operaciones y almacenamiento de datos, integración de datos e interoperabilidad, gestión de datos maestros, bodegas de datos, almacenamiento y Big Data y gestión de datos de referencia), así como habilitar y mantener (inteligencia de negocios, uso de datos maestros, gestión documental, ciencia de datos, monetización de datos y analítica predictiva) y finalmente en las actividades fundacionales se encuentra la gestión de riesgo de datos, la gestión de metadatos y la gestión de calidad de datos.

En cuanto al tema de interoperabilidad, señaló sobre el estado de avance de los proyectos implementados a través de la PDI 2020-2021, de acuerdo con los niveles de certificación del lenguaje común de Mintic en donde se ha presentado el servicio de intercambio de información de estadísticas del sector de hidrocarburos y minería alcanzando un nivel de 3; en el servicio de intercambio de información del registro sanitario de bebidas alcohólicas se alcanzó un nivel de 2 y está en proceso de finalización el presente año.

Continuó su intervención Marly De Moya indicando los proyectos de interoperabilidad GIT gestión de datos con geoportales en el 2020 donde se encuentra el servicio de interoperabilidad de indicadores EGIT y el servicio de interoperabilidad de indicadores de SIENA contando con una base de datos construida e interoperabilidad implementada.

Como retos de seguridad para la implementación de ejercicios de intercambio de información e interoperabilidad se encuentran:

- Políticas de seguridad de la información debido a que no todas las entidades manejan los mismos criterios de seguridad
- Organización en donde se requiere establecer planes progresivos que permitan la aplicación de los controles diseñados para la recolección y difusión de la información según los criterios de las políticas de seguridad de la información
- Entregas de información debido a que se manejan múltiples formatos y archivos que dificultan mantener una única estrategia de seguridad sobre los datos.
- Cumplimiento donde se busca que todos los intercambios de información con la entidad deben cumplir con los requisitos legales que garantice recepción y entrega
- Gestión de incidentes debido a que, al tener múltiples mecanismos de entrega de información, se dificulta el seguimiento y control frente eventos o incidentes de seguridad

Teniendo en cuenta el lago de datos local se encuentran los siguientes pilotos en curso:

- Proyecto de educación D4N para medir la accesibilidad del hogar a la escuela del estudiante mediante la exploración de los registros estadísticos como insumo y la caracterización del impacto en las tasas de abandono escolar y de asistencia.
- Anonimización – CEED y DSCN sector carbón para la transformación de datos individuales para la mitigación de riesgos.

- Cruces difusos para disminución de tiempos de procesamiento para mejorar la oportunidad en la entrega de la información entre los cruces de variables como razón social, representante legal, entre otras.

Como potenciales pilotos se encuentran:

- Automatización de la actividad económica CIIU rev. 4 AC
- Detección de anomalías en el censo económico
- Automatización de procesos de calidad
- Imputación de datos con cruce difuso
- Educación para el fortalecimiento del sistema de matrícula de estudiantes de educación básica y secundaria del Sistema Nacional de Información para la educación superior

Julieth Solano comentó que en el tema de educación se busca la medición de distancias, a partir de la infraestructura tecnológica se podrá profundizar en la integración de datos y analizar el tema de deserción.

Juan Daniel Oviedo indicó que se busca poner en servicio los nuevos esquemas de gobierno de datos, pero se presenta como desafío que el país no ha logrado ver a todos los sistemas de ciencia y tecnología integrados; el objetivo de trabajar con MinEducación es prestar un servicio y dejarle al país armado un modelo integrado con los registros de mercado laboral como el RELAB, para tener las trayectorias completas de educación y desempeño profesional y presentarle al país el primer panel de vías laborales del país para ver los efectos de las políticas en este tema de educación.

Marly De Moya señaló los retos para la implementación de ejercicios en el lago de datos y nuevas técnicas de analítica de datos donde se encuentran:

- Restricciones en los cambios a los paquetes de software
- Ambiente de desarrollo seguro
- Datos de prueba y producción
- Redundancias
- Registro y seguimiento
- Disposición segura o reutilización de equipos

En el lago de datos en Nube, se encuentra en DANE realizando evaluación de partners con Microsoft Azure, Google Cloud, Oracle Cloud, AWS.

En cuanto a potenciales pilotos en la Nube se encuentran:

- Detección de anomalías en el censo económico
- Indicadores sector carbón
- Pobreza monetaria y multidimensional
- ODS

Finalmente, se indican los retos de seguridad para la implementación de ejercicios en el lago de datos en la nube y nuevas técnicas de analítica de datos, los cuales son:

- Ambiente de desarrollo seguro de procesamiento en la nube
- Datos de prueba y producción
- Registro y seguimiento
- Disposición segura o reutilización de equipos

7. Aportes expertos frente a temáticas en el punto anterior (Expertos de la sala – 30 minutos)

León Darío Parra señaló que los retos y proyectos presentados son ambiciosos dentro de los cuales hay 5 a nivel interno y 5 a nivel externo, con un alcance amplio, teniendo en cuenta los problemas que suelen presentarse. Formuló la pregunta de cuál será el alcance de la sala, pues hay varias entidades con las que se va a hacer interoperabilidad que son externas y se pueden presentar fugas de información. También preguntó qué lineamientos se plantearían para que estas fugas no se presenten, y en cuanto a la función de los proyectos, qué plan de contingencia se establece para prevenir fugas de información por interacción entre bodega local de datos y otras empresas que no utilizan estándar CDMX. También indaga sobre cómo manejar la presencia de errores y ruido.

Julieth Solano comentó que el alcance si es ambicioso porque se estaba conceptualizando y definiendo esquemas de trabajo. Actualmente ya se cuenta con datos para aplicaciones de proyectos, sin embargo, se presentan desafíos metodológicos, en cuanto a capacidades de las otras entidades, adicional que se debe reconocer el grado de madurez y por tal razón se busca fortalecer articulaciones con el CAD.

Marly De Moya indicó que las iniciativas que se tienen establecidas se pueden abordar, y mencionó el ejemplo de los proyectos de geoportales, en donde se realizan reunión con entidades y se cuentan con capacidades para llevarlos a cabo; adicionalmente otro ejemplo es el de las estadísticas vitales donde se trabaja desde el 2020 para definir modelos para interoperar. En cuanto a las debilidades se espera aportes de los expertos para subsanarlas.

Jorge Gallego preguntó sobre cuál es el papel de la sala relacionándolo con los temas expuestos, pues si se profundizará sobre la línea de seguridad, está un tema transversal a todos los proyectos y por lo tanto hay dos posibles caminos, uno transversal y uno específico.

Julieth Solano respondió que se cuenta con elementos de seguridad que pueden abordarse desde lo transversal y al conocerse la experticia de los expertos, se cuenta con ellos para abordar las necesidades del DANE. Adicional, se prefiere tomar como enfoque el de abordar los pilotos en un contexto específico.

Juan Daniel Oviedo manifestó que se busca un feedback por parte de expertos, para los proyectos a realizar, adicional que es interesante trabajar a través de los pilotos y de integración basada en casos, de acuerdo a perfiles específicos de los expertos con el fin de generar alternativas para tener valor agregado en los análisis de casos.

Valerie Gauthier indicó estar de acuerdo en abordar pilotos en un contexto específico, esto con el fin de aportar en la toma de decisiones basados en datos. Adicional pregunta si el CAD existe y quién o quiénes realizarán los proyectos, si universidades, o ya está el equipo establecido.

Julieth Solano comentó al respecto que los proyectos ya tienen datos y equipos con conceptualizaciones y propuestas; estos equipos deben conectarse a la sala con los expertos y a partir de esto realizar aportes en cuanto a criterios de resultado de modelos, se trabaja con capacidades In House, pero se está abierto para trabajo conjunto entre académicos y técnicos del DANE. En cuanto al CAD, indica que ya existe, pero no se ha instaurado, y que hay varios temas a trabajar y el rol de la sala con el CAD es de asesoría.

León Darío Parra preguntó si los aportes se pueden documentar en un artículo por temas asociados a procesos de investigación y de academia. Otra pregunta que formula es si el lago de datos local y de la nube cuentan con capacidad de almacenamiento de información, y cómo se ha pensado manejar este tema con las contrapartes.

Marly De Moya comentó que en cuanto al lago de datos, ya está implementado 50 teras de almacenamiento, y hay servicios para seguridad, servicio y almacenamiento, y la integración con lago de datos en la nube está en trámite de evaluación para selección de proveedores.

Julieth Solano comenta que la publicación conjunta es el enfoque que se busca para este año y de acuerdo con el desarrollo de la sala se mirará cuál será el alcance.

8. Identificación de actores del ecosistema de datos relacionados a la línea de “seguridad de la información” invitados a participar en la sala (Integrantes de la sala – 10 minutos)

Julieth Solano manifestó que se trabajará con un enfoque de aplicaciones y de pilotos, con los equipos de trabajo para articulación. Adicional se buscará generar un producto de divulgación como webinar.

9. Conclusiones y cierre (Julieth Solano - 10 minutos)

Julieth Solano indicó que se compartirán fichas de proyectos y resultados y a partir de la revisión, se trabajarán en las siguientes reuniones de la sala. Adicional, se plantea realizar la siguiente reunión para activar agenda de trabajo en 2 o 3 semanas y se enviará Doodle para disponibilidad de agenda.

León Darío Parra solicitó que en los documentos compartidos de los proyectos, se relacione el cronograma, hitos, en qué momento se encuentra el proyecto, retos o principales desafíos para que los expertos detecten y puedan aportar sobre cómo afrontarlos.

Juan Daniel Oviedo concluyó que al DANE le interesa avanzar y generar impacto desde el punto de vista de la posibilidad de que algunos de los casos sean inspiradores en la literatura académica

y que el DANE está abierto para trabajar de la mano con los expertos, si este trabajo llevará a involucrar a otros equipos, se vuelve más valioso para acercar a la academia en estos proyectos. Finaliza indicando que fue satisfactorio el desarrollo de la reunión y felicita al equipo.

Valerie Gauthier comentó que la Universidad del Rosario se encuentra realizando el lanzamiento de la maestría en matemáticas aplicadas y ciencias de la computación y que sería interesante trabajar de la mano con los estudiantes para la generación de los productos de la sala.

Juan Daniel Oviedo manifestó estar totalmente de acuerdo e invita a vincular estudiantes de los demás expertos.

Compromisos

- 1. Tarea:** enviar el acta de la sesión
Responsable: DANE
Fecha entrega: a más tardar 21 de mayo
- 2. Tarea:** envío Doodle para agendar segunda reunión
Responsable: DANE
Fecha entrega: 13 mayo
- 3. Tarea:** envío documentación proyectos a trabajar en la sala
Responsable: DANE
Fecha entrega: a medida que se abarque la temática

Próxima reunión:

Responsable de convocar: Secretaría Técnica CASEN

Fecha: Por confirmar

Expertos CASEN

DANE