

PROYECTO INTERNO SIN FINANCIAMIENTO O AUTOGESTIONADO PII-DETRI-2019-01
"Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de anonimato con enfoque en la preservación de utilidad de los datos"

En la ciudad de Quito D.M., a los veinte y seis días del mes de junio del año dos mil veinte, comparecen a la celebración de la presente Acta de Finalización del Proyecto Interno sin Financiamiento o Autogestionado **PII-DETRI-2019-01 "Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de anonimato con enfoque en la preservación de utilidad de los datos"**, por una parte, la **Ph.D. Alexandra Patricia Alvarado Cevallos** en calidad de **Vicerrectora de Investigación, Innovación y Vinculación** de la Escuela Politécnica Nacional, y por otra el **Ph.D. Luis Felipe Urquiza** en calidad de **Director del Proyecto Interno Sin Financiamiento o Autogestionado**, al tenor de lo siguiente:

1. ANTECEDENTES:

- a) Mediante Memorando Nro. EPN-DETRI-2018-1803-M del 19 de diciembre del 2018, el Jefe del Departamento de Electrónica, Telecomunicaciones y Redes de Información, solicita al Vicerrectorado de Investigación y Proyección Social, que se asigne código y se registre el proyecto "**Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de anonimato con enfoque en la preservación de utilidad de los datos**" propuesto por el Ph.D. Luis Urquiza.
- b) Mediante Memorando Nro. EPN-DIPS-2019-0055-M del 28 de enero del 2019, la Dirección de Investigación y Proyección Social, notifica al Jefe del Departamento de Electrónica, Telecomunicaciones y Redes de Información que el proyecto de Investigación Interno sin Financiamiento o Autogestionado del Ph.D. Luis Urquiza ha sido registrado con el código PII-DETRI-2019-01 y que su fecha de inicio es el 28 de enero de 2019.

2. DATOS GENERALES DEL PROYECTO:

Código de Proyecto	PII-DETRI-2019-01
Nombre del Proyecto	Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de anonimato con enfoque en la preservación de utilidad de los datos
Director del Proyecto	Luis Felipe Urquiza Aguiar
Colaboradores del Proyecto	Ana Rodríguez José Estrada
Departamento	Departamento de Electrónica, Telecomunicaciones y Redes de Información
Líneas de Investigación	Seguridad y Privacidad
Objetivo	Analizar y adaptar mecanismos de anonimización de datos, basados en privacidad sintáctica, con el fin de garantizar la usabilidad y la utilidad de los datos.
Duración del Proyecto	<ul style="list-style-type: none"> • Inicio: 28 de enero del 2019 • Fin planificado: 27 de enero del 2020 • Duración total: 12 meses
Entrega del Informe Final	15 de mayo de 2020

3. INFORME FINAL:

Mediante Memorando Nro. EPN-DETRI-2020-0907-M del 15 de mayo del 2020 el Ph.D. Luis Urquiza, Director del Proyecto PII-DETRI-2019-01, presenta el Informe Final del Proyecto Interno sin Financiamiento, mismo que es revisado por la Dirección de Investigación, y que se anexa y forma parte integrante del Acta de Finalización, cuyas conclusiones y productos generados son:

CONCLUSIONES:

- A lo largo del proyecto se corroboró la intuición de que la captura y utilización de las propiedades estadísticas de los microdatos (e.g., la dependencia estadística entre cuasi-identificadores y atributos confidenciales) cuando se construyen algoritmos de microagregación, permiten reducir la degradación de la utilidad empírica. Esto se evidencia claramente en nuestros experimentos con MDAV con dependencia estadística. Lastimosamente, la métrica híbrida planteada para evaluar su rendimiento resulta no ser un adecuado predictor de la utilidad práctica de datos microagregados, como sería de esperar. Sin embargo, valores altos de distorsión medidos con dicha métrica sí sugieren alguna correlación con las métricas de utilidad empírica.
- Se confirmó también que los algoritmos de microagregación k-anónima permiten preservar gran parte de la utilidad de los datos mientras protegen la privacidad de cada sujeto en grupos de k individuos. Las operaciones de clustering y promediado parecen contribuir a filtrar, normalizar o consolidar la información estadística dentro de los microdatos.
- Existe un margen interesante que explotar en los algoritmos de microagregación k-anónima propuestos en la literatura, en lo referente a su complejidad computacional. Esto se ha corroborado con las estrategias algorítmicas y algebraicas propuestas a raíz de este proyecto y que permiten reducir significativamente el tiempo de ejecución, en este caso, del algoritmo MDAV. Este esfuerzo es importante a corto plazo pues muchos de los algoritmos de protección de privacidad son inviables en la práctica, al tratar de mantener la utilidad de los datos, debido a su complejidad computacional incrementada por el creciente volumen de datos (big data).
- Finalmente, y en la misma línea de la conclusión anterior, se ha abordado el problema, natural, de la degradación de la utilidad de los datos como consecuencia de su distorsión para proteger la privacidad. El método propuesto transforma los datos de tal manera que, luego de ser microagregados, los clústers k-anónimos resultantes permiten la construcción de modelos de aprendizaje más efectivos, incrementando la utilidad de los datos con respecto a aquellos microagregados con el algoritmo MDAV original. Esto ocurre mientras se mantiene intacta la garantía de privacidad (k-anonimato).

En la era del big data y de la vigilancia masiva, en la que se necesita tanto la utilidad que se extrae de los datos como la privacidad que se ofrece a los sujetos de esos datos, abordar ese compromiso es ineludible. Este proyecto hace, en ese sentido, un aporte interesante.

PRODUCTOS:

- Artículo en formato de la Revista Politécnica: "Evaluación de la Degradación de Utilidad de Estrategias de Microagregación k-Anónima de Datos Numéricos"
- Artículo publicado: "The Fast Maximum Distance to Average Vector (F-MDAV): An algorithm for k-anonymous microaggregation in big data"; Ana Rodríguez-Hoyos, José Estrada-Jiménez, David Rebollo-Monedero, Ahmad Mohamad Mezher, Javier Parra-Arnau, Jordi Forné; Engineering Applications of Artificial Intelligence; <https://doi.org/10.1016/j.engappai.2020.103531>; Enero 2020

- Artículo enviado: "Preserving Empirical Data Utility in k -Anonymous Microaggregation via Linear Discriminant Analysis"; Ana Rodríguez-Hoyos, David Rebollo-Moneder, José Estrada-Jiménez, Jordi Forné, Luis Urquiza-Aguar; Engineering Applications of Artificial Intelligence.
- Artículo enviado: "A Unified Comparison of the Utility Degradation of Widely Used Strategies for k -Anonymous Microaggregation of Numerical Data"; Ana Rodríguez-Hoyos, José Estrada-Jiménez, David Rebollo-Monedero, Jordi Forné, Javier Parra-Arnau, Luis Urquiza-Aguar; Engineering Applications of Artificial Intelligence.
- Presentación a la comunidad politécnica: "Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de ANONIMATO con enfoque en la preservación de UTILIDAD de los datos"; José Antonio Estrada; febrero 2020.

4. LIQUIDACIÓN ECONÓMICA:

El Proyecto Interno sin Financiamiento o Autogestionado PII-DETRI-2019-01 no contó con asignación presupuestaria.

5. FINALIZACIÓN:

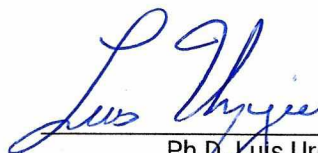
Con la presente Acta se declara finalizado y cerrado el Proyecto Interno sin Financiamiento o Autogestionado PII-DETRI-2019-01 "Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de anonimato con enfoque en la preservación de utilidad de los datos".

Para constancia de lo ejecutado y por estar de acuerdo con el contenido de la presente Acta, las partes libre y voluntariamente suscriben la misma, en tres ejemplares de igual contenido, tenor y valor legal.

Dado en la ciudad de Quito, D.M. a los veinte y seis del mes de junio del año dos mil veinte.



Ph.D. Alexandra Alvarado
Vicerrectora de Investigación,
Innovación y Vinculación



Ph.D. Luis Urquiza
Director del Proyecto
PII-DETRI-2019-01

np/cc

Recibido
30/09/2021
L. Urquiza

