



A. PROPUESTA PROYECTO DE INVESTIGACIÓN INTERNO SIN FINANCIAMIENTO

1. TIPO DE INVESTIGACIÓN:

Básica		Aplicada	X
--------	--	----------	---

2. **UNIDAD EJECUTORA** (*Departamento, Instituto o Estructura de Investigación*):
Departamento de Electrónica, Telecomunicaciones y Redes de Información (DETRI)

3. **LINEA(S) DE INVESTIGACIÓN:**
Seguridad y Privacidad

4. **TÍTULO DEL PROYECTO** (*mínimo 10 palabras*):
Protección de datos personales en Ecuador: Un enfoque desde las políticas de privacidad en Internet

5. **RESUMEN** (*máximo 200 palabras*)
Una política de privacidad es una declaración de intenciones frente al tratamiento de datos personales. Así, el análisis del contenido de este documento permite evaluar el nivel de compromiso de distintos actores con la protección de dichos datos. Además, este documento tiene efectos legales, comúnmente regulados por normativa de cada país, dirigidos a garantizar la privacidad de los usuarios. En Ecuador, un país sin legislación para protección de datos personales, no está claro cómo los controladores de estos datos actúan para salvaguardarlos.

Como primer paso para medir la protección de datos personales en Ecuador, planteamos analizar sistemáticamente las políticas de privacidad publicadas en Internet de los sitios más populares en este país, incluyendo aquellos que recolectarían gran cantidad de datos. Usando analítica de datos y una exploración manual, se estudiarán los principales componentes de dichas políticas, incluyendo: tipo de información personal tratada, motivaciones, destinatarios, mecanismos de protección, consentimiento, notificación, y uso de herramientas especializadas de rastreo. La interpretación de estos datos brindará una perspectiva del estado actual de la protección de datos personales, y una línea base para estudiar su evolución, por ejemplo, debido a nueva legislación de protección, o para su comparación con el estado en otros países.

6. **PALABRAS CLAVE** (4-6)
privacidad; políticas; datos personales; Ecuador

7. OBJETIVOS

7.1. OBJETIVO GENERAL

Evaluar el estado de la protección de datos personales en Ecuador mediante el estudio sistemático de las políticas de privacidad de sitios web ecuatorianos y la comparación de su estructura con las de otro país latinoamericano.

7.2. OBJETIVOS ESPECÍFICOS

- Estudiar las prácticas y patrones recomendados localmente y por organismos y regulación internacionales para la elaboración de políticas de privacidad en la información.
- Construir un *data set* de políticas de privacidad representativo de (al menos 50) sitios web ecuatorianos y de otro país latinoamericano, así como un registro de su localización en la Web que permita un monitoreo continuo de su evolución.
- Procesar el texto obtenido, utilizando técnicas automatizadas y manuales, para caracterizar el contenido y la estructura general de las políticas de privacidad en estos países.
- Evaluar el enfoque de protección de datos personales en Ecuador a partir de la caracterización de las políticas de privacidad de los sitios web locales.



8. HIPÓTESIS (opcional)

9. DETALLE DE LOS RESULTADOS ESPERADOS *(con relación a los objetivos)*

- Estado del arte del análisis de protección de datos personales, enfocado en el procesamiento de políticas de privacidad.
- Data sets* y registro de políticas de privacidad en Ecuador y otro país latinoamericano.
- Metodología de recolección y procesamiento de información no estructurada presente en políticas de privacidad.
- Análisis y diagnóstico preliminar del estado actual de la protección de datos personales en Ecuador.

10. IMPACTO DE LA INVESTIGACIÓN *(científico, social, económico u otros (máximo una carilla))*

El impacto de esta propuesta de investigación tiene varias aristas en el ámbito de la protección de privacidad de la información, particularmente en Ecuador. Primero, a nivel científico, la investigación en privacidad o protección de datos en el contexto ecuatoriano es casi inexistente. Así, este proyecto pretende no solamente motivar el desarrollo de investigación en esa línea, sino también introducir esta temática en la discusión de la opinión pública para crear conciencia en los individuos y las instituciones (impacto en lo social). Esta discusión es fundamental para un país en el que no se ha desarrollado una política clara frente a la protección de la privacidad que, además, es un derecho humano.

Motivada por esta limitación y por la falta de transparencia frente al tratamiento de datos personales en Ecuador, nuestra investigación busca revelar el estado de compromiso que tendrían distintos actores frente a su protección actualmente, incluyendo parámetros como justificación del procesamiento o tipo de información recolectada. Más allá del impacto científico de esta investigación, los resultados pueden tener un impacto en la esfera regulatoria de la protección de datos, pues podrán guiar ciertas disposiciones que se emitan en una futura normativa. De hecho, esta investigación puede servir de base en la discusión del proyecto de ley de protección de datos personales que actualmente se debate en la Asamblea Nacional.

Esta investigación tiene además un impacto a mediano plazo y posiblemente más amplio a nivel geográfico. La construcción de una línea base en términos de registro y almacenamiento de las políticas de privacidad usadas actualmente servirá para analizar, por ejemplo, la efectividad de una regulación futura de privacidad. Si los esfuerzos de protección resultan ahora deficientes, es natural pensar que la emisión de normativa específica al respecto mejoraría esa situación. Basados en la metodología e información recolectada en esta investigación será posible medir la evolución esperada de la protección de datos al comparar los resultados históricos.

Esta propuesta está relacionada estrechamente con la línea de investigación *Seguridad y Privacidad* del Departamento de Electrónica, Telecomunicaciones, y Redes de Información.

11. ESTADO DEL ARTE, E INVESTIGACIONES PREVIAS DEL EQUIPO *(máximo tres carillas)*

Las políticas de privacidad son un mecanismo importante de transparencia en un contexto tan sensible como la protección de datos personales. Por esa razón, su estudio se ha abordado desde distintos enfoques, aunque casi exclusivamente en Europa y Estados Unidos.

Por ejemplo, trabajos previos se han orientado a medir el nivel de adopción de políticas de privacidad a lo largo del tiempo. Como es de esperarse, conforme se ha masificado el acceso a Internet, ha crecido el nivel de adopción [4, 6, 8]. Dicha adopción se ha concentrado históricamente en sectores relacionados con la tecnología y no tanto con la industria. Además, ciertos eventos como la emisión de nuevas y más fuertes regulaciones han motivado un incremento en la adopción de políticas de privacidad [15].

Al estar dirigidas a usuarios finales, las políticas de privacidad se han estudiado también desde la perspectiva de su usabilidad [7], que resulta importante dada la complejidad que suele caracterizar a su contenido. En esa línea, se ha encontrado que se requeriría una gran cantidad de tiempo y un nivel de educación avanzado para poder alcanzar a leer y entender todas aquellas de los sitios que un usuario común visita [9]. En cualquier caso, muy pocos usuarios las leen, pero la mera existencia de una ya proporcionarían



cierta confianza de los usuarios en las entidades [1] que las publican, seguramente porque no las comprenden.

Debido a la complejidad inherente de este documento, hay esfuerzos para automatizar su análisis, a partir de la recolección del texto de políticas de privacidad para construir *corpuses* [12] que permitan, por ejemplo, entrenar algoritmos de aprendizaje. Los trabajos relacionados con el análisis de políticas de privacidad incluyen: estudiar el impacto de la longitud de estos documentos [5] o de su diseño, o construir herramientas para resumir o visualizar su contenido [11].

Desafortunadamente, una revisión de la literatura científica muestra una muy limitada atención a contextos en los que el idioma de las políticas de privacidad es distinto del inglés, y menos aún en ámbitos regionales diferentes a Estados Unidos y Europa.

En lo referente a las investigaciones previas del equipo, a continuación, se detallan los proyectos relacionados con la línea de investigación *Seguridad y Privacidad* en los que los proponentes de este proyecto han participado:

- “Desarrollo de Prototipos para la Investigación en Seguridad de la Información”, Código PIMI-14-21, Proyecto Multidisciplinario financiado por la Escuela Politécnica Nacional.
- CIPSEC (Enhancing Critical Infrastructure Protection with innovative SECURITY framework). European project - Código: H2020-EU.3.7. - Secure societies - Protecting freedom and security of Europe and its citizens.
- "Privacidad Sintáctica Funcional: Análisis y adaptación de mecanismos de anonimato con enfoque en la preservación de utilidad de los datos", Código: PII-DETRI-2019-01, Proyecto Interno auspiciado por la Escuela Politécnica Nacional.

Publicaciones previas relacionadas con el proyecto

J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, and J. Forné, "Online advertising: Analysis of privacy threats and protection approaches," Elsevier Computer Communications, vol. 100, pp. 32-51, March 2017. ISSN: 0140-3664. Impact factor 2017: 2.613.

J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, and J. Forné, "On the regulation of personal data distribution in online advertising platforms," Engineering Applications of Artificial Intelligence, vol. 82, pp. 13-29, June 2019. ISSN: 0140-3664. Impact factor 2019: 4.201.

J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, and J. Forné, "Measuring Online Tracking and Advertising in Iberoamerica," to be submitted to IEEE Access. ISSN: 2169-3536. Impact factor 2019: 3.745.

J. Estrada-Jiménez, J. Parra-Arnau, A. Rodríguez-Hoyos, and J. Forné, "Measuring Online Tracking and Privacy Risks on Ecuadorian Websites," IEEE Fourth Ecuador Technical Chapters Meeting (ETCM), Guayaquil, Ecuador, November 2019.

Rodríguez-Hoyos, D. Rebollo-Monedero, J. Estrada-Jiménez, J. Forné, and L. Urquiza-Aguiar, "Preserving Empirical Data Utility in k-Anonymous Microaggregation via Linear Discriminant Analysis," accepted to be published in Elsevier Engineering Applications of Artificial Intelligence, May 2020. ISSN: 0952-1976. Impact factor 2019: 4.201.

Rodríguez-Hoyos, J. Estrada-Jiménez, D. Rebollo-Monedero, J. Parra-Arnau, Ahmad Mohamad Mezher, and J. Forné, "The fast MDAV (F-MDAV) algorithm: An algorithm for k-anonymous microaggregation in big data," Elsevier Engineering Applications of Artificial Intelligence, vol. 90, no. 103531, April 2020. ISSN: Impact factor 2019: 4.201.

Rodríguez-Hoyos, J. Estrada-Jiménez, D. Rebollo-Monedero, J. Parra-Arnau, and J. Forné, "Does k-anonymous microaggregation affect machine learned macrotrends?," IEEE Access, vol. 6, pp. 28 258-28 277, May 2018. ISSN: 2169-3536. Impact factor 2018: 4.098.



E. Pallarès, D. Rebollo-Monedero, A. Rodríguez-Hoyos, J. Estrada-Jiménez, A. Mohamad Mezher, and J. Forné, "Mathematically optimized, recursive prepartitioning strategies for k-anonymous microaggregation of large-scale datasets," Elsevier Expert Systems with Applications, vol. 144, pp. 1-17, April 2020. ISSN: 0957-4174. Impact factor 2019: 5.452.

Rodríguez-Hoyos, J. Estrada-Jiménez, D. Rebollo-Monedero, J. Forné, R. Trapero, A. Álvarez, and R. Rodríguez, "Anonymizing cybersecurity data in critical infrastructures: The CIPSEC approach," in Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM), Valencia, Spain, May 2019.

Estrada-Jiménez, J. A., Estrada-Jiménez, J. C., Rodríguez-Hoyos, A., & Tipantuña, C. (2015). Ecuador y la Privacidad en Internet: Una Aproximación Inicial. Revista Politécnica, 36(1), 54-54.

Rodríguez-Hoyos, A., Estrada-Jiménez, J., Urquiza-Aguilar, L., Parra-Arnau, J., & Forné, J. (2018, April). "Digital hyper-transparency: leading e-government against privacy". In 2018 International Conference on eDemocracy & eGovernment (ICEDEG) (pp. 263-268). IEEE.

Estrada-Jiménez, J., Calva, M., Rodríguez-Hoyos, A., Tipantuña, C., "Seguridad de la telefonía IP en Ecuador: Análisis en internet." *Enfoque UTE 7.2* (2016): 25-40.

12. DESCRIPCIÓN DETALLADA DEL PROYECTO, INCLUIDO METODOLOGÍA (*máximo tres carillas*)

Este proyecto pretende aportar en la evaluación del estado de la *protección* de datos personales en Ecuador, más allá de las contribuciones que este mismo equipo ha hecho para revelar los riesgos vigentes para la privacidad de los usuarios [2,3,10] en este país. Para este análisis, recurrimos al análisis sistemático de las políticas de privacidad de los sitios web ecuatorianos más populares o que más información personal podrían estar recopilando. Ya que este documento tiene el rango de un contrato y dada la amplitud de su adopción, es posible extraer de éste información valiosa que revele la postura de las entidades frente a la protección de los datos de los usuarios. Proponemos extraer esta información mediante distintas técnicas de procesamiento de texto y comparar los resultados obtenidos con aquellos de otro país latinoamericano que posea legislación de protección de datos.

Para empezar, se estudiará la regulación referente a privacidad y protección de datos tanto a nivel local como internacional (en particular, la más moderna legislación europea). Esto permitirá revelar las principales líneas guía para la elaboración de políticas de privacidad presentes en el Reglamento General de Protección de Datos europeo. Aunque la legislación ecuatoriana relacionada es muy limitada y dispersa, se investigará también dichas bases en las regulaciones de telecomunicaciones y transmisión de datos. Asimismo, se revisará en la literatura científica aportes aplicados y sistemáticos relacionados con el estudio de políticas de privacidad o protección de datos.

Con el fin de analizar el contexto de protección de datos en Ecuador, se construirá un *data set* de políticas de privacidad a partir de su recolección en un grupo representativo de al menos 50 sitios web locales. Para esta recolección, se visitarán los sitios web más populares según el *ranking* de Alexa [13], incluyendo también algunos sitios que por su naturaleza podrían estar recolectando gran cantidad de datos personales. Se identificará la política de privacidad en cada sitio ya sea manualmente o de forma automática, y se registrará su ubicación en la Web para referencia futura. Entre los parámetros que se registrará están su URL y una marca de tiempo. Finalmente, cada política de privacidad será almacenada en línea mediante Internet Archive [14] y localmente en formato de texto claro.

Una vez disponible el *data set* de políticas de privacidad, se procesará el texto de todas, utilizando técnicas automatizadas y manuales que permitan analizar la estructura y el contenido de dichos documentos. El proceso automatizado de procesamiento de texto se hará en la medida de lo que el contenido de las políticas de privacidad lo permita; éste podría incluir técnicas tales como conteo de palabras, identificación de tópicos clave (como referencia a rastreo, publicidad o *cookies*), y análisis semántico. El procesamiento de texto manual se realizará para identificar los principales componentes estructurales de cada política de privacidad, por ejemplo: tipos de datos personales tratados, justificación del tratamiento, técnicas utilizadas, o entidades involucradas en dicho tratamiento. Cabe



anotar que dicho procesamiento se aplicará también al *data set* obtenido del otro país latinoamericano seleccionado.

Finalmente, en función de la información procesada y de la caracterización resultante de las políticas de privacidad, se analizará el estado de protección de datos personales en Ecuador para ofrecer un diagnóstico preliminar. En primera instancia, se medirá la prevalencia del uso de políticas de privacidad en el país en el dominio de Internet. En un enfoque más profundo, se analizará la concordancia de las políticas de privacidad locales con la regulación ecuatoriana e internacional relacionada, es decir la legislación ecuatoriana referente a telecomunicaciones y transmisión de datos, y al Reglamento General de Protección de datos europeo. Todo esto permitirá luego discutir y diagnosticar el nivel de protección de datos que las entidades ecuatorianas estarían aplicando en un entorno tan masivo como Internet. Además, se brindará a esta discusión más contexto al incluir en el análisis la situación de otro país latinoamericano que sí disponga de legislación referente a la protección de datos personales.

Sin duda, ésta es una de las varias aproximaciones que podrían hacerse para estudiar el problema de la protección de datos personales en un escenario tan particular como el ecuatoriano. Sin embargo, frente a la opacidad de las instituciones en el tratamiento de datos, consideramos que éste es un primer paso interesante para su estudio.

Bibliografía (Normas APA)

- [1] Ermakova, T., Baumann, A., Fabian, B., & Krasnova, H. (2014). Privacy policies and users' trust: does readability matter?
- [2] Estrada-Jiménez, J. A., Estrada-Jiménez, J. C., Rodríguez-Hoyos, A., & Tipantuña, C. (2015). Ecuador y la Privacidad en Internet: Una Aproximación Inicial. *Revista Politécnica*, 36(1), 54-54.
- [3] Estrada-Jiménez, J., Rodríguez-Hoyos, A., Parra-Arnau, J., & Forné, J. (2019, November). Measuring Online Tracking and Privacy Risks on Ecuadorian Websites. In 2019 IEEE Fourth Ecuador Technical Chapters Meeting (ETCM) (pp. 1-6). IEEE.
- [4] Federal Trade Commission, "FTC Releases Report on Consumers' Online Privacy," <https://www.ftc.gov/news-events/press-releases/1998/06/ftc-releases-report-consumers-online-privacy>, Jun. 1998.
- [5] Gluck, J., Schaub, F., Friedman, A., Habib, H., Sadeh, N., Cranor, L. F., & Agarwal, Y. (2016). How short is too short? Implications of length and framing on the effectiveness of privacy notices. In Twelfth Symposium on Usable Privacy and Security (SOUPS) 2016 (pp. 321-340).
- [6] Liu, C., & Arnett, K. P. (2002). Raising a red flag on global WWW privacy policies. *Journal of Computer Information Systems*, 43(1), 117-127.
- [7] McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *Isjlp*, 4, 543.
- [8] Nokhbeh Zaeem, R., & Barber, K. S. (2017). A study of web privacy policies across industries. *Journal of Information Privacy and Security*, 13(4), 169-185.
- [9] Proctor, R. W., Ali, M. A., & Vu, K. P. L. (2008). Examining usability of web privacy policies. *Intl. Journal of Human-Computer Interaction*, 24(3), 307-328.
- [10] Rodríguez-Hoyos, A., Estrada-Jiménez, J., Urquiza-Aguilar, L., Parra-Arnau, J., & Forné, J. (2018, April). Digital hyper-transparency: leading e-government against privacy. In 2018 International Conference on eDemocracy & eGovernment (ICEDEG) (pp. 263-268). IEEE.
- [11] Turow, J., Hennessy, M., & Draper, N. (2018). Persistent misperceptions: Americans' misplaced confidence in privacy policies, 2003–2015. *Journal of Broadcasting & Electronic Media*, 62(3), 461-478.



[12] Wilson, S., Schaub, F., Dara, A. A., Liu, F., Cherivirala, S., Leon, P. G., ... & Norton, T. B. (2016, August). The creation and analysis of a website privacy policy corpus. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1330-1340).

[13] Amazon, "Top Sites in Ecuador", <https://www.alexa.com/topsites/countries/EC>, Aug. 2020.

[14] Kahle, B., "Internet Archive ` Online Privacy,"<https://archive.org>, Aug. 2020.

[15] Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., & Holz, T. (2018). We value your privacy... now take some cookies: Measuring the GDPR's impact on web privacy. *arXiv preprint arXiv:1808.05096*.

13. INFRAESTRUCTURA Y EQUIPOS

- Indicar la infraestructura y equipos **disponibles** para la ejecución del proyecto, con la ubicación actual de los mismos

Infraestructura	Equipos	
Laboratorio ZZ	Nombre del Equipo	Ubicación del Equipo
Intel Core i7-7700 CPU 3.6 GHz, 16 GB RAM	Computador personal, director	
Intel Core i7-7700 CPU 3.6 GHz, 16 GB RAM	Computador personal, colaborador	



B. DATOS INFORMATIVOS

1. INFORMACIÓN DEL DIRECTOR, COLABORADOR (EPN o EXTERNO) Y COLABORADORES TÉCNICOS

Apellidos y nombres	No. de Cédula	HSS*	Departamento	Rol	Título de mayor nivel y mención.
Rodríguez Hoyos Ana Fernanda	1713027157	8	Departamento de Electrónica, Telecomunicaciones y Redes de Información - DETRI	Director	Doctora en Ingeniería Telemática
Estrada Jiménez José Antonio	1714672456	6	Departamento de Electrónica, Telecomunicaciones y Redes de Información - DETRI	Colaborador	Doctor en Ingeniería Telemática

* HSS =Horas Semana Semestre: Es el número de horas que se dedica por semana a la investigación. Este número de horas se mantiene para todo el semestre



C. DECLARACIÓN FINAL DECLARACIÓN DEL DIRECTOR DEL PROYECTO

El equipo de investigadores, representado por el Director del Proyecto declara lo siguiente:

- Que el presente proyecto es una creación original de mi autoría y del equipo de investigadores, y por tanto asumimos la completa responsabilidad legal en caso de que un tercero alegue la titularidad de los derechos intelectuales del proyecto, exonerando a la EPN de cualquier acción legal que se derive por esta causa.
- Que el presente proyecto no ha sido presentado en ninguna convocatoria de otra institución pública o privada. El incumplimiento será causal para que el proyecto no sea tomado en consideración.
- Que si el proyecto genera algún producto o procedimiento susceptible de obtener derechos de propiedad intelectual, de los cuales se deriven beneficios, aceptamos que éstos serán compartidos entre los investigadores y la institución o las instituciones participantes en el proyecto, conforme a lo establecido en el COESC.
- Que el equipo de investigadores y/o instituciones participantes se comprometen a mantener la confidencialidad de la información si ésta podría ser susceptible de protección por patentes, y solicitar la valoración de propiedad intelectual respectiva previa a cualquier publicación o difusión.
- Que para el caso de derechos de autor otorgamos una licencia de uso exclusivo con fines académicos para la o las instituciones participantes en el proyecto.
- Que aceptamos conocer y cumplir con la normativa vigente para la gestión de proyectos.

ANA FERNANDA 2021.02.19
RODRIGUEZ 02:01:35
HOYOS -05'00'

Firma de la Directora del Proyecto
Nombre: Ana Fernanda Rodríguez Hoyos
C.I.: 1713027157

