

Germán F. Rosati

Correo electrónico:
german.rosati@gmail.com

<https://ar.linkedin.com/in/germán-rosati-53637263>

Sitio personal: <https://gefero.github.io/>

LinkedIn:

Resumen

Es Investigador Adjunto del CONICET en la Escuela Interdisciplinaria de Altos Estudios Sociales (EIDAES-UNSAM), donde también se desempeña como docente. Allí coordina el laboratorio de Ciencias Sociales Computacionales, factor~data. Además, es investigador del PIMSA.

Anteriormente coordinó los programas de Data Science e Inteligencia Artificial de Digital House y el Área de Procesos de Producción de Datos del programa PASCAL-UNSAM. También se ha desempeñado como Analista Experto de Datos en el Ministerio de Trabajo, Empleo y Seguridad Social de la Nación. Ha dictado cursos de grado y posgrado sobre análisis cuantitativo y aprendizaje automático. Ha sido becario doctoral del CONICET e investigador Visitante de la Freie Universität Berlin y del Instituto Iberoamericano de Berlín.

Es doctor y licenciado en Sociología por la Universidad de Buenos Aires y magíster en Generación y Análisis de Información Estadística por la Universidad Nacional de Tres de Febrero.

Sus áreas de interés son: Ciencias Sociales Computacionales, métodos de Machine Learning aplicados a las ciencias sociales, procesamiento de lenguaje natural y el análisis de la estructura social y agraria.

1. Posiciones actuales

- *Investigador Adjunto*, Consejo Nacional de Investigaciones Científicas y Técnicas.
- *Profesor Titular*, "Métodos Cuantitativos de Investigación", Escuela Interdisciplinaria de Altos Estudios Sociales, Universidad Nacional de San Martín.
- *Coordinador* Factor~Data, laboratorio de Ciencias Sociales Computacionales, Escuela Interdisciplinaria de Altos Estudios Sociales, Universidad Nacional de San Martín.
- *Coordinador*, Diploma en Ciencias Sociales Computacionales y Humanidades Digitales, Escuela Interdisciplinaria de Altos Estudios Sociales, Universidad Nacional de San Martín.
- *Profesor a Cargo* "Visualización de la Información", *Maestría en Explotación de Datos y Descubrimiento de Conocimiento*, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

2. Educación

2015, Doctor en Ciencias Sociales, Universidad de Buenos Aires

Tesis: "Funcionalidad y diferenciación en un mercado de trabajo en metamorfosis: el caso de la cosecha de algodón en Chaco (1960-2010)", Calificación: Sobresaliente- Summa cum Laude (máxima)

2013, Magister en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero, Argentina

Tesis: "Estimación de la demanda de trabajo requerida en la cosecha de algodón y análisis de sus determinantes. Chaco (1960-2009)", Calificación: Aprobada con mención honorífica (máxima)

2007, Licenciado en Sociología, Universidad de Buenos Aires, Argentina.

Maestría en Explotación de Datos y Gestión del Conocimiento, Universidad Austral, Argentina (*sin finalizar*).

3. Cursos y formación

- *Introducción al Data Mining*, Facultad de Ingeniería, Universidad Austral, Ciudad de Buenos Aires; 1er. semestre 2015; 32 hs. cátedra.
- *Bases de datos*, Facultad de Ingeniería, Universidad Austral, Ciudad de Buenos Aires; 1er.

- semestre 2015; 32 hs. cátedra; aprobado.
- *Algoritmos y estructuras de datos*, Facultad de Ingeniería, Universidad Austral, Ciudad de Buenos Aires; 1er. semestre 2015; 32 hs. cátedra, aprobado.
- *Métodos estadísticos no paramétricos*, Universidad Nacional de Tres de Febrero, Ciudad de Buenos Aires; noviembre-diciembre de 2012; 24 hs.; aprobado.
- *Introducción al análisis de datos cualitativos con Atlas-Ti*, Facultad de Ciencias Sociales de la Universidad de Buenos Aires; noviembre de 2012; 15 hs. cátedra; aprobado.
- *Seminario de Sistemas de Información Geográfica*, Facultad de Filosofía y Letras de la Universidad de Buenos Aires; Octubre-Diciembre 2011; 35 hs. cátedra.
- *Modelos de regresión no lineales*, Instituto de Industria de la Universidad Nacional de General Sarmiento; Los Polvorines, Provincia de Buenos Aires; Octubre 2010; 24 hs. cátedra.
- *Análisis de series de tiempo macroeconómicas con STATA*, Instituto de Industria de la Universidad Nacional de General Sarmiento; Los Polvorines, Provincia de Buenos Aires; Agosto 2010; 24 hs. cátedra; aprobado con 8 (ocho).

4. Otra experiencia docente

Posgrado

2023, Profesor “Ciencias Sociales Computacionales, Big Data y Machine Learning. Un mapa de situación del giro computacional en las ciencias sociales”, Instituto de Desarrollo Económico/ Universidad Nacional de Tres de Febrero (IDES/UNTREF).

2021, Profesor “Introducción a la minería de textos y procesamiento de lenguaje natural para ciencias sociales”, *Doctorado en Sociología*, Escuela IDAES, Universidad Nacional de San Martín.

2021, Profesor “Visualización de Datos”, *Doctorado en Sociología*, Escuela IDAES, Universidad Nacional de San Martín.

2020, Profesor “Ciencia de Datos y Machine Learning. Aplicaciones en R”, *Maestría en Generación y Análisis de Información Estadística*, Universidad Nacional de Tres de Febrero.

2019-2020, Profesor “Machine Learning en Aplicaciones Espaciales”, curso de posgrado *Big Data e Inteligencia Territorial*, Facultad Latinoamericana de Ciencias Sociales (FLACSO).

2019, Profesor “Análisis de la Estructura Social Argentina”, *Doctorado en Ciencias Sociales*, Facultad de Ciencias Sociales, UBA.

2019, Profesor “Introducción a la Ciencia de Datos para Científicos Sociales”, posgrado en *Métodos y técnicas para la investigación aplicada a la realidad socioeconómica regional*, Universidad Nacional de Santiago del Estero.

2016-2019, Profesor *Fundamentos de la Programación Estadística y Data Mining en lenguaje R*, curso de posgrado/extracurricular, Maestría en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero.

Grado

2024, Profesor invitado, *Introducción a la Visualización y Simulación Interactivas e Inmersivas*, Departamento de Computación, Facultad de Ciencias Exactas y Naturales,

2012-2017, Profesor Titular Ordinario, “Estadística y Técnicas de Computación Aplicadas a la Investigación en Ciencias Sociales” en Facultad de Ciencias Sociales, Universidad del Salvador, Argentina

2016, Profesor invitado, *Investigación Electoral y de Opinión Pública Aplicada II*, Maestría en

Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero.

2016, Profesor, *Introducción a la Estadística*, curso de posgrado/extracurricular, Escuela de Política y Gobierno, Universidad Nacional de General San Martín.

2014, Jefe de Trabajos Prácticos, *Estadística Aplicada a las Ciencias Sociales*, Instituto de Conurbano, Universidad Nacional de General Sarmiento.

2014, Instructor, *Estadística Aplicada a las Ciencias Sociales*, Instituto de Conurbano, Universidad Nacional de General Sarmiento (modalidad virtual).

2014, Profesor (Jefe de Trabajos Prácticos), *Problemas Socioeconómicos Contemporáneos*, Instituto de Ciencias, Universidad Nacional de General Sarmiento.

2013, Profesor, *Introducción al Análisis Estadístico Multivariado*, curso extracurricular, Facultad de Filosofía y Letras, Universidad Nacional de Cuyo.

5. Intereses y campos de investigación

Estructura Social y Agraria; Data Mining; Machine Learning; Modelización de datos; Procesamiento de Lenguaje Natural; Programación Estadística; Datos Cualitativos

6. Habilidades y competencias

Modelización estadística y aplicación de algoritmos de machine learning (regresión, árboles de decisión, clustering, etc.); ensamble de modelos (algoritmos bagging y boosting, random forest, etc.), Natural Language Processing, Text Mining

Limpieza y análisis de consistencia en bases de datos.

Programación estadística: R, Python (Scikit-Learn, Pandas, numpy, scipy)

Paquetes estadísticos: SAS, SAS-Miner, SPSS, Stata, Knime

GIS: Quantum GIS, ArcGIS, ArcView

Bases de datos: programación en SQL

7. Publicaciones científicas

Libros

- *Los excedentes del campo. Superpoblación relativa y cuestión agraria en Argentina*, Buenos Aires: UNSAM Edita,, 2025 (en coautoría con Adriana Chazarreta)
- *Los que sobran (para el capital). Los trabajadores argentinos entre el activo y la reserva*, Buenos Aires: Grupo Editor Universitario, 2023 (en coautoría con Ricardo Donaire).

Artículos en revistas científicas con referato

Publicados

- "Configuraciones espaciales del agro argentino. Análisis de las estructuras sociales en Argentina mediante datos de alta resolución (2018-2020)" (en coautoría con Adriana Chazarreta), en *Estudios Rurales*, Vol. 15, N°32, <https://doi.org/10.48160/22504001er32.635>.
- "Movilidad laboral de los asalariados agropecuarios en Argentina", en *Estudios Del Trabajo*,

- Nº69, (en coautoría con Adriana Chazarreta), <https://ojs.aset.org.ar/revista/article/view/161>
- “Desarrollo capitalista, globalización y cuestión agraria en el siglo XXI”, *Revista Eutopía*, FLACSO-ANDES, Nº27, 2025, DOI: <https://doi.org/10.17141/eutopia.27.2025.6531> (en coautoría con Adriana Chazarreta).
 - “Gender biases and hate speech: Promoters and targets in the Argentinean political context”. *PLOS ONE* 20(1): e0317001, 2025. <https://doi.org/10.1371/journal.pone.0317001> (en coautoría con Domenech Burin L, Pérez JM, Rodrigues Pires M, Nanton M, Kozłowski, D.)
 - “Analizando trayectorias de uso del suelo. Una propuesta de clusterización.” en *Geograficando*, FaHCE-UNLP, 2023, (ISSN:2346-898X).
 - “Una aproximación a los temas acerca de la COVID-19. Aplicación de técnicas de procesamiento de lenguaje natural sobre comentarios de lectores de noticias digitales” en *Papeles de Trabajo*, EIDAES-UNSAM, nº15, Vol. 28,, (ISSN: 1851-2577), 2022 (en coautoría con Adriana Chazarreta, Laia Domenech y Tomás Maguire).
 - “La integración de fuentes diversas de información en los estudios territoriales. Reflexiones sobre dos investigaciones” en *Revista QUID* 16, nº17, 2022 (ISSN: 2250-4060).
 - “Procesamiento de Lenguaje Natural aplicado a las ciencias sociales. Detección de tópicos en letras de tango” en *Revista Latinoamericana de Metodología de la Investigación Social*, nº23, CIES, 2022, (ISSN: 1853-6190)
 - “Métodos de Machine Learning como alternativa para la imputación de datos perdidos. Un ejercicio en base a la Encuesta Permanente de Hogares” en *Estudios del Trabajo*, nº61, 2021, (ISSN: 2545-7756)
 - “Tipos de estructuras sociales agrarias en la formación social argentina. Un análisis a nivel departamental: 2001-2002” en *Revista Mundo Agrario*, nº21, Vol. 48, 2020, CEHR, Facultad de Humanidades y Ciencias de la Educación, UNLP, (ISSN: 1515-5994) (en coautoría con Adriana Chazarreta).
 - “Building a sanitary vulnerability map from open source data in Argentina (2010-2018)”, en *International Journal of Equity in Health*, nº20, 2020, (ISSN: 1475-9276).
 - “Ciclos de ocupación anuales y patrones de movilidad de los trabajadores agropecuarios: una aproximación a sus determinantes mediante técnicas de Ensemble Learning”, en *Estudios Rurales*, nº19, Vol. 10, CEAR, UNQUI, (ISSN: 2250-4001).
 - “Una aproximación a las trayectorias laborales de los trabajadores mecánicos y manuales en la cosecha algodonera. Chaco, Argentina”, en *Trabajo y Sociedad*, nº30, 2018, Facultad de Humanidades y Ciencias Sociales, Universidad Nacional de Santiago del Estero, (ISSN: 1514-6871).
 - “Construcción de un modelo de imputación para variables de ingreso con casos perdidos a partir de ensemble learning. Aplicación a la Encuesta Permanente de Hogares”, en *Revista SaberEs*, Vol. 9, nº1, 2017, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario, (ISSN: 1852-4184).
 - “El Qualitative Comparative Analysis (QCA) como herramienta analítica. Dos aplicaciones para el análisis de entrevistas” en *Revista Latinoamericana de Metodología de las Ciencias Sociales*, Centro Interdisciplinario de Metodología de las Ciencias Sociales, Vol. 7, nº, 2017 Universidad Nacional de La Plata, La Plata, (ISSN: 1853-7863), (en coautoría con Adriana Chazarreta).
 - “Transformaciones vitivinícolas recientes: estructura productiva y patrones de movilidad de las bodegas. Mendoza (Argentina), 2004-2011”, en *Cuadernos FhyCS-UNJu*, nº50, 2016, Facultad de Humanidades y Ciencias Sociales, Universidad Nacional de Jujuy, San Salvador de Jujuy, (ISSN: 0327-1471) (en coautoría con Adriana Chazarreta).
 - “Dos dimensiones de las lógicas de funcionamiento de un mercado de fuerza de trabajo agropecuario: niveles y formas de remuneración en la cosecha de algodón, (Chaco 1990-2012)”, en *Revista Población y Sociedad*, Vol. 22, nº2, 2015, Instituto Superior de Estudios Sociales, Universidad Nacional de Tucumán, Tucumán (ISSN: 1852-8562).
 - “Cambios en la estructura social agraria pampeana. Un acercamiento a la caracterización de los pequeños propietarios rentistas del sur de Santa Fe”, en *Pampa. Revista Interuniversitaria de Estudios territoriales*, nº9, 2013, Universidad Nacional del Litoral, Santa Fe, (ISSN: 1669-3299) (en coautoría con Masello, D.)
 - “Conflicto agrario y desarrollo del capitalismo en el campo. Un análisis exploratorio del

- conflicto por la resolución 125 (Argentina, marzo-julio 2008)", en *Revista América Latina* n°13/14, 1er semestre de 2013, Doctorado en Procesos Sociales y Políticos en América Latina (PROPSAL), Universidad ARCIS, Santiago de Chile, (ISSN: 0717-7097)
- "La fijación de trabajadores indígenas en un territorio productivo. Documentos para el análisis de la 'Masacre de Napalpí', Chaco, 1924", en *Estudios del Trabajo*, n°41/42 Asociación de Estudios del Trabajo, Buenos Aires, 2011 (ISSN: 0327-5744)
 - "Sobre el supuesto de 'homogeneidad' en el análisis de la estructura social. Reflexiones a partir de un ejercicio empírico", en *Revista de la Carrera de Sociología. Entramados y perspectivas*, Facultad de Ciencias Sociales, Universidad de Buenos Aires, Buenos Aires, (ISSN: 1853-6484) (coautoría con R. Donaire)
 - "La captación estadística de los asalariados agropecuarios. Reflexiones en torno a sus problemas mediante un ejercicio de comparación entre censos de población y agropecuarios (Argentina, 2001/2002)", en *Revista Mundo Agrario*, N°23, 2do. semestre 2011, Centro de Estudios Histórico Rurales (CEHR), Facultad de Humanidades y Ciencias de la Educación, UNLP, La Plata, (ISSN: 1515-5994).
 - "Consideraciones sobre algunas metodologías habituales para el análisis de tablas de movilidad social. Ventajas de una alternativa basada en cadenas de Markov", en *Empiria. Revista de Metodología de Ciencias Sociales*, N°22, julio-diciembre 2011, Facultad de Ciencias Políticas y Sociología de la UNED, Madrid, (ISSN: 1139-5737).

Artículos en libros

- "Frontera, deforestación y despoblamiento. Discutiendo algunas hipótesis sobre la expansión reciente de la frontera agraria en Argentina" en Porcaro, Tania; Salizzi, Esteban; Martirén, Juan Luis; Lanteri, Sol (comps.). *Fronteras: aportes para la consolidación de un campo de estudios*. Buenos Aires: TeseoPress.
- "Trabajador algodonero", en Muzlera, J. y Salomón, A. (comps.). *Diccionario del Agro Iberoamericano*, Editorial Teseo.
- "Estructuras sociales agrarias y conflicto en el sector. El caso del llamado 'conflicto por la 125' en 2008", en Panero, Marcelo (comp.), *Actores, políticas públicas y conflicto agropecuario. A 10 años de la Resolución 125*, 2019, Editorial Universitaria de Villa María.
- "Problemas teóricos en torno a las funciones de la población sobrante en el capitalismo dependiente: un ejercicio de comparación entre las formaciones sociales argentina y brasileña, 1980-2010", en Cangussu de Souza, D. (comp.) (2016): *Desemprego e movimentos sociais*, Editora FAP-UNIFEPS, Brasil.
- "Los cambios en la estructura social agraria argentina, 1988-2012", en Kessler, Gabriel (comp.) (2015): *La sociedad argentina hoy. Radiografía de una nueva estructura*, Siglo XXI Editores- Fundación OSDE, Buenos Aires (en coautoría con Adriana Chazarreta).
- "Patrones espaciales de expansión de la frontera agrícola: la expansión sojera en Argentina (1987/88-2009/10)" en Gras, C. y Hernández, V. (comps.) (2013): *El agro como negocio: producción, sociedad y territorios en la globalización*, Biblos, Buenos Aires.
- "Crisis del algodón, cambio tecnológico y expansión sojera en el Chaco (1988-2009). Impactos sobre la demanda de fuerza de trabajo en la cosecha algodonera" en Gras, C. y Hernández, V. (comps.) (2013): *El agro como negocio: producción, sociedad y territorios en la globalización*, Biblos, Buenos Aires.

Documentos de trabajo (no publicados, internos y pre-prints)

- "The undesirable dependence on frequency of gender bias metrics based on word embeddings" disponible en <https://arxiv.org/abs/2301.00792> (en coautoría con Francisco Valentini, Diego Fernandez Slezak, Edgar Altszyler)
- "On the interpretation and significance of bias metrics in texts: a PMI-based approach", disponible en <https://arxiv.org/abs/2104.06474>, 2021 (en coautoría con Francisco Valentini, Damián Blasi, Diego Fernandez Slezak, Edgar Altszyler)
- "Detecting Areas of Potential High Prevalence of Chagas in Argentina", disponible en <https://arxiv.org/abs/2001.00604>, 2019 (en coautoría con AA.VV.)
- "Construcción de un Mapa de Vulnerabilidad Sanitaria en Argentina a partir de datos públicos", disponible en <https://arxiv.org/abs/1901.08105>, 2019 (en coautoría con Antonio Vázquez Brust y Tomás Olego).

- “Factores asociados a la tercerización en empresas del AMBA. Aplicación de un clasificador basado en random forest”, *Documento interno MTEySS – BIRF*, 2016.
- “Propuesta de modelo de imputación de casos perdidos para variables de ingreso en encuestas a hogares. Ejercicio sobre EPH”, *Documento interno MTEySS – BIRF*, 2016.
- “Caracterización del ciclo ocupacional de los asalariados agropecuarios a partir de la utilización de clasificadores random forest”, *Documento interno MTEySS – BID*, 2015.

8. Otras posiciones relevantes

- 2017-2018, *Coordinador* Programa de Data Science en Digital House -cursos Data Analytics, Data Science y Artificial intelligence/Deep Learning
- 2015-2018, *Analista Experto de Datos (consultor externo)*, Subsecretaría de Políticas, Estadísticas y Estudios Laborales, Ministerio de Trabajo, Empleo y Seguridad Social, Argentina
- 2014-2015, *Coordinador Área de Procesos de Producción de Datos*, programa PASCAL, Universidad Nacional de San Martín
- 2014-2015, *Consultor Estadístico*, Banco Interamericano de Desarrollo (proyecto ejecutado por el MTEySS)
- 2014, *Investigador Principal*, Centro de Investigaciones en Estadística Aplicada Universidad Nacional de Tres de Febrero.
- 2009-2014, *Becario Doctoral*, Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET)
- 2006-2009, *Investigador Asistente*, Centro de Investigaciones en Estadística Aplicada, Universidad Nacional de Tres de Febrero.

9. Formación de Recursos Humanos

Dirección de Tesis de Grado

- Ignacio Rodríguez Sanudo, Licenciatura en Ciencia de Datos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (en curso).
- Santiago Rodríguez Rimedio, Licenciatura en Antropología Social, EIDAES, Universidad Nacional de San Martín.
- Matías Grinberg, Licenciatura en Psicología, Universidad Favaloro (aprobada con calificación máxima)
- Tomás Maguire, Licenciatura en Sociología, Universidad Nacional de San Martín (aprobada con calificación máxima)
- Florencia Piñeyrúa, Licenciatura en Sociología, Universidad Nacional de San Martín (aprobada con calificación máxima)

Dirección de Tesis de Maestría

- Paolo Donizetti, Maestría en Data Mining, Universidad de Bologna (aprobada con calificación máxima)
- Gabriel Kofman, Maestría en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero (aprobada con calificación máxima)
- Jorge Miceli, Maestría en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero (en curso)
- María Giselle Galli, Maestría en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero (en curso)

Co-dirección de Tesis de Maestría

- Juan José Iguarán Fernández, Maestría en Explotación de Datos y Descubrimiento de Conocimiento, Facultad de Ciencias Exactas y Naturales, UBA (finalizada, calificación Sobresaliente)
- Julia Gentile, Maestría en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero (en curso)

Co-dirección de Tesis de Doctorado

- María Leonor Pérez Bruno, Doctorado en Demografía, Universidad Nacional de Luján

10. Actividad de Investigación (básica y aplicada)

Proyectos de Investigación Básica

Actual

- *Director* “Frontera agropecuaria y procesos demográficos. Argentina 1988-2018”, Proyecto de Investigación Bianual para investigadoras/es Asistentes y Adjuntas/os de reciente ingreso (PIBAA).
- *Director* “Usos de Natural Language Processing para aplicaciones en Ciencias Sociales”, factor~data, Instituto de Altos Estudios Sociales, Universidad Nacional de San Martín.
 - *Análisis de conversaciones en foros de noticias sobre COVID-19 en cinco diarios nacionales* (Web scraping, Topic Modeling, fasttext, word embeddings)
 - *Los temas del tango y el rock en Argentina* (Web scraping, Topic Modeling).
- *Investigador* “La sustentabilidad de la producción agroalimentaria como imperativo global. Respuestas y controversias en torno a la cuestión ambiental” (financiado por Agencia Nacional de Promoción Científica y Tecnológica), Directora: Dra. Carla Gras.
Subproyecto: “Patrones espaciales de la frontera agrícola en Argentina: el caso del boom de la soja”

Finalizados

2020-2021, *Investigador* “Proyección de tendencias y evaluación de escenarios de intervención para la epidemia COVID-19 en Argentina mediante Modelado y Simulación Computacional”, (financiado por Agencia Nacional de Promoción Científica y Tecnológica), Director: Dr. Rodrigo Castro.

2020-2021, *Investigador* “Efectos económicos y productivos del aislamiento social sobre PyMES de industria y servicios en la Provincia de Buenos Aires (conurbano y otras localizaciones de alta densidad industria”, (financiado por Agencia Nacional de Promoción Científica y Tecnológica), Directora: Dra. Verónica Robert.

2017-2020, *Investigador* en “Modelos de desarrollo rural, dinámicas territoriales y cuestión ambiental en la Argentina contemporánea” (financiado por Agencia Nacional de Promoción Científica y Tecnológica), Directora: Dra. Carla Gras.

2018-2020, *Director*: “Aplicación de técnicas de Ensemble Learning para la imputación de datos perdidos en encuestas sociodemográficas”, financiado por la Secretaría de Investigación de la Universidad Nacional de Tres de Febrero.

2010-2013, *Investigador*: “Modelo de desarrollo: actores, disputas y escenarios en la Argentina Contemporánea” (PICT 2008 N°1216), Directora: Dra. Maristella Svampa.

2007-2009, *Investigador* del “Observatorio sobre la cultura política de la Ciudad de Buenos Aires” con asiento en Centro de Investigaciones en Estadística Aplicada (CINEA); Programación Científica y Tecnológica 2007- 2008 Universidad Nacional de Tres de Febrero; Director: Mg. Miguel Oliva.

2006-2007, *Asistente de investigación*: “Articulación y conflicto entre el movimiento obrero organizado sindical y políticamente y los desocupados en Argentina: 1930- 1935 y 1991-2004” Proyecto de Investigación Plurianual 2005-2006, CONICET (CONICET- PIP 5761).

Proyectos de Investigación Aplicada

FundAr

2024, *Director*, “Evaluación, análisis y propuesta de mejoras del sistema de visualizaciones de Argendata-FUNDAR”.

Objetivo: Análisis forense de visualizaciones en el sitio, propuesta de mejoras y diseño de pautas de estilo para las visualizaciones.
Modalidad: STAN-CONICET

Information Technology & Services S.A.

2023, Director, “Uso de Large Language Models para extracción de datos de recetas médicas”.

Objetivo:
Modalidad: STAN-CONICET

Centro Interdisciplinario de Estudios en Ciencia, Tecnología e Innovación (CIECTI-MINCyT).

2023, “Generación de un Índice de Potencial Bioeconómico (IPBE)”.

Objetivos: 1) generar un índice que resuma la información contenida en las 49 variables del Mapa de Potencial Bioeconómico; 2) identificar regiones en función del tipo y nivel de potencial bioeconómico con alta resolución; 3) generar y publicar un mapa interactivo con la información construida.

Modalidad: Asesoría individual CONICET

Municipio de Vicente López / Fundación Conocimiento Abierto / FUNINTEC

2020, Director, “Construcción de un clasificador automático de reportes sobre la situación de Ancianos Mayores en el Municipio de Vicente López”

Objetivos: 1) entrenar un modelo que permita automatizar la clasificación de reportes en texto libre acerca de la situación de los adultos mayores en el Municipio; 2) diseñar una metodología de evaluación continua y diagnóstico periódico del modelo; 3) desarrollar una herramienta (API) para uso de los analistas del municipio.

Organización Internacional del Trabajo / FUNINTEC

2020, Investigador, “Estudio de percepción sobre el trabajo infantil en el sector agropecuario”

Objetivos: estudio de percepción acerca del Trabajo Infantil y Adolescente en el sector agropecuario entre los diferentes públicos participantes de la exposición agroindustrial “Expoagro”

Ministerio de Trabajo, Empleo y Seguridad Social –MTEySS

2017, “Segmentando hogares rurales para la Encuesta Nacional de Actividades de Niños y Adolescentes”

Objetivo: análisis e implementación de varios métodos de clustering para segmentar hogares rurales georreferenciados y optimizar el trabajo de campo y el recorrido de los encuestadores.

2017, “Aplicación de Estimadores de Área Pequeña (*Small Area Estimators*) para el cálculo de indicadores sociolaborales a nivel departamental en base a datos de la Encuesta Nacional de Protección y Seguridad Social”

Objetivo: evaluar la aplicabilidad de estimadores de área pequeña para la construcción de información sociolaboral (tasas de actividad, empleo, desempleo y subempleo) para unidades censales departamentales/municipales.

2016, “Aplicación de algoritmos de *ensemble learning* para la imputación de datos perdidos en variables de ingreso en encuestas a hogares”

Objetivo: diseño e implementación de un método de imputación de datos perdidos en variables de ingreso que mejore la performance de otros métodos habituales (ej. “hot deck”) para ser aplicado en diversos relevamientos del Ministerio.

2016, “Uso de técnicas de remuestreo para la estimación del error muestral en un diseño complejo”

Objetivo: desarrollar un método de estimación del error muestral alternativo

2016, “Análisis de la tercerización en empresas del AMBA. Aplicación de un clasificador basado en Random Forest”

Objetivo: desarrollar un método para analizar la prevalencia de las prácticas de tercerización en empresas del AMBA.

Fundación Bunge y Born

2018, “Construcción de un monitor de discurso anti-vacuna en medios gráficos y en redes sociales”

Objetivo: puesta en producción de un sistema que detecte noticias y tuits con contenido anti-vacuna en tiempo real. Modelado de tópicos y clasificación automática de textos mediante word embeddings y topic modeling

2018, “Uso de Big Data y Datos Abiertos para la Detección de Zonas con Alto Riesgo de Chagas”

Objetivo: generación de un mapa de riesgo potencial de Chagas con alta resolución espacial

Como Consultor Independiente

“Diseño e implementación de un sistema de clusterización automático de encuestas”, realizado para QSocialNow, S.A.

“Diseño e implementación de un detector de anomalías para series de tiempo de menciones en redes sociales”, realizado para QSocialNow, S.A.

“Modelo predictivo sobre temas de agenda en base a datos de redes sociales (Twitter y Facebook)”, realizado para QSocialNow, S.A.

Diseño muestral de “Estudio de Opinión Pública en Goya, Corrientes”, realizado para Ibarómetro, S.A.

“Perfilado de votantes al Frente Alianza País, Ecuador, clasificador construido en base a Random Forest”, realizado para Demos, S.A.

11. Actividad de evaluación

- Jurado de tesis, Juan Manuel Ortiz de Zárate, Doctorado en Ciencias de la Computación, FCEyN-UBA, 2024
- Jurado de tesis, Juan Manuel Pérez, Doctorado en Ciencias de la Computación, FCEyN-UBA, 2022
- Miembro del Comité Académico del Doctorado en Sociología, EIDAES-UNSAM, 2025
- Miembro del Comité Científico de *Association for Computational Linguistics*, 2024
- Miembro del Comité Científico de *Use R! 2021 y 2022. The R Conference*.
- Miembro del Comité Científico de *AGRANDA 2020, 2021 y 2022 Simposio Argentino de Ciencia de Datos y Grandes Datos*.
- Miembro del Comité Científico de *Latin-R 2019, 2020, 2021, 2022, 2023 y 2024. Conferencia Latinoamericana sobre Uso de R en Investigación + Desarrollo*.
- Evaluador de artículos en revistas *Journal of Agrarian Change*, *Revista Latinoamericana de Estudios del Trabajo*, *@H-industria*, *Población y Sociedad*. *Revista Regional de Estudios Sociales*; *Kairos. Revista de Temas Sociales*; *Multiciencias (Venezuela)*; *Revista Corpoica (Colombia)*, *Revista Saberes (Argentina)*, *Folia Histórica del Nordeste (Argentina)*.
- 2016-2018, Coordinador del Comité Editorial de la *Revista Argentina de Estadística Aplicada*, Universidad Nacional de Tres de Febrero.

12. Charlas, clases públicas, exposiciones

- 2023, "Abordajes ágiles para el análisis de trayectorias de uso del suelo", exposición en el *Ateneo sobre "Economía Agraria"* del Centro Interinstitucional en Ciencia de Datos, Facultad de Ciencias Exactas, UBA - Ministerio de Ciencia y Técnica (MINCyT).
- 2023, "Del giro lingüístico al giro (lingüístico) computacional. Aprendizaje automático y NLP en Ciencias Sociales", exposición en el 2do. Encuentro "Ética en la investigación e Inteligencia Artificial", Comité Nacional de Ética en la Ciencia y la Tecnología Actividades, Ministerio de Ciencia y Técnica (CECTE-MINCyT).
- 2023, "¿Qué son las Ciencias Sociales Computacionales?", charla en el Programa de Investigación en Clases Sociales del Instituto Gino Germani, Facultad de Ciencias Sociales-UBA.
- 2023, "Usos del aprendizaje automático en las Ciencias Sociales. Problemas y potencialidades", clase en seminario del Sistema de Educación a Distancia (SIED-UNSAM).
- 2023, "Un abordaje aplicado a los cambios en el uso del suelo y sistemas de alerta temprana de deforestación.", seminario
- 2023, "Dos casos de aplicación de técnicas de aprendizaje automático en el ámbito estatal", exposición en Simposio *Gestión de datos en organizaciones 1er Encuentro Interdisciplinar*, Facultad de Ciencias Económicas, UBA.
- 2022, "¿Qué son las Ciencias Sociales Computacionales?" charla en el marco de la Semana de la Ciencia, EIDAES-UNSAM.
- 2022 *¿De qué habla la música popular argentina? Generando un corpus de letras de rock y analizándolo con minería de texto*, taller en el marco de la Semana de la Ciencia, EIDAES-UNSAM.
- 2022, *Desarrollo capitalista y pobreza. Otra forma de medirlos*, charla en Desprogramando la desigualdad, factor~data / Ecofeminita.
- 2022, *El "estado de ánimo" de les adultes mayores durante la pandemia. Un caso de aplicación de NLP*, charla en Big Data para el desarrollo sostenible - Fortalecimiento de capacidades técnicas en Latinoamérica y el Caribe, Agencia Española de Cooperación Internacional para el Desarrollo (AECID).
- 2021, *Discursos, mapas y problemas en torno a la pandemia*, charla en panel Hechos, datos, decisiones. La dimensión tecnológica del cruce pandemia-sociedad, Centro Ciencia y Pensamiento de la UNSAM.
- 2021, *Ciencias Sociales y Ciencia de Datos. Nuevas herramientas, problemas clásicos*, charla en Jornadas sobre Herramientas Estadísticas, UNTREF
- 2021, *Ciencias Sociales y Ciencia de Datos: nuevas herramientas aplicadas a problemas clásicos*, charla en Diplomatura en Ciencia de Datos, Inteligencia Artificial y sus aplicaciones en Economía y Negocios, Universidad Nacional de Córdoba
- 2020, Los "temas" de la pandemia. Un análisis de los diálogos de lectores de medios nacionales, Seminario IC/ICC: La ciencia ante el COVID-19, Instituto de Ciencias de la Computación / Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, UBA.
- 2020, ¿Qué comenta "la gente" cuando piensa en COVID-19? Un análisis automático de los discursos de lectores en noticias sobre COVID-19 de cinco diarios argentinos, Instituto de Altos Estudios Sociales, Universidad Nacional de San Martín.
- 2020, *Ciencias Sociales y Ciencia de Datos. Nuevas herramientas, problemas clásicos*, Seminario General, Instituto de Altos Estudios Sociales, Universidad Nacional de San Martín.
- 2019, *Presentación del paquete de R eph*, Instituto de Investigaciones Gino Germani, Facultad de Ciencias Sociales, Universidad de Buenos Aires.
- 2019, *Procesamiento de Lenguaje Natural como herramienta para el análisis del conflicto social. Algunas ideas*, III Taller sobre Conflictividad Laboral y Social, Instituto de Investigaciones Gino Germani, Facultad de Ciencias Sociales, Universidad de Buenos Aires.
- 2019, *Introducción (simple) a Machine Learning*, Taller en II Workshop Big Data en Ciencias Económicas, Facultad de Ciencias Económicas, Universidad Nacional de Córdoba.,
- 2019, *Modalidades de expansión de la frontera agrícola en Argentina. Una aproximación en base a fuentes secundarias*, Seminario del Instituto de Matemática Aplicada, Universidad Nacional de San Luis.
- 2019, "Cantando el tango como ninguna. Una aplicación de topic modeling y algunos otros experimentos en letras de tango", *PyData San Luis Meetup*, Instituto de Matemática Aplicada, Universidad Nacional de San Luis.

- 2018 “Mapeando la Enfermedad de Chagas con Big Data y datos públicos”, Digital Summit, Digital House.
- 2018 “Machine Learning en las Ciencias Sociales. Algunas nociones y posibilidades de uso”, expositor invitado al *Seminario del Programa de Investigación sobre Clases Sociales*, Instituto de Investigaciones Gino Germani, Facultad de Ciencias Sociales, UBA.
- 2018 “Cuando los algoritmos fallan. ¿Qué es el ‘sesgo algorítmico’?”, orador en *Empleos del Futuro*, evento organizado por el Gobierno de la Ciudad Autónoma de Buenos Aires y Digital House.
- 2018, “Un aporte a la confusión: Machine Learning, Inteligencia Artificial y Data Science.”, expositor en *Arduino Day 2018*.
- 2018, “Maximizando el ROI de los datos mediante técnicas de imputación de datos perdidos con Ensamble Learning”, Coordinador de Workshop en *Globant Converge 2018*.
- 2018, “Data Science, Machine Learning e Inteligencia Artificial. Delimitando los campos”, expositor invitado en el *Colegio de Graduados de la Facultad de Ingeniería de la Universidad de Buenos Aires*.
- 2017, “Imputando datos perdidos con métodos de Ensamble Learning”, expositor en evento *Data Science Aplicada a Problemas Sociodemográficos*, Digital House.

13. Premios y distinciones

- 2023, Programa de producción académica de libros TITSHONHY-UNSAM
- 2017, 9no. lugar (de 40) en *Competencia de Predicción de Datos*, *Escuela de Ciencias Informáticas 2017*, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.
- 2012-2014, *Beca Interna Doctoral* Tipo II, Consejo Nacional de Investigaciones Científicas y Técnicas.
- 2011, *Beca para estancia de investigación* en “desiguALdades”, Freie Universität Berlin
- 2009-2012, *Beca Interna Doctoral* Tipo I, Consejo Nacional de Investigaciones Científicas y Técnicas.
- 2009, *Mejor promedio* de la primera promoción del año 2009 en la Facultad de Ciencias Sociales, Universidad de Buenos Aires.
- 2008, *Diploma de honor* otorgado por la Universidad de Buenos Aires, Licenciatura en Sociología.
- 2001, *Mejor promedio* de la promoción 2001 en el Colegio Nacional de Buenos Aires.