

## Taller de Introducción a los Métodos Computacionales para las CC.SS

### Sumilla

El objetivo del taller es que los/as estudiantes se familiaricen con los distintos métodos computacionales y su aplicación en las ciencias sociales. En este sentido, el curso no busca el desarrollo de habilidades específicas sobre un software u método en específico, por lo contrario, la idea es poder desarrollar una vista panorámica de las posibilidades (y limitaciones) así como su aplicación concreta a preguntas/fenómenos sociales relevantes.

El taller busca ser teórico-práctico. La idea es que a través del ejemplo de investigaciones empíricas, el/la participante pueda generar una noción sobre cómo plantear una investigación desde una mirada “computacional”. Antes que dar un “conjunto” de herramientas técnicas, el taller busca desarrollar la capacidad de los/as asistentes para incorporar métodos computacionales en sus propios temas/preguntas de investigación.

### Objetivos específicos

Al finalizar el taller el/la estudiante podrá realizar lo siguiente:

1. Incrementar su conocimiento sobre las posibilidades y limitaciones de los métodos computacionales para el análisis social.
2. Identificar una pregunta de investigación social y cómo abordarla aplicando métodos computacionales.
3. Aprender nociones básicas sobre cómo extraer información de la web.
4. Aprender nociones básicas sobre cómo analizar la información extraída.

### Distribución de sesiones

Sesión	Día
Sesión 1: Introducción y ejemplos de investigaciones sociales con métodos computacionales.	14 de septiembre del 2024 2 horas
Sesión 2: Pregunta, Teoría y Metodología: ¿Cómo se vinculan? (análisis de casos)	15 de septiembre del 2024 2 horas
Sesión 3: Técnicas de recojo de datos en la web	21 de septiembre del 2024 3 horas
Sesión 4: Análisis de los datos recogidos	22 de septiembre del 2024 3 horas
Sesión 5: Aplicación de lo aprendido a los propios proyectos de los/as participantes	28 de septiembre del 2024 2 horas

### Uso de softwares

Si bien el objetivo del taller no es el desarrollo técnico de algún software en particular, el curso requiere el uso de softwares para ejemplificar lo que se discute en las sesiones teóricas. Para ello, se utilizarán los softwares RStudio y Anaconda. Los códigos se proporcionarán al inicio de la sesión de modo que los/as participantes puedan seguir el desarrollo de la clase sin mayores dificultades.

**Material del curso**

Todo el material del curso estará disponible en el siguiente repositorio de Github:  
<https://github.com/rsquevaraa/intro-met-comp>

**Sobre el docente**

Renato Guevara Ayón

Soy sociólogo de la PUCP con una maestría en análisis de datos sociales por la Universidad de Lund (Suecia). Tengo experiencia en uso de metodologías cuantitativas y cualitativas en temáticas de educación, género y migración. Interesado en aprender, aplicar y desarrollar nuevas formas de recoger y analizar datos en ciencias sociales.

**Información de contacto**

Para cualquier consulta o duda referente al contenido del curso pueden escribirme al siguiente correo [rsquevaraa@gmail.com](mailto:rsquevaraa@gmail.com)

Para cualquier duda respecto a temas logísticos o de certificados contactarse directamente con [les.pucp@gmail.com](mailto:les.pucp@gmail.com) .

## Sesiones detalladas

### Sesión 1: Introducción y ejemplos de investigaciones sociales con métodos computacionales.

Objetivo: Introducir la definición de métodos computacionales y cómo se aplican en las ciencias sociales. Para ello, se presentarán dos investigaciones empíricas de cómo estos métodos son usados para responder preguntas sociales.

Material de apoyo:

#### Lecturas teóricas

- Hox, J. J. (2017). Computational Social Science Methodology, Anyone? *Methodology*, 13(Supplement 1), 3–12. <https://doi.org/10.1027/1614-2241/a000127>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Computational Social Science. *Science*, 323(5915), 721–723. <https://doi.org/10.1126/science.1167742>

#### Trabajos empíricos (ejemplos)

- Dömsödi, L. B. (2024). *Exploring the Role of Online Social Support in Eating Disorder Recovery: A Quantitative Computational Study* [Lund University]. <https://lup.lub.lu.se/student-papers/search/publication/9152004>
- Guevara Ayón, R. (2024). *Representing the “Other”: How do Peruvian right-leaning newspapers represent Venezuelan immigrants (2017-2023)?* [Lund University]. <https://lup.lub.lu.se/student-papers/search/publication/9158373>
- Malik, M. (2024). *Deliberation in the Age of Deception: Measuring Sycophancy in Large Language Models* [Lund University]. <https://lup.lub.lu.se/student-papers/search/publication/9151763>

### Sesión 2: Pregunta, Teoría y Metodología: ¿Cómo se vinculan? (análisis de casos)

Objetivo: En esta sesión volveremos a los ejemplos empíricos expuestos en la sesión anterior y los analizaremos a detalle. Esto implica un análisis de cómo se relacionan la pregunta planteada, el marco teórico y la metodología. La idea es poder entender cuál es la lógica detrás de las decisiones que los/as autores toman para luego poder (re)pensar nuestras propias investigaciones desde una mirada crítica.

Material de apoyo:

- Dömsödi, L. B. (2024). *Exploring the Role of Online Social Support in Eating Disorder Recovery: A Quantitative Computational Study* [Lund University]. <https://lup.lub.lu.se/student-papers/search/publication/9152004>
- Guevara Ayón, R. (2024). *Representing the “Other”: How do Peruvian right-leaning newspapers represent Venezuelan immigrants (2017-2023)?* [Lund University]. <https://lup.lub.lu.se/student-papers/search/publication/9158373>

### Sesión 3: Técnicas de recojo de datos en la web (3 horas)

**Para esta sesión y la siguiente utilizaremos Python y R, por lo que recomiendo familiarizarse con estos programas antes de las sesiones.**

Objetivo: En esta sesión veremos diferentes formas de obtener información desde la web. Específicamente, aprenderemos 1) cómo interactuar con APIs, y 2) cómo extraer información directamente de las páginas cuando no existen APIs disponibles. Finalmente, conversaremos sobre las consideraciones éticas que se deben tener al momento de utilizar estas técnicas.

Material de apoyo:

Lecturas

- Li, F., Zhou, Y., & Cai, T. (2021). Trails of Data: Three Cases for Collecting Web Information for Social Science Research. *Social Science Computer Review*, 39(5), 922–942. <https://doi.org/10.1177/0894439319886019>
- Luscombe, A., Dick, K., & Walby, K. (2022). Algorithmic thinking in the public interest: Navigating technical, legal, and ethical hurdles to web scraping in the social sciences. *Quality & Quantity*, 56(3), 1023–1044. <https://doi.org/10.1007/s11135-021-01164-0>

API's

- MediaWiki API oficial de Wikipedia: [https://www.mediawiki.org/wiki/API:Main\\_page](https://www.mediawiki.org/wiki/API:Main_page)
- Reddit API: <https://www.reddit.com/dev/api/>

### Sesión 4: Análisis de los datos recogidos (3 horas)

Objetivo: En esta sesión aprenderemos cómo procesar los datos recogidos de la web. Para ello, nos centraremos específicamente en el aprendizaje de técnicas básicas para analizar texto: Procesamiento de Lenguaje Natural (NLP por sus siglas en inglés). Veremos normalización de texto, vectorización, TF-IDF, uso de word embeddings, así como formas de visualizar los datos recogidos a través de gráficos.

Material de apoyo:

- Hovy, D. (2020). *Text analysis in Python for social scientists: Discovery and exploration*. Cambridge University Press.
- Kedia, A., & Rasu, M. (2020). *Hands-On Python Natural Language Processing: Explore tools and techniques to analyze and process text with a view to building real-world NLP applications*. Packt.

### Sesión 5: Aplicación de lo aprendido a los propios proyectos de los/as participantes

Objetivo: Esta sesión es de tipo seminario, ello quiere decir que la participación de los/as asistentes es vital. Cada participante debe llegar con una idea/pregunta relacionada a un tema de investigación en ciencias sociales de su interés; y en grupos, diseñaremos una propuesta de investigación para responder esta pregunta usando métodos computacionales.

**IMPORTANTE:** Es indispensable que el/la asistente al taller tenga una pregunta o tema específico que quiera investigar. Cuanto más delimitado esté el tema es mejor. Los/as participantes son bienvenidos a poder traer sus ideas de tesis o de investigaciones en otros cursos previos.

Material de apoyo:

- Todo el material previo.