

Cómo presentar un proyecto a LatinX in NLP

Diana Galván Sosa
Research Associate @ University of Cambridge
LXNLP 2024 General co-chair
NAACL 2024 Ethics co-chair

 dianags@latinxinai.org

 [dianags.github.io](https://github.com/dianags)

¿Qué es LXAI (LatinX in AI) ?

- Una comunidad global de profesionales en Inteligencia Artificial (IA) de origen latinoamericano
- La misión de LXAI es impulsar la investigación y el desarrollo académico y profesional de sus miembros
- LXAI ofrece programas de mentoría, workshops y bolsa de trabajo



¿Qué es NAACL?

- Capítulo Norteamericano de la Asociación (internacional) de Lingüística Computacional
- Las conferencias de ACL son de las más importantes en el área de Procesamiento de Lenguaje Natural (PLN)



Association for
Computational Linguistics

¡Por primera vez NAACL será en México!

LXAI @ NAACL



Conferencia principal (NAACL): 16 al 21 de Junio de 2024

Workshop (LXNLP): 16 de Junio (por confirmar*)

¿Cómo participar en el workshop?

- Formando parte del comité organizador (aplicaciones cerradas)
- Formando parte del equipo de revisores (aplicaciones abiertas)
- Formando parte del equipo de voluntarios (aplicaciones abiertas)
- Mandando un trabajo (aplicaciones abiertas)

Los autores son los protagonistas de todo workshop 😁

Beneficios de participación

- ★ Exposición y divulgación de sus proyectos en una comunidad especializada e internacional
- ★ Contribuir a hacer de la comunidad de PLN una comunidad más diversa



Latin American Meeting
In Artificial Intelligence

- ★ Oportunidad de ganar una de las **2 becas** (vuelo, hospedaje y registro) para asistir a Khipu 2025 en Chile

¿Qué ayuda ofrece LXAI?

- **Traducción** (Español -> Inglés, Portugués -> Inglés)
- **Apoyo económico** (1 autor por equipo)
 - Transporte
 - Hospedaje
 - Comidas
 - Registro a la conferencia principal

Hacemos todo lo posible por cubrir el 100% de gastos 

Recepción y revisión de trabajos

1. Preparar un extended abstract

- Formato: PDF
- El PDF debe ser **anónimo** (i.e., sin nombres de los autores)
- Se debe usar el template de Latex para generar el PDF
- Mandar el PDF a través de nuestra página de Microsoft CMT

2. Revisión doble ciego (double blind)

- Nombre de los autores: anónimo
- Nombre de los revisores: anónimo

3. Notificación de aceptación

- Se notificará por correo y/o a través de CMT
- Los revisores asignaran una calificación e incluirán comentarios

4. Preparar la versión final del extended abstract (camera-ready version)

- Incluir los cambios recomendados por los revisores

5. Presentación

- En línea o en persona
- Idioma: Inglés
- Formato: póster y/o diapositivas (por confirmar)

Más info en el sitio web del workshop

Características de un extended abstract

	Abstract	Extended Abstract	Full paper
Extensión	200 ~ 300 palabras	2 o 3 páginas	4 páginas (corto)* 7 páginas (largo)*
Objetivo	<u>Resumir</u> la motivación, estructura y contribuciones del proyecto	<u>Describir en términos generales</u> la motivación, estructura y contribuciones del proyecto	<u>Describir detalladamente</u> la motivación, estructura y contribuciones del proyecto
Estructura	Introducción Objetivo Método Resultado principal Conclusiones	Introducción Antecedentes Método Resultados Discusión Conclusiones	Introducción Antecedentes Método Resultados Discusión Conclusiones
Figuras y/o tablas	No	Si	Si
Referencias	No	Si	Si

**De acuerdo a los lineamientos de ACL. Puede variar en otras disciplinas/conferencias*

Estructura de un extended abstract

- Introducción
 - Objetivo
 - Antecedentes
 - Método
 - Resultados
 - Discusión
 - Conclusiones
- ¿De qué trata el proyecto?
 - ¿Qué tema se está abordando?
 - ¿Cuál es la mayor contribución de este trabajo?

Estructura de un extended abstract

- Introducción
 - Objetivo
 - Antecedentes
 - Método
 - Resultados
 - Discusión
 - Conclusiones
- ¿Qué problema se está tratando de resolver?
 - ¿Cuál es la motivación detrás del proyecto?

Estructura de un extended abstract

- Introducción
 - Objetivo
 - Antecedentes
 - Método
 - Resultados
 - Discusión
 - Conclusiones
- ¿Por qué es relevante trabajar en el tema?
 - ¿Quiénes han tratado de resolver el mismo problema?
 - ¿Cómo han tratado de resolverlo?
 - ¿En qué es diferente este trabajo a los anteriores?

Estructura de un extended abstract

- Introducción
- Objetivo
- Antecedentes
- Método
 - ¿Cómo propone este trabajo solucionar el problema?
 - ¿Cómo se diseñaron los experimentos?
 - ¿Cómo se hizo la recolección de datos?
- Resultados
- Discusión
- Conclusiones

Estructura de un extended abstract

- Introducción
 - Objetivo
 - Antecedentes
 - Método
 - Resultados
 - Discusión
 - Conclusiones
- ¿Cuáles eran los resultados esperados?
 - ¿Cuáles fueron los resultados obtenidos?

Estructura de un extended abstract

- Introducción
 - Objetivo
 - Antecedentes
 - Método
 - Resultados
 - **Discusión**
 - Conclusiones
- ¿Cómo se pueden interpretar los resultados?
 - ¿Cuáles fueron las lecciones aprendidas (hallazgos)?

Estructura de un extended abstract

- Introducción
 - Objetivo
 - Antecedentes
 - Método
 - Resultados
 - Discusión
 - Conclusiones
- Resumen del objetivo, método y resultados
 - ¿Hay nuevas preguntas por responder?
 - ¿Cuáles son los siguientes pasos a seguir en este proyecto?

Consideraciones generales

- **Sean concisos**
 - Un extended abstract es un texto científico. Eviten “dejar lo mejor para el final”
- **No traten de sonar “inteligentes”**
 - Usen lenguaje simple. No es necesario decir “utilizar” en lugar de “usar”
- **No exageren las contribuciones de sus trabajos**
 - No se trata de publicidad. Lo importante es ver un trabajo bien estructurado
- **Usen ejemplos siempre que sea posible**
 - Contribuye a que su texto sea más fácil (e interesante) de leer
- **Determinen el orden de los autores**
 - Si el trabajo es aceptado, será necesario saber los nombres
 - El orden indica el nivel de contribución
- **Incluyan agradecimientos**
 - Si el trabajo es aceptado, es buena práctica agradecer a los revisores
 - En general, se agradece a quienes hicieron el proyecto posible (apoyo económico y/o recursos computacionales)
- **Citen trabajos publicados; no pre-prints**
 - Una referencia de una conferencia/journal es preferible a arXiv
 - Exepción: el trabajo que quieren citar sólo existe en arXiv

Buenos ejemplos de descripción de un LM

▼ Crea una base de datos y entrena tu propio LLM

Al unirse a este hackathon tendrás la oportunidad de colaborar con la creación de LLMs de calidad e inclusivos en tu lengua. Aplica tus conocimientos para superar los retos de cada etapa de desarrollo de tu LLM: creación del corpus, entrenamiento y evaluación.

- CamemBERT: <https://camembert-model.fr/>
- BERTurk: <https://github.com/stefan-it/turkish-bert>
- BETO: <https://github.com/dccuchile/beto>

Materiales interesantes

- **Tips for writing NLP papers (blogpost de Vered Shwartz)**
<https://medium.com/@vered1986/tips-for-writing-nlp-papers-9c729a2f9e1f>
- **Primera edición de este taller (presentado por Javier Turek)**
https://www.youtube.com/live/QziYfITvGrA?si=e0b5SnRCGTnG_aXy&t=1180
- **Create a dataset card (docs de Hugging Face)**
https://huggingface.co/docs/datasets/dataset_card

Contacto



latinxinai.org



linkedin.com/company/latinx-in-ai



latinxinai.org/naacl-2024



[@lxnlp](https://twitter.com/lxnlp)



lxnlp@latinxinai.org

¡Esperamos sus trabajos!