



UNIVERSIDAD COMPLUTENSE MADRID

Proyecto de Innovación
Convocatoria 2023/2024

Nº de proyecto: 172

Análisis crítico del uso de la inteligencia artificial para resúmenes
científicos

Responsable del Proyecto: Ángel Cuesta Martínez

Facultad de Farmacia

1. Objetivos propuestos en la presentación del proyecto

A lo largo de esta memoria se utiliza el género gramatical masculino y femenino indistintamente para referirse a colectivos mixtos, siguiendo el principio lingüístico de la economía expresiva y sin pretender discriminar a nadie por razón de sexo.

El objetivo principal del proyecto 172 “Análisis crítico del uso de la inteligencia artificial para resúmenes científicos”, es el desarrollo de las capacidades de lectura y análisis crítico (con rigor y fundamentado) de un texto de contenido científico generado por la Inteligencia Artificial (IA).

Brevemente, para que el estudiante pueda tener capacidad crítica de análisis debe haber adquirido previamente unos conocimientos científicos concretos que le sirvan de base para analizar, madurar y comparar/juzgar. Durante el grado (de Farmacia en el caso del proyecto 172), una alumna media puede carecer de la formación suficiente para acceder a las fuentes de información y del tiempo necesario para formarse en un tema específico. Sin embargo, y basándonos en las competencias básicas, generales y transversales indicadas en la ficha docente del Trabajo de Fin de Grado (TFG) en el grado de Farmacia, este alumnado de TFG es formado y evaluado por estos mismos conceptos, por lo que sí van a adquirir dichos conocimientos. Sucede, además, que estos TFGs son de carácter bibliográfico, con lo que la IA les puede proporcionar un texto similar al que el estudiante debe elaborar. Como alternativa y a modo de ensayo piloto a menor escala, intentamos la realización de un pequeño número de TFGs de carácter experimental tutorizados por profesoras de la Facultad de Medicina, con el fin de explorar su escalabilidad a trabajos experimentales. Lamentablemente, carecemos de respuestas en la parte final, una vez acabado el TFG, por lo que no se van a reflejar ni analizar en esta memoria final.

En este tipo de TFGs bibliográficos que se planea en el proyecto de innovación docente, las alumnas tienen que completar las tareas habituales de un trabajo de este tipo más un resumen que incluye conclusiones originales de los textos analizados o resultados obtenidos.

Los objetivos concretos del proyecto 172 son:

- 1) Familiarizar al alumnado y profesorado con las herramientas de IA y su uso para la escritura de resúmenes científicos.
- 2) Desarrollar y afianzar en el alumnado la capacidad de análisis crítico de textos científicos.
- 3) Evaluar el impacto de la IA como herramienta docente para la escritura de resúmenes científicos.

2. Objetivos alcanzados

Objetivo general: “desarrollo de capacidades de lectura y análisis crítico de un texto de contenido científico generado por la IA”.

20 estudiantes han participado en la encuesta inicial y, de ellos, 14 han completado la encuesta final, no contamos con las 9 alumnas del TFG experimental llevado a cabo en la Facultad de Medicina al no responder la encuesta final.

Aún con todo, estos números indican un alto porcentaje de participación en las encuestas, de carácter voluntario y anónimas, pues contamos con 16 docentes como miembros del proyecto.

Basándonos en la contestación a las encuestas realizadas antes y después del comienzo del uso de la IA en su TFG, las respuestas demuestran que los alumnos han descubierto flaquezas en la generación de textos por parte de la IA, se muestran críticos en aspectos como el rigor científico y calidad gramatical de los textos generados. Estos datos serán mostrados y analizados en el punto 5 “Desarrollo de las actividades”.

Por ello consideramos que este objetivo general ha sido alcanzado.

Objetivos concretos:

1) Familiarizar al alumnado y profesorado con las herramientas de IA y su uso para la escritura de resúmenes científicos.

Alumnas y profesoras, una vez contestada la encuesta inicial sobre sus impresiones acerca de la IA, recibieron una charla inicial describiendo el proyecto y las potenciales herramientas de la IA (Actividad 3 “Taller formativo para el aprendizaje del uso de la IA para la creación de textos científicos”), llevada a cabo por Dr. Daniel Pattier Bocos, profesor Doctor en Educación y Profesor y director del Máster en Programación e Implementación de Proyectos Educativos (UCM) el 27 de noviembre de 2023. Hay que indicar que es muy importante que el taller se realizara después de la encuesta, para no sesgar la opinión de las participantes (tanto alumnas como profesores).

En este taller se informó sobre los distintos sistemas de acceso gratuitos a la IA. Así mismo, se proporcionaron instrucciones/indicaciones básicas para conversar con el chat de manera eficaz y, por último, se indicaron los potenciales riesgos, sesgos o errores que han sido detectados por expertos en el uso de la IA.

Por ello consideramos que este objetivo concreto ha sido alcanzado.

2) Desarrollar y afianzar en el alumnado la capacidad de análisis crítico de textos científicos.

Esta parte del objetivo fue desarrollada de manera particular entre cada tutor de TFG con sus estudiantes. Los profesores instruyen sobre el tema de trabajo a desarrollar, proporcionándoles inicialmente una bibliografía básica sobre la cual las alumnas deben desarrollar el texto, así como acudir a otras fuentes bibliográficas en donde se contrastan datos científicos y clínicos y técnicas experimentales. En estos momentos es donde el alumno comienza a asentar conocimientos, en ocasiones enfrentados, y a desarrollar su capacidad de análisis crítico.

Durante el desarrollo del TFG, tutoras y alumnos han tenido varias reuniones en las cuales se ponen en común dudas y se supervisa el progreso del TFG. Finalmente, el estudiante de TFG analiza un resumen generado por la IA y decide si lo acepta, lo

rechaza y sustituye por uno propio o lo modifica parcialmente. La encuesta final de los estudiantes indica el uso y la valoración del texto que la IA ha generado sobre el tema en el que se ha formado.

Teniendo en cuenta las contestaciones iniciales y finales de las estudiantes con respecto a la potencia de la IA generando textos de calidad gramatical y científica, podemos indicar que se ha producido un cambio, una evolución, de las respuestas, de la opinión, lo que pone de manifiesto que se ha producido un aprendizaje de un tema científico concreto y que esa formación ha venido acompañada de un proceso de análisis crítico con el propio texto y el generado por la IA. Estos datos serán mostrados y analizados en el punto 5 “Desarrollo de las actividades”.

Así, consideramos que este objetivo concreto ha sido alcanzado. Ciertamente, el número de muestra (20 iniciales y 16 finales), no es representativo como para sacar conclusiones de toda la comunidad universitaria. Por este motivo, para aumentar el número, se solicitó (y concedió) la extensión de este proyecto por un año más (2024-2025).

3) Evaluar el impacto de la IA como herramienta docente para la escritura de resúmenes científicos.

Esta parte incluye no solo las contestaciones y comentarios de las alumnas, cuenta también con las respuestas de sus tutores, tanto antes como después de haber tutorizado el TFG.

Con respecto al profesorado, con capacidad crítica desarrollada, la clave está en la confianza que la IA ha generado después de su uso en el TFG. Se observa un cambio en la aplicabilidad de la IA para la generación de textos científicos, pasando de un 43% de opinión neutral o negativa en la encuesta inicial a un 24% en la final, que se corrobora al reducirse a la mitad el porcentaje de profesoras que no recomendarían su uso. De la misma manera, su uso y conocimiento ha incrementado de un 93,4% a un 100% la idea de que las alumnas deben ser informadas y supervisadas en el uso de la IA.

Así, un 37,5% se encontraban preparados para informar y guiar a sus alumnas al comienzo, siendo un notable 75% al final de la encuesta.

Sorprendentemente, después de su uso, el profesorado mantiene su desconfianza sobre su capacidad de generar un texto con mayor rigor científico.

Los estudiantes carecen de la experiencia del profesorado, pero cuentan con una alta capacidad de uso de las nuevas tecnologías que facilita el uso y comprensión de estas nuevas tecnologías. Quizá por ello sí creen que se generan textos con mayor rigor científico después del TFG (de un 10% a un 21%) y un 28% consideran que la IA realizó un mejor texto que ellos, habiéndolo cambiado o modificado un 50%.

Un 72% de ellas recomendaría el uso de la IA para la generación de textos científicos.

Por ello consideramos que este objetivo concreto ha sido alcanzado parcialmente, pues sucede que el número total de profesoras y alumnas encuestadas es pequeño como para obtener conclusiones con peso estadístico definitivo. La posibilidad de aumentar el número depende de la posibilidad de ampliación de colaboradores a profesoras de la facultad de Farmacia en la continuación del proyecto en el año 204-2025.

3. Metodología empleada en el proyecto

Como se describe en el proyecto, se han realizado 6 actividades principales. Para ello, la metodología ha sido una combinación de reuniones iniciales para definir y consensuar los criterios de trabajo y fijar entre todos la mejor manera de plantear la consecución de los objetivos propuestos. Se encuadrarían en la Actividad 1 “Reunión inicial de coordinación del profesorado” y la Actividad 3 “Taller formativo para el aprendizaje del uso de la IA para la creación de textos científicos”, realizada el 27 de noviembre de 2023, en donde participaron tanto profesoras como alumnos.

El alumnado, al tratarse de un TFG, ha trabajado de forma individual, aunque ha sido supervisado por su tutor en momentos puntuales a lo largo del año.

De manera individual, pero en un mismo período de tiempo, se desarrolló la Actividad 2 “Evaluación del conocimiento de las herramientas de IA entre las estudiantes y profesoras que participan en el proyecto”. En donde todas completaron una encuesta inicial en la que responderán preguntas en base a una escala de Likert (1, Totalmente en desacuerdo - 5, Neutral y 10, Totalmente de acuerdo), respuestas binarias (sí/no) o pregunta abierta. Los estudiantes desarrollaron de manera individual y a expensas del tutor, la Actividad 4 “Generación de resúmenes y evaluación del texto creado por la IA en base a su rigor científico y la calidad lingüística”.

Una vez creado el TFG y consultado a la IA, las alumnas completarán una breve encuesta final en la que responderán las siguientes preguntas en base a una escala de Likert (1, Totalmente en desacuerdo - 5, Neutral y 10, Totalmente de acuerdo), referentes a su experiencia en el uso de la IA.

El equipo que ha llevado a cabo el proyecto ha logrado llevar a cabo la Actividad 6 “Diseminación de resultados en congresos, publicaciones y redes sociales”.

Se ha participado en la sesión “Innovació en metodologies docents per al professorat dels àmbits docents vinculats amb el Departament de Bioquímica i Biotecnologia” (octubre de 2023, Universitat Rovira i Virgili); con la ponencia “Presentació d'un projecte d'innovació docent que vol utilitzar la IA en l'autoavaluació dels TFGs i pretén ensenyar als estudiants com interactuar amb aquesta tecnologia de manera crítica i reflexiva.

Se ha participado en la Semana de la Ciencia 2023 con la ponencia “La IA para la redacción de trabajos científicos” (noviembre de 2023, Facultad de Educación de UCM); ponencia vinculada al proyecto PIE205 ‘ChatGPTeaching: Usos éticos y pedagógicos de la IA en la docencia universitaria’.

Se ha participado en forma de póster y presentación en el III Congreso Internacional de Innovación Docente, Educación y Transferencia del Conocimiento (CIINECO). Debates y experiencias en torno a la innovación docente: visión actual y prospectiva (abril de 2024); con la ponencia “Hacia una pedagogía innovadora y actual: evaluación del uso de la inteligencia artificial en estudiantes de farmacia de la Universidad Complutense de Madrid”. La publicación formará parte del libro: "Revolucionando la docencia universitaria: Innovación educativa en la era de la IA y la Gamificación", con ISBN reservado "978-84-1070-329-2", que será publicado en diciembre de 2024.

4. Recursos humanos

Un total de 16 profesores (Catedráticas, titulares, Contratados doctores/ Permanentes laborales, Ayudantes Doctores, Asociados e investigadores Ramón y Cajal) de dos facultades (Farmacia y Medicina) han participado en este proyecto.

Un total de 20 estudiantes del grado de Farmacia (más 9 del grado de Medicina), han participado total o parcialmente en este proyecto.

Responsable Cuesta Martínez, Ángel. angcuest@ucm.es

Miembro Escribano Illanes, Óscar. PDI Complutense oescriba@ucm.es

Miembro Esteras Gallego, Noemi. PDI Complutense nesteras@ucm.es

Miembro García Redondo, Alberto. PDI Complutense alberg04@pdi.ucm.es

Miembro Gómez Hernández, María De La Almudena. PDI Complutense algomez@ucm.es

Miembro González Gálvez, Beatriz. PDI Complutense bggalvez@ucm.es

Miembro Hernández Sánchez, María. PDI Complutense marher36@ucm.es

Miembro Iniesta Serrano, María Pilar. PDI Complutense insepi@farm.ucm.es

Miembro De Juan Chocano, María Del Carmen. PDI Complutense juchoca@ucm.es

Miembro Linares Gómez, María. PDI Complutense mlinares@ucm.es

Miembro Martín De Saavedra, María Dolores. PDI Complutense dolores.ms@ucm.es

Miembro Martínez Ruiz, Antonio. PDI Complutense amartinezruiz@ucm.es

Miembro Musteanu, Mónica Andrea. PDI Complutense mmustean@ucm.es

Miembro Navarro González De Mesa, Elisa. PDI Complutense elisnava@ucm.es

Miembro Oset Gasque, María Jesús. PDI Complutense mjoset@ucm.es

Miembro Sánchez Muñoz, Aránzazu. PDI Complutense munozas@farm.ucm.es

5. Desarrollo de las actividades

Consideramos que en los puntos 2 “Objetivos alcanzados” y 3 “Metodología empleada en el proyecto”, de la presente memoria, han quedado explicadas y analizadas la Actividad 1 “Reunión inicial de coordinación del profesorado”; la Actividad 3 “Taller formativo para el aprendizaje del uso de la IA para la creación de textos científicos”; la Actividad 4 “Generación de resúmenes y evaluación del texto creado por la IA en base a su rigor científico y la calidad lingüística”; y la Actividad 6 “Diseminación de resultados en congresos, publicaciones y redes sociales”.

Presentamos aquí el análisis de los resultados obtenidos en la Actividad 2 “Evaluación del conocimiento de las herramientas de IA entre las estudiantes que participan en el proyecto” y la Actividad 5 “Análisis final de las encuestas completadas por las alumnas” que se corresponde con las encuestas iniciales y finales de profesoras y alumnos.

Como comentario general, cabe destacar que se han detectado algunos errores en las contestaciones de profesores y de alumnas (ver punto 6 Anexos). Son errores en respuestas que no siguen la lógica pues deberían repetirse o cambiar en un único sentido entre la encuesta inicial y final. Sin embargo, en el cuestionario final ha disminuye o variado de manera opuesta. Como ejemplo tenemos las respuestas a las preguntas del conocimiento de herramientas de IA en donde en la encuesta inicial nos encontramos con un 50% de profesoras que han utilizado la IA y un 62% en la final, cuando el 100% de los profesores del proyecto han realizado tareas con la IA.

El análisis de las respuestas de las profesoras indica los siguientes datos.

Han participado profesoras de todo rango (desde Catedráticas hasta investigadoras Ramón y Cajal), siendo el rango de edad mayoritaria los 45-54 años.

El uso de la IA ha aumentado al final del proyecto y esto ha generado un ligero aumento en la presunción de la IA como una buena herramienta docente. Porcentualmente no se demuestra al bajar el número de respuestas finales, pero las máximas calificaciones pasan de un 29 a un 37%. En ambos momentos consideran que un buen uso puede contribuir a ser una buena herramienta, siendo incluso aumenta la opinión de permitir su uso en trabajos universitarios.

Al mismo tiempo, demuestran su preocupación al apoyar en una gran mayoría en ambas encuestas (> 90%) que es necesario informar y supervisar a las estudiantes en el uso de la IA. Sin embargo, aumenta de un 37,5 a un 50% los profesores que intentarían no influir en la recomendación, aunque disminuye el de aquéllos que no lo recomendarían.

La preparación mostrada por los profesores muestra un aumento total hacia la mejora en la formación después del año de desarrollo del TFG. Bajo este aprendizaje, se mantiene la desconfianza en el rigor científico de la IA sin encontrar ninguna impresión positiva, como queda comprobado al demostrar el conocimiento de los inconvenientes que la IA ofrece. En los comentarios indican la falta de precisión, errores en las citas bibliográficas y de actualizaciones en la información (este hecho se debe al uso de sistemas gratuitos, los de pago sí están actualizados). No sucede igual cuando se consulta sobre la calidad gramatical, la división de la opinión permanece en ambas encuestas, siendo menos extremas en la segunda encuesta.

Finalmente, todos encuentran el beneficio de su uso como herramienta de apoyo que ofrezca variantes o una base sobre la que trabajar y mejorar un texto científico.

El análisis de las respuestas de las alumnas indica los siguientes datos.

Han participado de manera mayoritaria, en más de un 70%, mujeres. Porcentaje que se correlaciona con el porcentaje de mujeres matriculadas en el grado de Farmacia en los últimos 3 años (73,7% de media en el período 2017-2020, fuente UCM).

Inicialmente, un pequeño porcentaje no conocía ninguna herramienta de IA, pero sorprende que la mitad no hubiera utilizado ninguna herramienta, porcentaje que disminuye ligeramente, cuando debería ser un 100%, otro ejemplo de errores en las contestaciones.

Desconocen en un 60% los potenciales inconvenientes o problemas que puede presentar la IA, quizá esto explica la heterogeneidad en las respuestas sobre la calidad gramatical, donde queda distribuido de igual manera la percepción positiva como negativa, esta percepción se revierte hacia una mejor consideración en la encuesta final, con un 56% de apoyo y neutralidad. Los principales inconvenientes que describen son las inexactitudes, el plagio de frases o la dificultad de asentar conocimientos al ser proporcionados.

Con respecto a la calidad científica, se muestra desconfianza en un 60%, aunque la percepción positiva es de un 10% (puntuado con un 7 sobre 10). Esta desconfianza se mantiene similar una vez realizado el TFG y analizado el resumen generado, solamente un 21% considera que puede hacerlo mejor.

Estos datos tan heterogéneos se repiten cuando se pregunta en particular por el texto generado por la IA para su trabajo de TFG. Hay un ligero aumento en cuanto a la mejora en el rigor científico (un 63% muestra que ha sido igual o mejor) y más acusado en la calidad gramatical con un 72% de opinión favorable.

Estas repuestas se correlacionan con la contestación al mantenimiento o cambio total o parcial del resumen generado por la IA. Aquí, se mantiene la distribución igual en opiniones favorables y desfavorables sobre el rigor científico, pero aumenta en cuanto a los cambios en la calidad gramatical.

Esto último puede que indique un problema más de fondo (formativo de base) que el mero hecho del uso de la IA y el asentamiento de conocimientos académico-científicos.

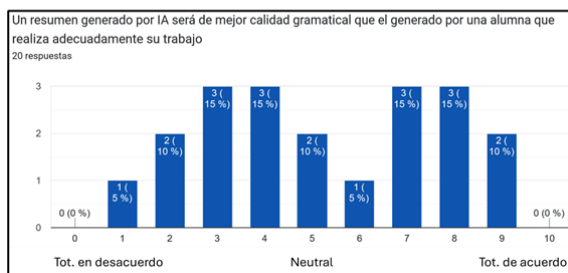
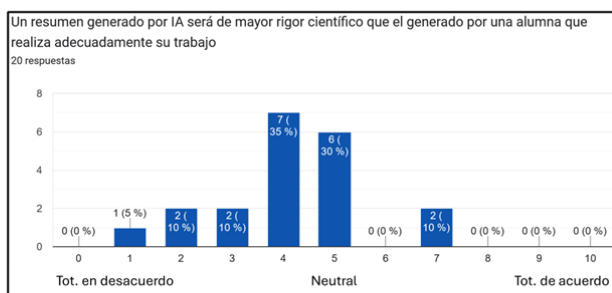
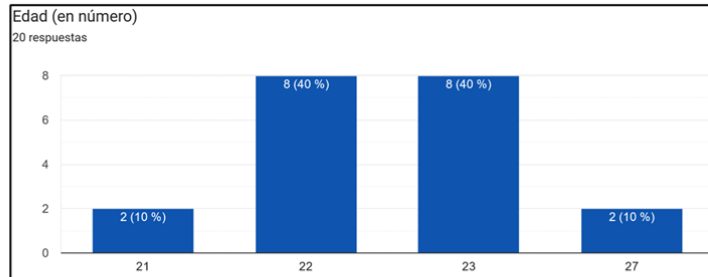
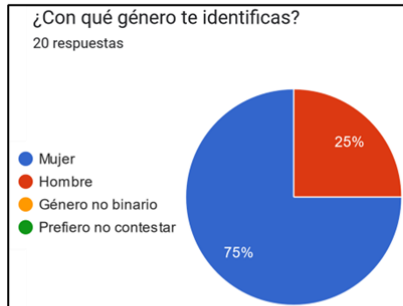
Como ejemplo de la percepción final sobre la IA, un 85% de las alumnas utilizarían en el futuro la IA para la generación de textos científicos y, salvo un 28% que se muestra neutral o ligeramente en desacuerdo (4 sobre 10), el resto están de acuerdo en animar a estudiantes al uso de la IA en la generación de trabajos científicos.

Lo más importante, es que el 100% de las alumnas no están en desacuerdo en cuanto a que ha aumentado su preparación acerca del uso de la IA y su estimación de las ventajas e inconvenientes de su uso.

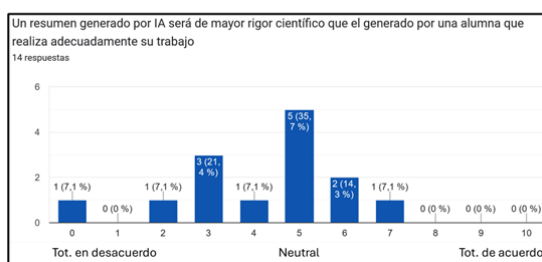
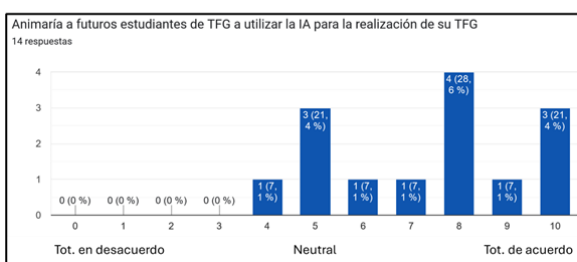
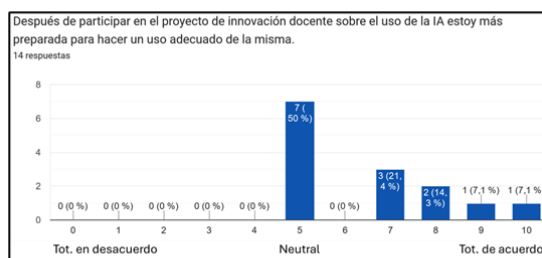
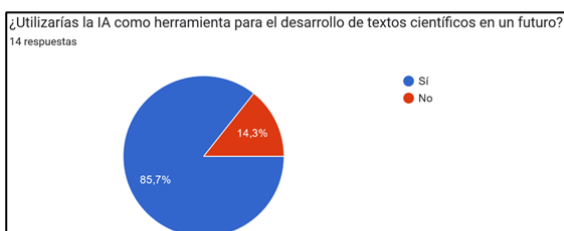
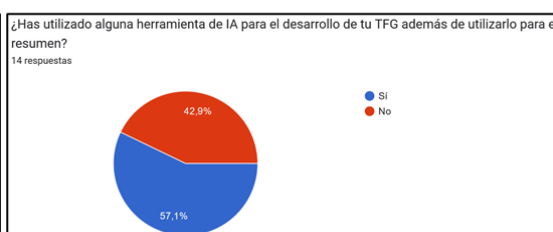
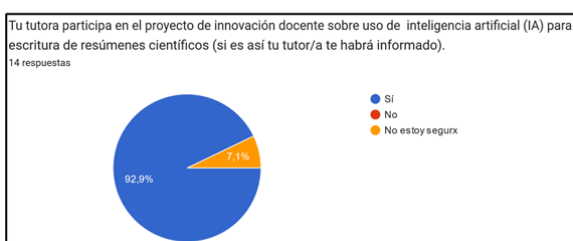
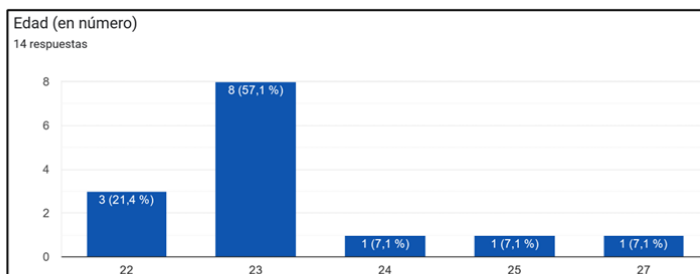
6. Anexos

En las siguientes páginas aparecen los gráficos de datos de las repuestas proporcionadas en las consultas iniciales y finales tanto a alumna como a profesoras. Como se ha indicado anteriormente, no se reflejan las encuestas realizadas en los TFGs experimentales al no tener ninguna respuesta en la parte final.

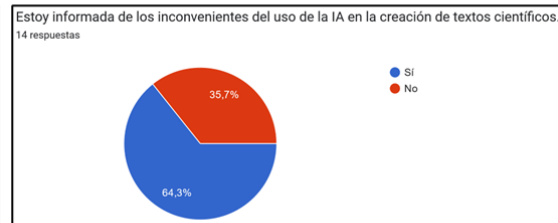
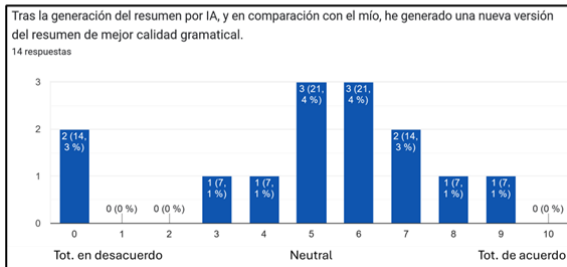
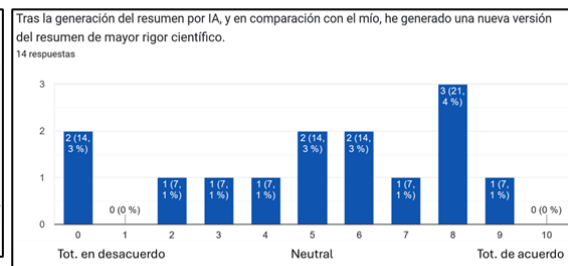
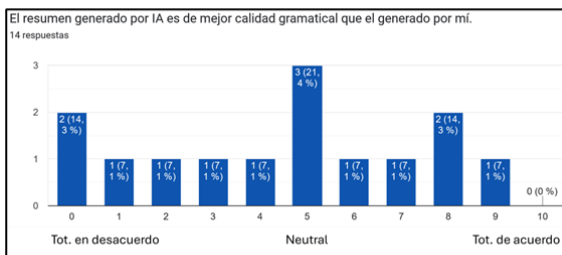
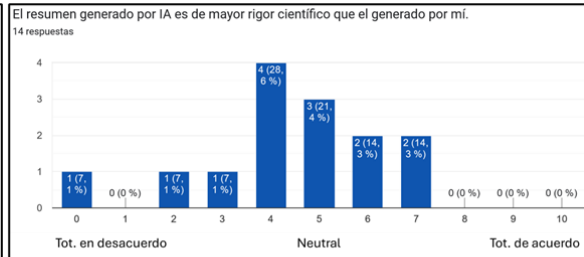
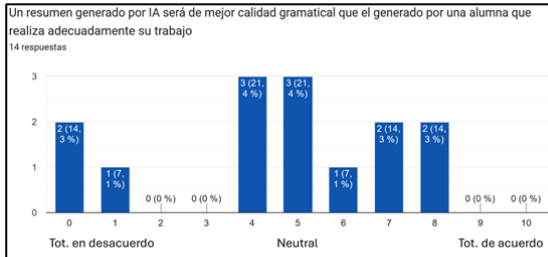
Encuesta inicial IA - Alumnas



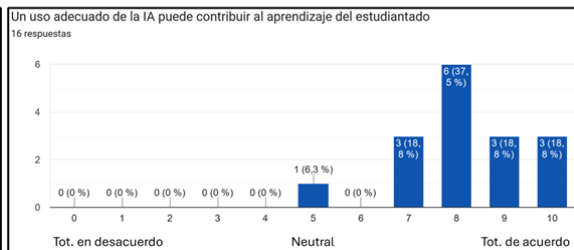
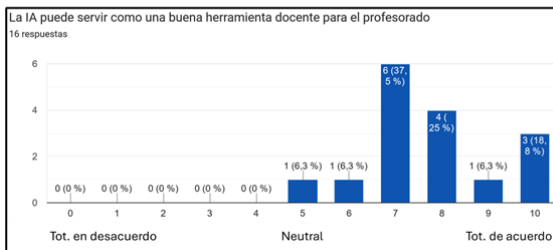
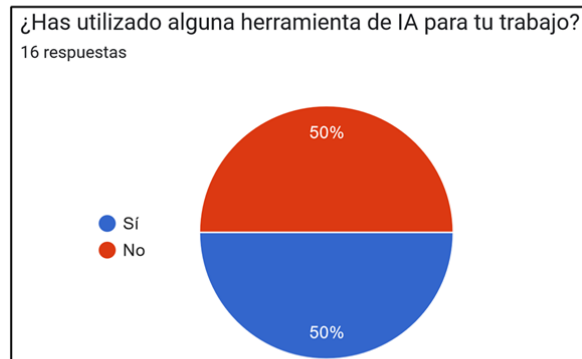
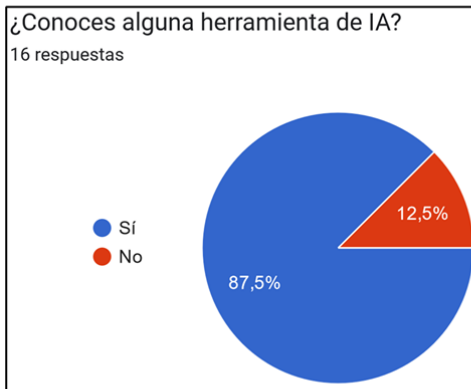
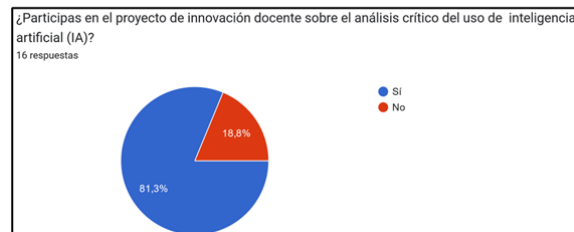
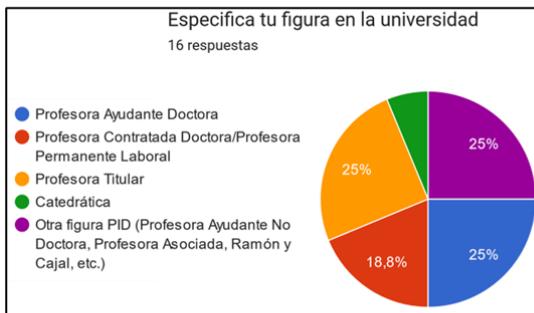
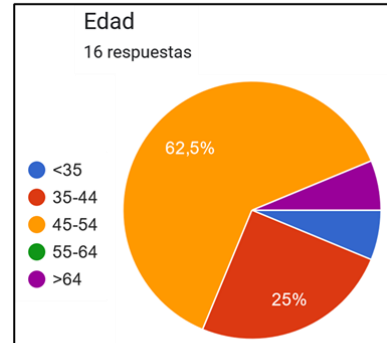
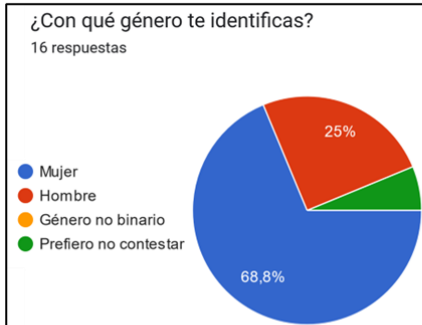
Encuesta final IA - Alumnas



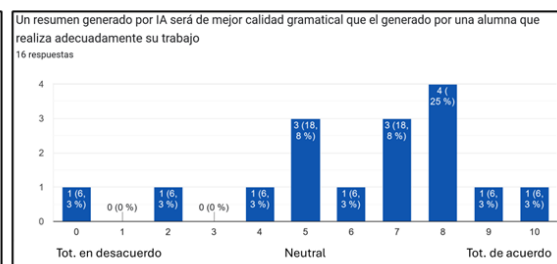
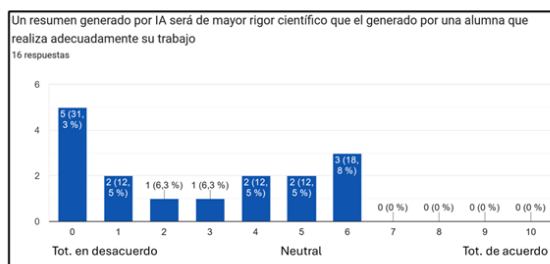
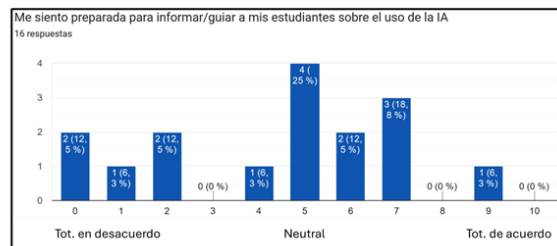
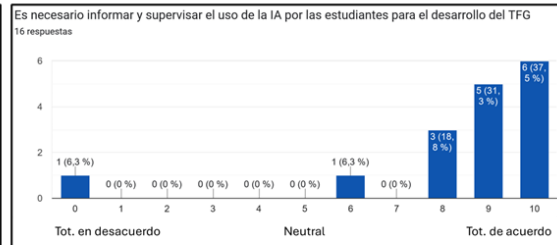
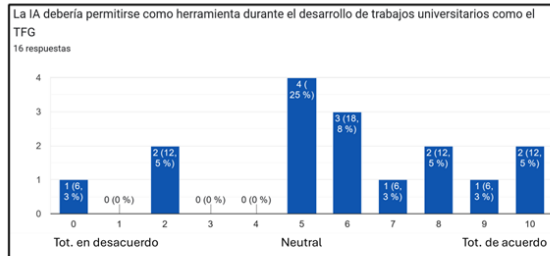
Encuesta final IA – Alumnas (cont.)



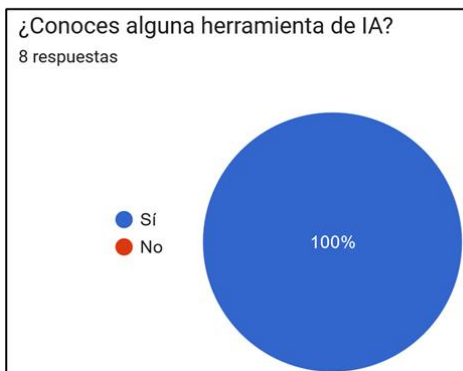
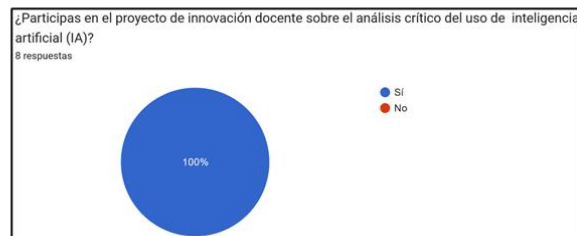
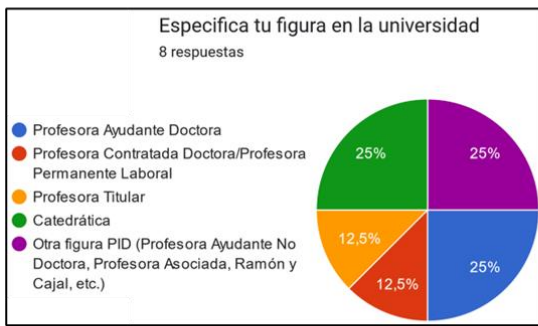
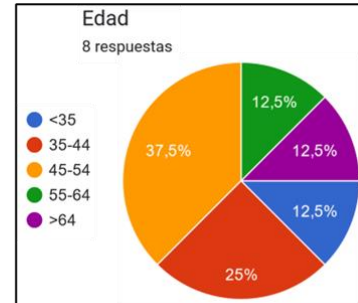
Encuesta inicial IA – Profesoras



Encuesta inicial IA – Profesoras (cont.)



Encuesta final IA – Profesoras



Encuesta final IA – Profesoras (cont.)

