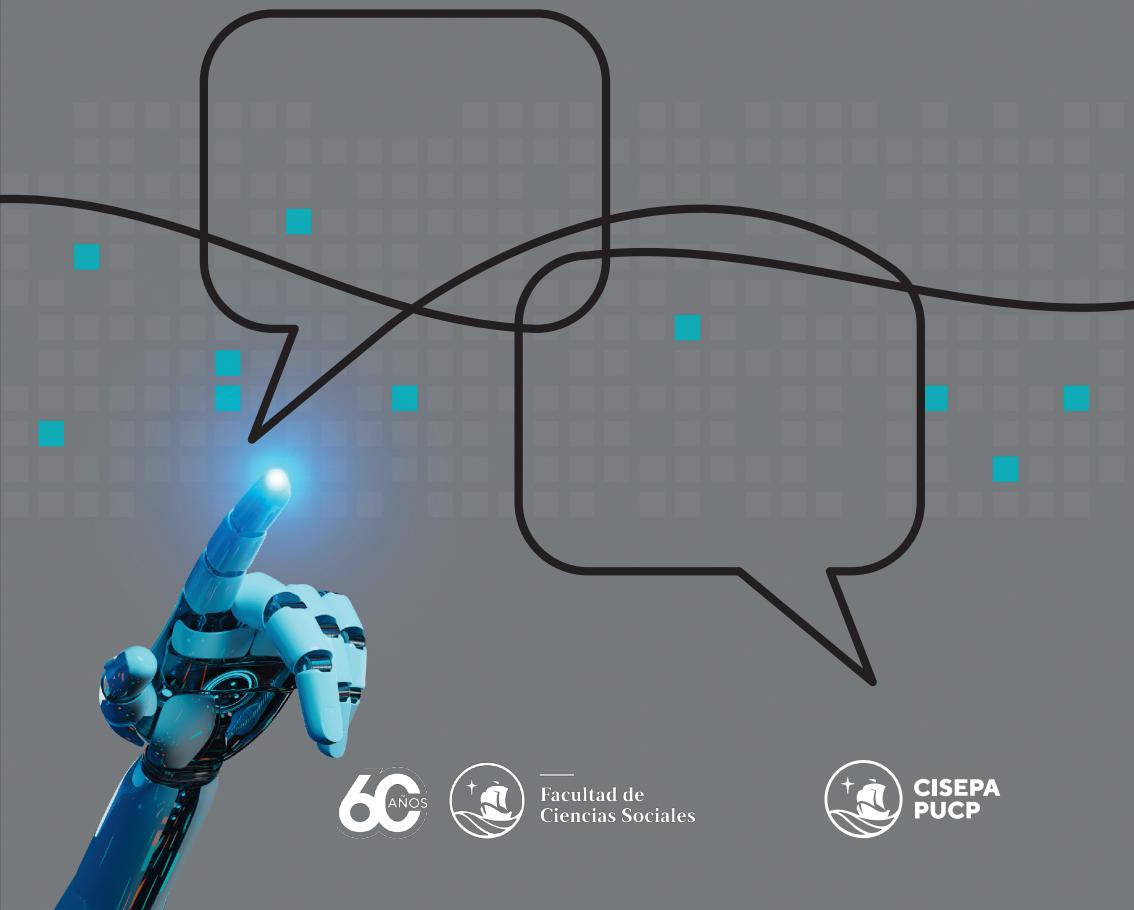


INTRODUCCIÓN AL ANÁLISIS DE DATOS CUALITATIVOS CON INTELIGENCIA ARTIFICIAL

Guía práctica para usar ChatGPT en
la investigación social y educativa

Manuel Etesse



Facultad de
Ciencias Sociales



CISEPA
PUCP

**INTRODUCCIÓN AL ANÁLISIS DE DATOS
CUALITATIVOS CON INTELIGENCIA ARTIFICIAL**
**Guía práctica para usar ChatGPT
en la investigación social y educativa**

Manuel Etesse



Facultad de
Ciencias Sociales



**CISEPA
PUCP**

Introducción al análisis de datos cualitativos con inteligencia artificial: Guía práctica para usar ChatGPT en la investigación social y educativa

Autor

Manuel Etesse, 2024 (manuel.etesse@pucp.edu.pe)

Pontificia Universidad Católica del Perú

Facultad de Ciencias Sociales

<https://facultad.pucp.edu.pe/ciencias-sociales/>

Centro de Investigaciones Sociológicas, Económicas, Políticas y Antropológicas, Cisepa

<https://cisepa.pucp.edu.pe/>

Av. Universitaria 1801, Pueblo Libre, Lima 15088, Perú



Licencia: Esta obra está bajo una licencia de Creative Commons Reconocimientos. No Comercial. Compartir Igual 4.0 Internacional (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Se debe reconocer explícitamente la autoría de los contenidos, no se puede utilizar el material para una finalidad comercial, y si se mezcla, transforma o crea a partir del material, no se puede difundir el material modificado.

Ilustraciones: el conjunto de figuras, recuadros y tablas, diagramas y recuadros son elaboración propia del autor, a menos que se indique lo contrario. Se incluyen capturas de pantalla provenientes de ChatGPT 3.5.

Género y lenguaje: se buscaron términos neutros en relación al género masculino, el uso del género masculino hace referencia a todas las personas de una población independientemente de sus características individuales.

Posibles errores: de corresponder, el señalamiento de errores se realizará a través de la página web del autor (<https://sites.google.com/pucp.edu.pe/etesse/inicio>)

Corrección de texto: Eleana Llosa

Primera edición digital: junio de 2024

Hecho el depósito legal en la Biblioteca Nacional del Perú N° 2024-05084

e-ISBN: 978-612-4355-17-2

DOI: [HTTPS://DOI.ORG/10.18800/978-612-4355-17-2](https://doi.org/10.18800/978-612-4355-17-2)

ÍNDICE GENERAL

Prólogo	11
Introducción	15
¿Inteligencia artificial para potenciar el análisis cualitativo?	
Capítulo 1	
La estrategia Aexco y el uso ético de ChatGPT	
1.1. Apuesta metodológica	21
1.2. Uso ético de la inteligencia artificial	25
Capítulo 2	
Preparar el análisis: obtención de una asistencia útil para mi proyecto	
2.1. Primeros pasos: entrar a ChatGPT	31
2.2. El entrenamiento CORI-F	33
Capítulo 3	
Explorar mis datos paso a paso con ChatGPT	
3.1. Primeros pasos: orientaciones básicas	41
3.2. Técnica 1. Selección de citas llamativas	49
3.3. Técnica 2. Síntesis casuística	53
3.4. Técnica 3. Recurrencia de términos	54
3.5. Técnica 4. Constatación de inconsistencias	55
3.6. Evaluación de los resultados	58

Capítulo 4

Codificar contenidos paso a paso con ayuda de ChatGPT

4.1. Primeros pasos: orientaciones básicas	61
4.2 Técnica 1. Codificación predefinida	71
4.3 Técnica 2. Codificación emergente	74
4.4 Evaluación crítica de los resultados	77

Cierre	81
--------	----

Referencias bibliográficas	83
----------------------------	----

Siglas	87
--------	----

Índice de consejos

Consejo 1. Evalúa el entrenamiento por medio de preguntas	37
Consejo 2. ¡Usa los pulgares!	39
Consejo 3. Recuerda frecuentemente el OE	52
Consejo 4. Usa <i>regenerate</i>	57
Consejo 5. Haz una retroalimentación a ChatGPT	60
Consejo 6. El uso de formatos ordenados para la codificación	76
Consejo 7. Haz una retroalimentación a ChatGPT	79

Índice de figuras

Figura 1.1. Ilustración de la secuencia inicial de la estrategia Aexco para el análisis de datos cualitativos	24
Figura 1.2. Diagrama de flujo para el uso seguro de ChatGPT	27
Figura 1.3. Circuito cerrado de investigación	28
Figura 2.1. Página principal de ChatGPT	31
Figura 2.2. Interfaz inicial de ChatGPT	32
Figura 2.3. <i>Prompt</i> del contexto del trabajo	35

Figura 2.4. <i>Prompt</i> del OE del trabajo de investigación	35
Figura 2.5. <i>Prompt</i> del rol en el trabajo de investigación	36
Figura 2.6. <i>Prompt</i> del uso de la información proporcionada a ChatGPT	36
Figura 2.7. Confirmación de los <i>prompts</i> del contexto y objetivos	37
Figura 2.8. Confirmación de los <i>prompts</i> del rol y el uso de la información	38
Figura 2.9. Resultados de preguntas acerca del proyecto de investigación	38
Figura 2.10 Ubicación de los pulgares de retroalimentación en la interfaz de ChatGPT	39
Figura 3.1. Secuencia de preguntas motivadoras para la exploración inicial de datos cualitativos	42
Figura 3.2. Entrenamiento CORI-F para la exploración	49
Figura 3.3. <i>Prompt</i> para realizar la selección de citas relevantes	50
Figura 3.4. Resultados de la selección de citas relevantes en ChatGPT	51
Figura 3.5. Señalamiento del OE en la indicación (<i>prompt</i>) del procedimiento de selección de citas	52
Figura 3.6. <i>Prompt</i> para realizar la síntesis casuística en ChatGPT	53
Figura 3.7. Resultados de la síntesis del caso por parte de ChatGPT	53
Figura 3.8. <i>Prompt</i> para realizar un recuento de las palabras más frecuentes en ChatGPT	54
Figura 3.9. Resultados del recuento de términos por parte de ChatGPT	55
Figura 3.10. <i>Prompt</i> para identificar inconsistencias en ChatGPT	56
Figura 3.11. Resultados de la identificación de inconsistencias por parte de ChatGPT	56
Figura 3.12. Opción <i>regenerate</i> en ChatGPT	57
Figura 4.1. Entrenamiento CORI-F para la codificación	72

Figura 4.2. <i>Prompt</i> para la creación de códigos predefinidos en ChatGPT	73
Figura 4.3. Resultados de la codificación predefinida en ChatGPT	73
Figura 4.4. <i>Prompt</i> para realizar la codificación emergente	74
Figura 4.5. Resultados de la codificación emergente en ChatGPT	75
Figura 4.6. Indicar formatos para la codificación inicial en ChatGPT	76

Índice de recuadros

Recuadro 3.1. Caso ilustrativo: Discurso presidencial de un expresidente	45
Recuadro 4.1. Guía de entrevista del proyecto «Trayectorias laborales y educativas de personas con discapacidad»	66
Recuadro 4.2. Fragmento de entrevista a Celia, proyecto «Trayectorias laborales y educativas de personas con discapacidad»	67

Índice de tablas

Tabla 3.1. Términos del proyecto de investigación «Discursos presidenciales»	45
Tabla 3.2. Evaluación del desempeño de ChatGPT en la exploración de datos cualitativos	59
Tabla 4.1. Términos del proyecto de investigación «Trayectorias laborales y educativas de personas con discapacidad»	65
Tabla 4.2. Fortalezas y debilidades identificadas en el proceso de codificación inicial con ChatGPT	78

PROLOGO

La inteligencia artificial ya ha llegado, pero aún queda mucho camino por recorrer.

Mustafa Suleyman y Michael Bhaskar¹

La presente guía práctica para el uso de ChatGPT² en la investigación social y educativa es resultado de una iniciativa del profesor Manuel Etesse, docente de la Facultad de Educación de la Pontificia Universidad Católica del Perú (PUCP), quien desde 2016 cuenta con una vasta trayectoria en la enseñanza de cursos y talleres vinculados al procesamiento y análisis de datos impartidos en la Facultad de Ciencias Sociales. En los últimos años, Etesse ha innovado en el campo de la inteligencia artificial (IA), en base a lo cual propuso el «Taller de introducción al análisis social de datos cualitativos usando inteligencia artificial». A partir de dicha experiencia, elaboró este manual, que ofrece una explicación sobre el sentido y el amplio alcance de ChatGPT, aplicación desarrollada por Open AI en 2022, la cual, a través de sistemas informáticos, procesa información y data requerida en distintos campos: salud, comercio, educación, comunicaciones, investigación, planeamiento, traducción, artes y creación de imágenes, entre otros. En el caso

1 Ver *Tecnología, poder y el gran dilema del siglo XXI. La ola que viene* (2023, p. 73; Penguin Random House).

2 Chat Generative Pre-Trained Transformer; es un *chatbot*, es decir, un *bot* conversacional, muy potente.

de ChatGPT, como en el de toda herramienta, es importante conocer las potencialidades de su uso, así como sus limitaciones, lo cual se muestra en esta publicación.

Como señala el autor, la IA tiene un enorme potencial para la investigación social. En el caso del análisis de datos cualitativos, por ejemplo, ChatGPT abre muchas posibilidades en la investigación social y educativa, considerando la riqueza de datos y contenidos que proceden de entrevistas, normas, grupos focales y documentación que se requiera analizar. Para mostrarlo, Etesse presenta de manera didáctica y práctica ejemplos de investigaciones realizadas por él mismo para procesar y analizar información primaria. Tales ejemplos ayudan a comprender y aplicar las técnicas de investigación que se pueden emplear con ChatGPT.

No obstante, también hay que advertir el uso ético de este tipo de programas y estar en permanente vigilancia epistemológica, como señalaba Pierre Bourdieu³. No se trata de que el investigador o la investigadora delegue a la aplicación los resultados de su estudio, porque ChatGPT no reemplaza el trabajo que realiza el profesional, sino que lo complementa, acercándose, quizás, al tipo de colaboración que se puede tener de un o una asistente de investigación.

El o la investigadora requiere realizar previamente una selección inicial de datos resaltantes y debe poseer ideas sobre cómo organizar la información, así como haber llevado a cabo un potencial análisis. Así podrá saber que la información arrojada por

3 «La práctica de la sociología reflexiva» (en Bourdieu, P. y Wacquant, L., *Una invitación a la sociología reflexiva*; 2005, pp. 266-317; Siglo XXI).

ChatGPT es verosímil, con base en los datos de investigación. Así, pues, resalta la centralidad del investigador en el proceso.

De acuerdo a Noam Chomsky⁴, este tipo de aplicaciones no son propiamente «inteligentes», al menos como lo son los seres humanos. A diferencia de lo que ellas hacen, los humanos podemos almacenar pequeñas cantidades de información a través de la cual no solo se busca generar correlaciones entre puntos, sino más bien y sobre todo crear posibles explicaciones. En otras palabras, dichas aplicaciones no poseen la capacidad que tiene como especie el ser humano de razonar y generar reflexión crítica.

Mediante la publicación de este manual, desde la Facultad de Ciencias Sociales, así como desde el Centro de Investigaciones Sociológicas, Económicas, Políticas y Antropológicas (Cisepa), buscamos contribuir a que tanto los y las estudiantes de la facultad como otras personas interesadas puedan conocer la utilidad de la aplicación ChatGPT y los nuevos modelos conversacionales. Asimismo, esperamos aportar a la reflexión y el debate sobre la implicancia de esta tecnología en la transformación de todos los niveles de la vida, incluida la labor de investigación.

Fanni Muñoz Cabrejo

Decana

Facultad de Ciencias Sociales

Pontificia Universidad Católica
del Perú

Paola Patiño Rabines

Directora ejecutiva

Centro de Investigaciones

Sociológicas, Económicas,
Políticas y Antropológicas

Pontificia Universidad Católica
del Perú

4 «The false promise of ChatGPT» (en *The New York Times*, 8 de marzo de 2023; <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>).

INTRODUCCIÓN

¿Inteligencia artificial para potenciar el análisis cualitativo?

Desde el lanzamiento de ChatGPT a finales del año 2022, el avance de la inteligencia artificial (IA) generativa está transformando la vida y el trabajo a nivel global. En la esfera de la creación científica y humanística, este fenómeno toca todos los ámbitos del quehacer: desde la ciencia matemática hasta la creación en artes plásticas. Su desarrollo es exponencial, por lo que hoy se hace difícil imaginar el futuro sin IA.

ChatGPT es una aplicación que pertenece a la organización sin fines de lucro OpenAI Incorporated y a la corporación OpenAI Limited Partnership, empresa con sede en el estado de California, en los Estados Unidos (OpenAI, 2023). Esta aplicación rompió los esquemas debido a que cuenta con una interfaz tan simple y amigable como lo es una ventana virtual de conversación, un *chat*. Quizás por ello, cinco días después de su lanzamiento ya contaba con más de un millón de usuarios inscritos (Firat, 2023). Lo que más atrae al público hacia ChatGPT se explica por el significado de sus siglas: *generative pretrained transformer*, es decir, es un «transformador generativo preentrenado» basado en un modelo de lenguaje natural a gran escala (LLM) con capacidad de sostener una con-

versación, o «chatear», con una persona y brindar sorprendentes grados de coherencia y relevancia en sus respuestas. Asimismo, cuenta con tres valiosas capacidades: recuerda el contenido de la conversación, provee información a partir de grandes fuentes de información y tiene el potencial de transformar textos. Sin embargo, como veremos más adelante, algunas de estas capacidades presentan limitaciones que se deben tratar con sumo cuidado y entendimiento.

La IA tiene un gran potencial en la investigación social al ofrecer la capacidad de analizar grandes cantidades de datos de manera eficiente y de identificar patrones complejos. La intersección entre la innovación tecnológica y la investigación plantea oportunidades y desafíos para el análisis empírico de datos (Vargas-Bianchi, 2023). Así, en particular, resulta útil el uso de ChatGPT en el análisis de datos cualitativos en la investigación social y educativa, datos caracterizados por la riqueza de los contenidos provenientes de entrevistas, discursos, normativas, materiales didácticos y grupos focales, entre muchos otros (Lopezosa y Codina, 2023; Lopezosa et al., 2023). Se ha podido ver que, a partir de experiencias prácticas, ChatGPT puede ser una herramienta valiosa especialmente en el análisis de documentos y entrevistas cuando se le asigna un rol pertinente y se proporcionan instrucciones adecuadas (Anis y French, 2023; Rahman et al., 2023). No obstante, es esencial que el investigador lleve una constante supervisión o «auditoría» para garantizar la pertinencia y validez de los resultados.

Los contenidos que aquí se presentan son resultado de diversas rondas de revisiones de la literatura naciente sobre el tema. En paralelo, provienen de experiencias concretas de investigación y enseñanza, cuyo rasgo principal ha sido la experimentación constante y la reflexión crítica sobre los alcances de cada procedimiento. Por

ello, esta guía propone, para cada técnica presentada, una evaluación precisa de las potencialidades y limitaciones identificadas.

El objetivo de este trabajo es presentar un marco claro y útil para el uso de la IA generativa en el estudio de contenidos cualitativos, así como detallar, paso a paso, procedimientos específicos que pueden contribuir a su aplicación en una gran diversidad de temáticas específicas. En este sentido, la presente propuesta tiene vocación multidisciplinaria, dado que busca ser útil para un amplio espectro de disciplinas, epistemologías y metodologías en la práctica del análisis de datos cualitativos.

En particular, esta propuesta de análisis asistido por ChatGPT presenta seis técnicas eficaces y creativas para las fases de exploración y codificación en el marco de un diseño metodológico elaborado por personas (no por IA).

El descubrimiento y la clasificación de los contenidos objeto de análisis pueden beneficiarse mucho de insumos producidos por la IA con el objetivo de enriquecer las nacientes perspectivas del investigador —no para reemplazarlas o crearlas— acerca de contenidos que él ya conoce. Estos insumos, como aquellos que produce un asistente de investigación humano, pueden servir para hallar nuevas pistas analíticas, apoyar la identificación de patrones y brindar elementos para fortalecer la abstracción del investigador responsable. Es importante resaltar que tales insumos han de ser principalmente un complemento a la mirada propia del investigador —a través de la lectura y el análisis directo de contenidos— y no un sustituto de ella.

Fue en las aulas de la Facultad de Ciencias Sociales de la Pontificia Universidad Católica del Perú (PUCP) y en la Escuela de In-

vestigadores del Centro de Investigaciones Sociológicas, Económicas, Políticas y Antropológicas (Cisepa) donde esta propuesta tomó forma. Por ello, agradezco a la Dirección de Estudios por posibilitarme impartir el curso interdisciplinario Taller de Herramientas Cualitativas para el Análisis Social en el verano de 2023, en el que se intercambiaron las primeras reflexiones sobre ChatGPT aplicado a la investigación, tras su reciente lanzamiento a finales de 2022. Asimismo, quiero agradecer al Cisepa, que poco después acogió mi propuesta de dictar un Taller Introductorio al Análisis Social de Datos Cualitativos usando Inteligencia Artificial y por el apoyo recibido durante cada reedición de este curso de extensión.

Los procedimientos que aquí se presentan se beneficiaron de los valiosos aportes y reflexiones de los estudiantes¹ de pregrado y educación continua en los cursos mencionados, así como en el Taller de Procesamiento de Datos de la carrera de Sociología en el primer semestre de 2023. En particular, cabe destacar que la redacción de esta guía recibió la valiosa asistencia del estudiante Piero Beretta Vidal. Asimismo, se añade un especial agradecimiento a la Facultad de Ciencias Sociales, en particular a la decana doctora Fanni Muñoz, por el apoyo recibido para la publicación de mi manuscrito.

Por último, corresponde destacar que la presente guía práctica está compuesta de cuatro capítulos y una sección de reflexiones y consejos finales. El primer capítulo presenta la estrategia metodológica Aexco, en el marco de la cual se desarrollan los procedimientos de esta guía, así como los planteamientos éticos esen-

1 En la redacción de este documento, se buscaron términos neutros e inclusivos, no obstante, cuando fue necesario, como aquí, se usó el género masculino para hacer referencia a todas las personas de una población, independientemente de sus características individuales.

ciales para el uso de ChatGPT en el análisis de datos. El segundo desarrolla los primeros pasos en el uso del programa ChatGPT, así como el importante entrenamiento CORI-F (contexto, rol, información y formato). El tercer capítulo brinda una presentación sobre la exploración de datos cualitativos y, posteriormente, guía al lector en la aplicación de dos técnicas básicas, paso a paso. Asimismo, provee un análisis de las fortalezas y debilidades identificadas en cada técnica. El cuarto capítulo introduce los principios de la codificación como proceso analítico, y, más adelante, propone al lector dos técnicas básicas para iniciar la clasificación del contenido cualitativo. Al igual que el capítulo anterior, brinda un análisis de las fortalezas y debilidades identificadas en ambas técnicas. Finalmente, como se ha anotado, se presenta un conjunto de reflexiones y consejos finales para ser más creativos, eficientes y responsables en el uso de la IA para el análisis de datos cualitativos.

CAPÍTULO 1

La estrategia Aexco y el uso ético de ChatGPT

1.1. Apuesta metodológica

La apuesta metodológica de esta guía se inspira libremente en la teoría enraizada o teoría fundamentada (del inglés, *Grounded theory*) a través de autores como Barney Glaser y Anselm Strauss (2006) y Kathy Charmaz (2014). Sin embargo, también se basa en la práctica heterodoxa de análisis de datos cualitativos y en la búsqueda de perspectivas de utilidad multidisciplinaria (Elder y Giele, 2009; Creswell y Poth, 2016; Etesse, 2019). De hecho, se ha nutrido de la práctica investigativa y la enseñanza de metodología en disciplinas como Educación, Antropología, Sociología y Ciencias Políticas.

En este sentido, cuatro elementos esenciales dan forma a la apuesta metodológica de esta guía. En primer lugar, un interés particular por contenidos textuales no estructurados, es decir, que no poseen una estructura estandarizada (como sí existe en los datos cuantitativos), como transcripciones de entrevistas, grupos focales, canciones, discursos y sesiones de clase. En segundo lugar, se busca fomentar el descubrimiento de elementos nuevos, que

difícilmente pueden ser advertidos durante el trabajo de campo o tras una relectura inicial de los datos. En tercer lugar, se quiere responder a preguntas específicas de investigación por medio de conceptualizaciones, por lo que se fomenta la abstracción y la comparación. Finalmente, se establecen vínculos claros entre el resultado del análisis y los datos primarios por medio de procedimientos que aseguran su trazabilidad y transparencia.

De esta apuesta metodológica surge la estrategia analítica llamada Aexco. Antes de profundizar en ella, es importante considerar la fase de diseño de investigación. En esta etapa, las preguntas específicas definen la orientación, responsabilidad que recae en el investigador. La formulación de tales preguntas, más detalladas y precisas que la pregunta general, requiere una atención minuciosa, ya que el investigador debe ser plenamente consciente de sus implicancias, dado que esta fase de diseño es fundamental al orientar las decisiones metodológicas a lo largo del estudio. En este sentido, se sugiere que cada pregunta específica de investigación se sitúe con claridad en tres dimensiones clave (Gibbs, 2008; Creswell y Poth, 2016): el tipo de razonamiento predominante (inductivo o deductivo), el tipo de evidencia predominante a analizar (objetiva o subjetiva) y el enfoque principal de descubrimiento (nomotético o idiográfico). Con preguntas específicas de investigación claras y precisas, el investigador podrá tomar decisiones metodológicas coherentes, efectivas y particulares en las fases de preparación, recolección y generación de datos; registro y transcripción de información; así como en el análisis e interpretación de los datos.

La estrategia Aexco propone generar descubrimiento y análisis útiles a través de tres procesos: la anotación, la exploración y la codificación. Cabe recordar que la efectividad de esta estrategia se basa en su orientación a responder preguntas específicas de

investigación planteadas con claridad y precisión. Esta estrategia proporciona una orientación práctica sobre la organización e interpretación de datos cualitativos, lo que resulta beneficioso para investigaciones que adopten enfoques diversos.

En cuanto al primer proceso, la anotación, desempeña un papel esencial en el transcurso de la investigación al permitir a los investigadores registrar sus impresiones, reflexiones y decisiones en relación con los datos recolectados. Este proceso se erige como una práctica que facilita la exploración de datos, la identificación de patrones y tendencias emergentes, y la movilización de conceptos sólidos (Birks et al., 2008). Además, es un recordatorio constante de los objetivos de la investigación, lo que ayuda a mantener a los investigadores en el camino correcto a medida que avanzan en su estudio, así como a generar sinergias entre las diferentes etapas del trabajo de investigación.

El segundo proceso, la exploración, es la práctica crucial para reconocer las características específicas al interior de los casos, y entre ellos, con la finalidad de identificar patrones y tendencias a partir de los elementos presentes en el material de estudio (Flick, 2014). Este proceso involucra una relectura minuciosa y una familiarización profunda con los contenidos, lo que implica no solo comprender su estructura, sino también evaluar su calidad y destacar las particularidades significativas. Durante esta fase, se inicia la búsqueda de contenidos y se trazan pistas analíticas que pueden resultar especialmente interesantes para responder a las preguntas específicas de investigación. Es así, principalmente, porque Aexco promueve nuevas formas de aproximarnos a nuestros datos, con apertura al descubrimiento, lo que allana el camino para una exploración más detallada y un análisis en profundidad de los datos recopilados.

Por último, la codificación es un proceso de clasificación de segmentos de contenidos bajo uno o más códigos. Dicho de otra manera, este proceso implica seleccionar contenidos –detallados y situacionales– y asignarles conceptos o palabras más generales. Los códigos pueden predefinirse o surgir durante el proceso, según sea la orientación prioritaria de la pregunta específica de investigación. La codificación suele realizarse en dos grandes ciclos (Saldaña, 2021): en el primero, se categorizan los datos en uno o más códigos; y en el segundo, se realiza una reorganización de los códigos con el fin de establecer ejes analíticos encaminados a una proposición final. En esta guía trataremos únicamente sobre el primer ciclo de codificación.

La estrategia Aexco tiene un carácter tanto sistemático como flexible, ya que se basa en una práctica iterativa y concomitante. Así, como se aprecia en la figura 1.1, el mayor potencial analítico se alcanza a través de secuencias de continua retroalimentación

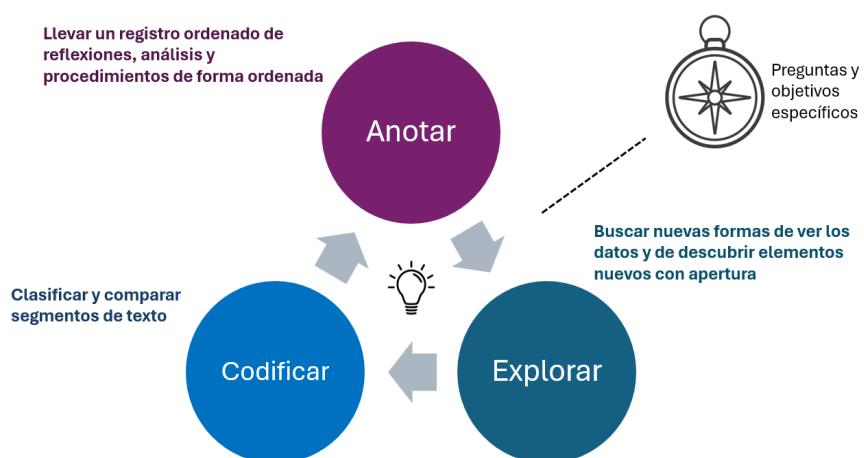


Figura 1.1. Ilustración de la secuencia inicial de la estrategia Aexco para el análisis de datos cualitativos.

y retrospección entre anotación, exploración y codificación. Cabe señalar que esta guía se enfoca precisamente en los procesos de exploración y codificación de los datos cualitativos.

1.2. Uso ético de la inteligencia artificial

Es importante destacar que el rol y el alcance de la IA en la investigación social y educativa debe contemplar tales marcos regulatorios institucionales y de carácter ético para promover un uso adecuado, pertinente y seguro de dicha herramienta. A pesar de que tales marcos se encuentran en pleno desarrollo, cabe destacar los lineamientos del *Libro blanco sobre la IA: un enfoque europeo orientado a la excelencia y la confianza*, publicado por la Comisión Europea (2020), en donde se hace hincapié en la ética, la privacidad y la seguridad en el uso de algoritmos para la toma de decisiones y la investigación. Asimismo, se debe considerar la *Recomendación sobre la ética de la inteligencia artificial* de la Unesco (2021), que promueve una reflexión normativa sistemática sobre el surgimiento y la aplicación de la IA, reflexión basada en un marco integral, global, multicultural y evolutivo de valores, principios y acciones. Esta reflexión debe guiar el desarrollo y uso responsable de esta tecnología bajo principios que garanticen los derechos humanos, la privacidad, la protección de datos, la transparencia y la auditabilidad. Así como la atención a estos lineamientos y consideraciones elaborados por parte de entes internacionales, cabe señalar que el uso dado a las herramientas de IA debe realizarse a través de los estándares éticos de cada disciplina y del entorno académico y laboral específico en el que se desenvuelve el investigador.

Dentro de los elementos principales en la discusión ética sobre el uso de la IA en la investigación, emerge la potencial autoría de la IA (Xames y Shefa, 2023). Sobre este punto, es importante establecer

que, si bien ChatGPT se usa cada vez más en distintos procesos de la investigación, no ostenta estatus legal, por lo que tal programa no puede poseer derechos de autor ni ser sujeto de demanda ante posibles incidencias derivadas de la investigación (Committee on Publication Ethics, COPE, 2020). En particular, resulta clave indicar que ChatGPT no es un autor, y que la autoría solo puede recaer en el investigador, que asume toda la responsabilidad del tratamiento y de los resultados finales de la investigación. A la fecha, algunos editores, como es el caso de Springer-Nature, Elsevier, y Taylor & Francis, han fijado lineamientos para la inclusión del uso de la IA en la sección de metodología o sobre uso de herramientas (Rahman et al., 2023). En la misma línea, esta guía sugiere el señalamiento del uso de IA generativa en la sección de métodos al calificarse como una herramienta, al igual que otros programas tradicionales de análisis de datos.

Por otra parte, aquí se considera fundamental que el investigador adopte una postura clara en cuanto al rol de la IA en el proyecto de investigación. En particular, se asigna a ChatGPT el rol de herramienta de asistencia al análisis de datos cualitativos. La finalidad de esta herramienta será transformar información y producir instrumentos para enriquecer la mirada del investigador sobre los datos. En este sentido, resulta indispensable poseer un buen conocimiento previo de los contenidos a analizar con la ayuda de la IA.

Teniendo en consideración estos elementos, la figura 1.2 presenta un diagrama de flujo para tomar la decisión de usar ChatGPT. Esta decisión implica estar apto para seguir el camino trazado con línea verde: primero, debe importar que el resultado sea verdadero; segundo, se requiere disponer de experiencia suficiente para verificar que el resultado sea preciso; y, tercero, es necesario que el investigador esté dispuesto a asumir la completa responsabilidad por los resultados finales. Si él no está en condiciones de asumir

todas y cada una de estas tres condiciones, más vale no usar dicha herramienta para el análisis de datos sociales y educativos.

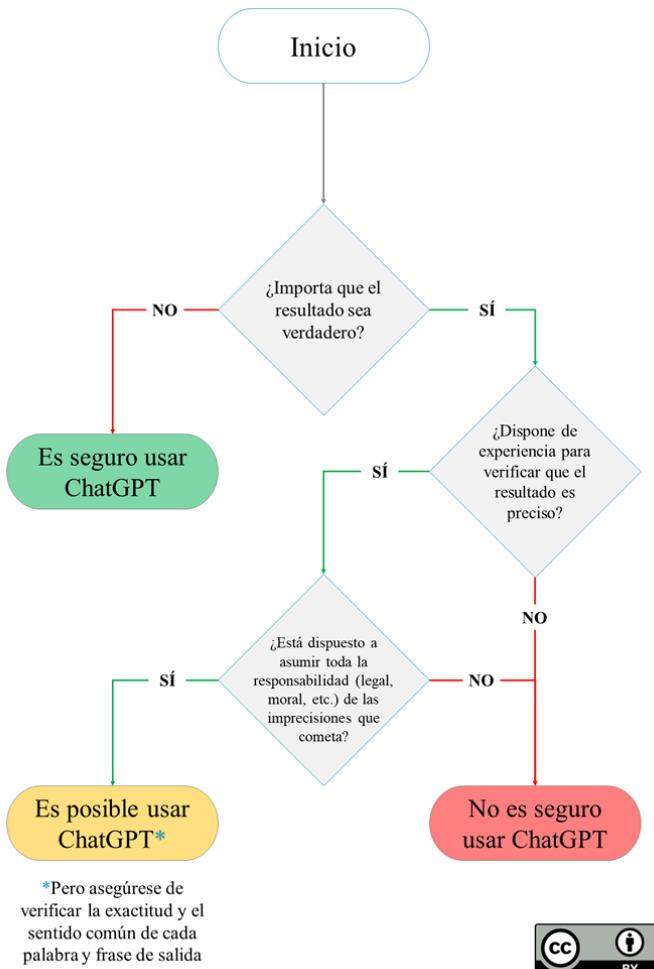


Figura 1.2. Diagrama de flujo para el uso seguro de ChatGPT.

Fuente: Sabzalieva y Valentini (2023; diseño de Aleksandr Tiulkov; rediseño de Fraidy Alonso Alzate Pamplona; bajo licencia de Atribución CC BY; fuente: AI and Data Policy Lawyer).



Como se muestra en la figura 1.3., el trabajo de investigación debe desarrollarse bajo la idea de un circuito cerrado de información. Puesto que ChatGPT no será usado para acceder a información desconocida, es importante establecer que utilice únicamente la que se le proveerá. Ello cobra una relevancia mayor ante un uso cada vez más extendido de la IA generativa como motor de búsqueda de información que se desconoce. Esta guía no sugiere usar ChatGPT como motor de búsqueda, tampoco para consultar información teórica, metodológica o de otro tipo. Las preguntas o solicitudes que se hagan a la herramienta comprometerán única y exclusivamente los datos de la investigación que el investigador ya conoce bien. Así, el trabajo se centrará en que la herramienta transforme información que ya él ya conoce. En este sentido, el circuito cerrado de información permitirá crear una actitud específica en el investigador con respecto a la herramienta.

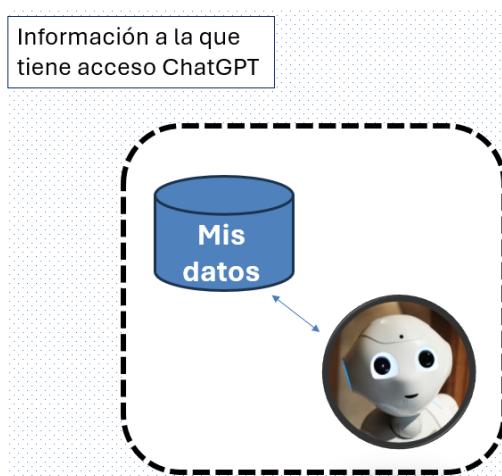


Figura 1.3. Circuito cerrado de investigación.

Dentro del ámbito de la investigación social y educativa, es de suma importancia establecer protocolos de manejo seguro de la información de personas, grupos y organizaciones. A la fecha no se tiene un conocimiento certero acerca del uso que ChatGPT da a los datos que le proveen los usuarios, por lo que resulta imprescindible brindar datos anónimos. Así, toda la información sometida al análisis de IA debe ser previamente sometida a un exhaustivo proceso de anonimización. Este procedimiento busca salvaguardar la identidad de los participantes y su confidencialidad, un pilar ético fundamental para garantizar la integridad de la investigación (PUCP, 2016). Cabe considerar que la cuidadosa gestión de la confidencialidad contribuye a respetar los derechos y la privacidad de los involucrados en un estudio, y en muchos casos fortalece la validez de los resultados obtenidos.

Se debe tener en cuenta que los resultados que da ChatGPT están lejos de ser siempre neutros y pertinentes. Esta herramienta posee sesgos basados en la forma en que ha sido programada, particularmente aquellos que reflejan los que tiene el equipo que la ha desarrollado. Diversos estudios recientes han detectado en ChatGPT sus tendencias y posturas, las cuales no evidencian la diversidad cultural, social y política que existe en las sociedades y grupos de personas alrededor del mundo (Hartmann et al., 2023; Rutinowski et al., 2023). Asimismo, existen serias limitaciones en la herramienta para tomar en cuenta contextos e idiomas diferentes a aquellos en los que ha sido desarrollada y entrenada (Gosh y Caliskan, 2023). Debido a ello, resulta importante que el investigador establezca procedimientos minuciosos de supervisión, lo que incluye evaluar cada uno de sus resultados en términos de la pertinencia, la coherencia y el contexto específico que es estudiado.

CAPÍTULO 2

Preparar el análisis: obtención de una asistencia útil para mi proyecto

2.1. Primeros pasos: entrar a ChatGPT

A continuación, se presentan de forma ilustrada los dos primeros pasos para comenzar a trabajar con ChatGPT: el registro y la iniciación del trabajo en la interfaz.

Paso 1. Ingreso a la página web y creación de una cuenta

Ingresamos a la web oficial de ChatGPT² y creamos una cuenta gratuita (versión GPT 3.5). La figura 2.1 muestra la ventana de ini-

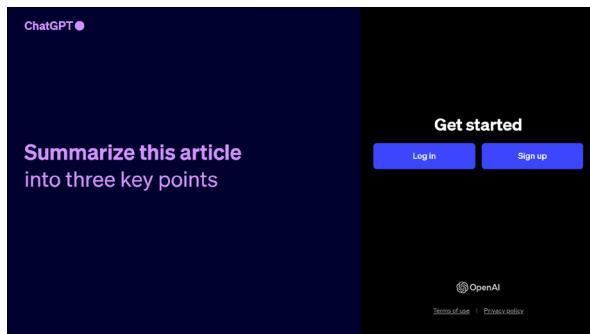


Figura 2.1. Página principal de ChatGPT.

2 Ver <https://chat.openai.com>

cio de ChatGPT. Para registrarnos, hacemos clic en *Sign up* y completamos nuestros datos.

Paso 2. Acceso a la interfaz

Tras la creación de la cuenta, accedemos a la interfaz de ChatGPT, donde encontraremos sus distintos componentes. Como nos muestra la figura 2.2 y los números en ella indicados, la interfaz se compone de:

1. La barra de indicaciones o *prompts* (en la parte inferior). Aquí escribimos los mensajes, indicaciones y solicitudes a la herramienta, que llamaremos de aquí en adelante *prompts*.
2. El historial de conversaciones o *chats* (en la franja lateral izquierda). Aquí se encontrarán repatriadas cronológicamente las conversaciones realizadas.
3. El botón para crear una nueva conversación o *chat* (en la franja lateral izquierda). Al hacer clic en este espacio se genera una nueva conversación.
4. El espacio de conversación (en el área central). En este espacio, se visualizan nuestros mensajes y las respuestas de ChatGPT.

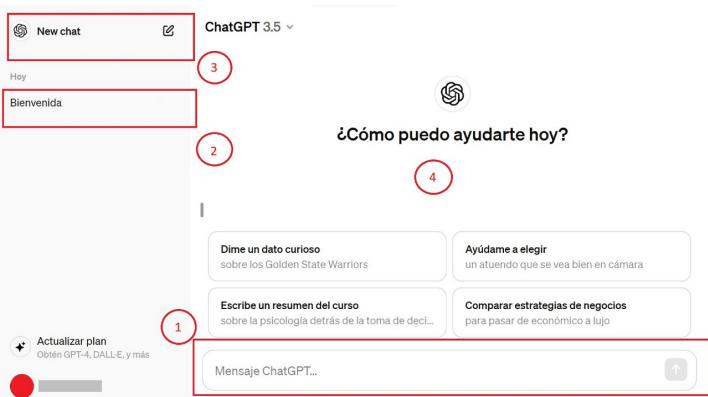


Figura 2.2. Interfaz inicial de ChatGPT.

2.2. El entrenamiento CORI-F

Resulta esencial configurar o «entrenar» a ChatGPT para mejorar su asistencia al análisis específico de nuestro proyecto de investigación particular. Con ello, la herramienta podrá realizar un análisis más valioso y preciso. Este entrenamiento consiste en establecer ciertos parámetros clave para el tipo de trabajo que se le solicitará. De esta manera, para que la herramienta pueda generar resultados pertinentes, se deben proporcionar instrucciones iniciales acerca de la naturaleza y finalidades del trabajo a realizar (Zhang et al., 2023). Esto es importante porque un entrenamiento bien hecho influye significativamente en las respuestas de ChatGPT, especialmente en lo que respecta a la exploración y la codificación de datos.

Esta guía propone el entrenamiento CORI-F como procedimiento esencial clave para una preparación adecuada de la herramienta. Tal entrenamiento resulta de informar a la herramienta acerca de: el Contexto (C) del trabajo, su Objetivo (O), la adopción de un Rol (R), la forma de tratamiento de la Información (I) y, adicionalmente, las características del Formato (F) de la respuesta que se solicita. Este enfoque se inspira en la metodología R.O.C.E.F. (Hoyl, 2023), pero no se agota en ella, ya que incorpora la idea de un trabajo en circuito cerrado. En este sentido, el entrenamiento CORI-F surge como un método efectivo y especializado para la utilización adecuada de ChatGPT en la investigación de datos cualitativos.

De esta manera, se establece un marco integral para guiar la interacción con ChatGPT al asegurar que se tenga claro el contexto de trabajo, los objetivos específicos (OE), el rol que debe adoptar la herramienta, el uso exclusivo de información conocida y el formato óptimo de salida. Llegados a este punto, cabe destacar que el

establecer un entrenamiento inicial no implica la imposibilidad de solicitudes de reajuste de las respuestas y/o recordatorios acerca de los elementos básicos del proyecto. De esta manera, se permite que esta interacción desemboque en un virtuoso ajuste de resultados con respecto a nuestras necesidades particulares de investigación.

Ejemplos de entrenamiento inicial

C - Contexto

La C en CORI-F se refiere al contexto, lo que implica proporcionar la fundamentación necesaria para abordar de manera pertinente el proyecto de investigación, esto es, que la herramienta contemple el marco de trabajo de los datos.

Al respecto, se establecen al menos cuatro elementos clave de información:

1. Tipo de trabajo a realizar

Ejemplos: investigación científica; análisis de información.

2. Tipo de datos a trabajar

En este caso, datos cualitativos. Puede ser: transcripciones de entrevistas, clases o grupos focales.

3. Disciplina o enfoque disciplinario de estudio

Ejemplos: sociología política; educación.

4. Ubicación geográfica y/o temporalidad

Ejemplos: Perú actual; Argentina en 1996.

La figura 2.3 muestra un ejemplo de este tipo de *prompt*.

El *prompt* se puede formular también de otras formas y con diferentes temas de investigación. Por ejemplo: «Me encuentro

P **Tú**

El objetivo específico por trabajar es identificar la forma en que se plantean los principales problemas y la crisis política en el contenido de los discursos.

○

Figura 2.3. *Prompt* del contexto del trabajo.

desarrollando una investigación universitaria en educación en el Perú. Para ello, tengo datos cualitativos, que son entrevistas a personas con discapacidad. Estos datos deben ser analizados para entender sus experiencias educativas y laborales».

O - Objetivos

La O de CORI-F se refiere a los objetivos, lo que implica especificar claramente los lineamientos que guían la investigación y, por ende, orientar el análisis con la dirección previamente establecida. Se recomienda trabajar con las preguntas específicas o los OE de investigación como punto de partida para los diferentes *prompts* que se realicen. La figura 2.4 presenta un ejemplo de *prompt* acerca del objetivo.

P **Tú**

El objetivo específico por trabajar es identificar la forma en que se plantean los principales problemas y la crisis política en el contenido de los discursos.

○

Figura 2.4. *Prompt* del OE del trabajo de investigación.

Otro ejemplo podría ser el siguiente: «El primer objetivo específico es analizar cómo las personas con discapacidad entrevistadas recuerdan su experiencia educativa a lo largo de las tres etapas: educación primaria, educación secundaria y educación superior».

R - Rol

La R implica delimitar bien el rol que adoptará ChatGPT, lo que implica definir el papel que desempeña tal herramienta en el proceso de investigación. Esto ayudará a esclarecer su función, que debe estar centrada en la generación de insumos y la asistencia al investigador principal. Un ejemplo de este tipo de *prompt* se ilustra en la figura 2.5.

Tú
En este proyecto te desempeñas como asistente de investigación experto en el análisis de datos cualitativos.

Figura 2.5. *Prompt* del rol en el trabajo de investigación.

Una alternativa al *prompt* puede ser la siguiente: «Para este proyecto, tu rol será el de un analista en datos cualitativos que brinda asistencia en la investigación».

I - Información

La I es la sección información, la cual permite delimitar el conjunto de información que será utilizada en el diálogo con ChatGPT. Ello es esencial para establecer el ya mencionado circuito cerrado de información y utilizar únicamente los datos proporcionados por el investigador. Como se ha visto en el primer capítulo, trabajar únicamente con los datos conocidos y anonimizados resulta esencial desde el punto de vista ético (ver sección 1.2). A continuación, un ejemplo de este tipo de *prompt* se ilustra en la figura 2.6.

Tú
Debes basar tus respuestas exclusivamente en la información que te proporcionaré.

Figura 2.6. *Prompt* del uso de la información proporcionada a ChatGPT.

Una alternativa al *prompt* puede ser: «Para el análisis de los datos, debes usar únicamente la información que te suministraré a lo largo de la investigación».

F - Formato de los *prompts*

La sección F de CORI-F hace referencia al formato de *prompts* y se centra en brindar instrucciones específicas sobre la presentación de las respuestas generadas por ChatGPT. Esto garantizará que los resultados se entreguen de la manera más idónea en función de las necesidades del investigador. El formato de *prompts* abarca el tipo de lenguaje deseado, la presentación de resultados y la organización de los mismos. Por ejemplo, se puede requerir un lenguaje formal o informal, ordenado en una lista, una tabla, un párrafo, en guiones o en otras formas.

En las siguientes secciones, se presentan ejemplos concretos que ilustran estos *prompts* sobre el formato para orientar de manera específica la generación de respuestas por parte de ChatGPT.

Consejo 1. Evalúa el entrenamiento por medio de preguntas

Es útil realizar preguntas de validación a ChatGPT para asegurarnos de que comprende claramente lo que se requiere. Su respuesta nos permitirá, eventualmente, especificar o completar nuestro *prompt*.

Por ejemplo, tras indicar el entrenamiento CORI-F, es útil formularle las siguientes preguntas incluidas en las figuras 2.7 y 2.8.



Tú

¿Comprendes el contexto de esta investigación y el objetivo específico a trabajar?



Figura 2.7. Confirmación de los *prompts* del contexto y objetivos.

P Tú

¿Comprendes cuál es tu rol en la investigación y cómo debes utilizar la información proporcionada?



Figura 2.8. Confirmación de los prompts del rol y el uso de la información.

Adicionalmente, una alternativa útil es realizar algunas preguntas sobre lo que ha retenido ChatGPT tras el entrenamiento. Para ello, la figura 2.9 presenta un ejemplo de pregunta general e incluye la respuesta de ChatGPT.

P Tú

Hasta este punto, ¿Qué has entendido del proyecto de investigación?

ChatGPT

Hasta ahora, he entendido que estás llevando a cabo un proyecto de investigación en sociología política centrado en el análisis de discursos de investidura de los últimos presidentes peruanos. El objetivo específico que me has compartido se centra en identificar la forma en que se plantean los principales problemas y la crisis política en estos discursos.

La metodología que estamos discutiendo implica un enfoque cualitativo, donde se espera examinar cómo los presidentes abordan y comunican los desafíos políticos a través de sus discursos formales. Se mencionaron áreas clave de análisis, como la identificación de problemas, el análisis de la retórica utilizada, la jerarquía de problemas, las soluciones propuestas, entre otros.

Mi papel en este contexto es actuar como un asistente de investigación virtual, ofreciendo sugerencias y orientación basadas en la información que proporcionas sobre tu proyecto. Estoy aquí para ayudarte con preguntas específicas, proporcionar ideas y asistir en el desarrollo del análisis cualitativo de los datos. Si hay detalles adicionales o preguntas específicas que deseas abordar, por favor, házmelos saber.



Figura 2.9. Resultados de preguntas acerca del proyecto de investigación.

Consejo 2. ¡Usa los pulgares!

ChatGPT permite que los usuarios le brinden retroalimentación a través de los pulgares que se ubican en la parte inferior de sus respuestas (figura 2.10). Esto es esencial para comprender la preferencia de los usuarios hacia las respuestas de ChatGPT. Los pulgares –uno hacia arriba, positivo; otro hacia abajo, negativo– moldean el aprendizaje de la herramienta, lo que impulsa mejoras en la calidad y relevancia de las respuestas generadas para nuestro proyecto.

 ChatGPT

Hasta ahora, he entendido que estás llevando a cabo un proyecto de investigación en sociología política centrado en el análisis de discursos de investidura de los últimos presidentes peruanos. El objetivo específico que me has compartido se centra en identificar la forma en que se plantean los principales problemas y la crisis política en estos discursos.

La metodología que estamos discutiendo implica un enfoque cualitativo, donde se espera examinar cómo los presidentes abordan y comunican los desafíos políticos a través de sus discursos formales. Se mencionaron áreas clave de análisis, como la identificación de problemas, el análisis de la retórica utilizada, la jerarquía de problemas, las soluciones propuestas, entre otros.

Mi papel en este contexto es actuar como un asistente de investigación virtual, ofreciendo sugerencias y orientación basadas en la información que proporcionas sobre tu proyecto. Estoy aquí para ayudarte con preguntas específicas, proporcionar ideas y asistir en el desarrollo del análisis cualitativo de los datos. Si hay detalles adicionales o preguntas específicas que deseas abordar, por favor, házmelos saber.



Figura 2.10. Ubicación de los pulgares de retroalimentación en la interfaz de ChatGPT.

CAPÍTULO 3

Explorar mis datos paso a paso con ChatGPT

3.1. Primeros pasos: orientaciones básicas

Objetivo

El objetivo principal del proceso de exploración de datos cualitativos es conocer de forma más profunda nuestros datos y fomentar la generación de nuevas miradas acerca de los contenidos estudiados. La exploración acoge impresiones acerca de los contenidos y admite lo que es «llamativo», ya que necesita apertura para descubrir cosas nuevas y encontrar caminos alternativos para aproximarnos a nuestros datos. Por estas razones, se recomienda abordar este proceso con una actitud abierta, receptiva y espontánea. Este proceso consta de dos aproximaciones distinguibles: casuística o intra-caso; y comparativa o inter-caso. En esta guía nos centramos en el primer tipo de aproximación, para el cual se presentan cuatro técnicas concretas. En la figura 3.1, se ilustra una forma de explorar datos cualitativos a través de una secuencia de cuatro preguntas.

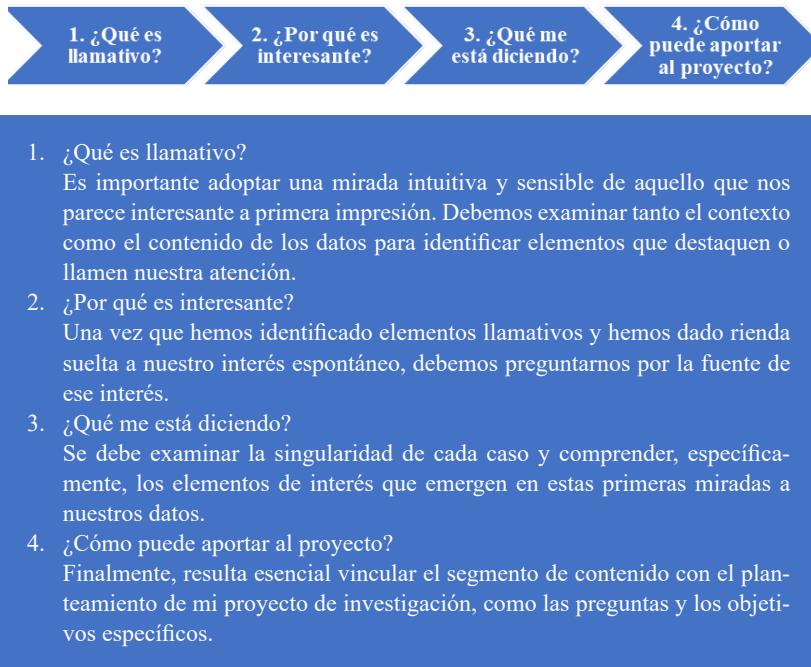


Figura 3.1. Secuencia de preguntas motivadoras para la exploración inicial de datos cualitativos.

Fuente: elaboración propia, libremente inspirada en Richards (2015).

Técnicas

Considerando estas interrogantes que guían nuestra mirada exploratoria a los datos, cabe presentar las cuatro técnicas concretas que tocamos en este capítulo. Si bien estas pueden ser realizadas en cualquier orden, se recomienda hacerlo en el que aparece a continuación.

Selección de citas llamativas

Se basa en identificar segmentos que nos resultan particularmente interesantes por su significado, composición, contexto o alguna otra razón que no necesariamente logramos explicar.

Buscamos seleccionar algunas citas llamativas en todos nuestros documentos analizados, para lo cual preguntamos: ¿Qué expresiones resaltan? ¿Qué segmentos parecen importantes para el emisor del mensaje? ¿Qué citas representan mejor cada caso?

Síntesis casuística

Es el procedimiento de creación de resúmenes que, de forma concisa, reflejen la información considerada importante para cada caso. Así, para cada uno podemos realizar una síntesis de aquello que nos ha llamado la atención espontáneamente como investigadores o, alternativamente, sintetizarlo en base al interés de nuestras preguntas específicas de investigación. Dos preguntas pueden ayudar a este procedimiento de síntesis: ¿Cómo se podría condensar este caso? ¿Qué elementos importantes son específicos de este caso?

Recurrencia de términos

Está basada en el conteo del uso de ciertos términos o palabras dentro de un texto estudiado. Se trata de una exploración léxica para trazar primeras pistas de semejanzas, patrones y/o temas recurrentes. Dos preguntas que pueden guiar este procedimiento son: ¿Cuáles son las palabras o expresiones que se repiten más en este caso? ¿Existen términos que se repiten más al tratar ciertos temas o en ciertos momentos del texto?

Constatación de inconsistencias

Implica poner en perspectiva diferentes segmentos del texto para identificar incongruencias o incoherencias al interior de cada caso. Se trata de un procedimiento de triangulación que permite descubrir la lógica de la articulación de los contenidos textuales en un nivel más detallado. Dos preguntas que po-

demos hacernos para emprender esta forma de explorar son:
¿Existe coherencia al interior de las diferentes partes del texto?
¿Por qué parece haber incongruencias en este caso?

Exploración e inteligencia artificial

Antes de empezar a explorar, es importante que estemos familiarizados con los contenidos de interés, es decir, asegurarnos de que comprendemos su estructura, características y limitaciones. Para ello, debemos leer detenidamente todos los contenidos al menos una vez. En caso de que se haya realizado una transcripción, antes de comenzar el trabajo de exploración es importante confirmar la calidad de la gramática, ortografía y puntuación de los textos. El potencial heurístico del proceso de exploración se incrementa considerablemente cuando conocemos previamente los contenidos y estos han pasado por una revisión de forma.

La IA puede ser una herramienta de gran utilidad para complementar el trabajo exploratorio del investigador, ya que, utilizada de forma adecuada, potenciaría el valor heurístico del proceso. En este sentido, el mayor valor que brinda trabajar con ChatGPT es la generación de insumos que enriquezcan, complementen y evalúen lo que el investigador elabora previamente. Para ello, se debe considerar que, para cada técnica de exploración, pediremos a ChatGPT que nos brinde insumos específicos que complementen aquellos que como investigadores hemos elaborado previamente.

Caso ilustrativo

En este capítulo, trabajaremos con el proyecto «Discursos presidenciales» como caso ilustrativo para la explicación de cada una de las cuatro técnicas descritas. Como se detalla en la tabla 3.1, se trata de un proyecto cuyo objetivo es explicar cuáles fueron los aspectos prioritarios en la visión de país y su situación en los discursos de investidura ante el Congreso de la República de los

Tabla 3.1. Términos del proyecto de investigación «Discursos presidenciales»

Pregunta general	¿Cuáles son los elementos esenciales de la visión de país que expresaron los presidentes peruanos en su primer mensaje presidencial?
Objetivo general	Explicar cuáles fueron los aspectos prioritarios en la visión de país y su situación en los discursos de investidura de los últimos presidentes.
Pregunta específica 1	¿Cómo se aborda la crisis política y los principales problemas que aquejan al país?
Objetivo específico 1 (OE1)	Comprender la forma en que los últimos presidentes plantean los principales problemas y crisis política que enfrenta el país.

últimos presidentes peruanos. Aquí se utilizará el OE1 para realizar la exploración de datos cualitativos. Asimismo, del conjunto de casos estudiados, se utilizará uno en particular, presentado en el recuadro 3.1, para ilustrar los procedimientos de cada técnica de exploración.

Recuadro 3.1. Caso ilustrativo: Discurso presidencial de un presidente peruano

Recibo con humildad y profundo fervor patriótico el cargo de presidente de la República. Declaro ante el Congreso, ante los presidentes amigos aquí reunidos y ante el pueblo peruano que, fiel al mandato de las urnas y en pleno respeto al Estado de Derecho, dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la exclusión y la pobreza construyendo un Perú para todos, atento siempre, en los más frágiles de nuestros hermanos. Exigiré el mismo compromiso y la misma energía a todo el equipo que me acompaña en el Ejecutivo. La democracia peruana será plena cuando la justicia y la paz social, la soberanía nacional y la seguridad de nuestras familias constituyan el zócalo de nuestra Nación, cuando la igualdad sea patrimonio de todos y la exclusión social desaparezca aun en los lugares más remotos del país.

Queremos que la expresión misma de «exclusión social» se borre para siempre de nuestro lenguaje y de nuestra realidad. Asumiré este reto con mi palabra y con mi vida. Hace casi un siglo, en 1914, Víctor Andrés Belaunde, uno de los grandes intelectuales y políticos del siglo XX, al terminar un discurso en la Universidad Nacional Mayor de San Marcos, lanzó a los jóvenes una proclama que era, al mismo tiempo, un grito de batalla y una demanda: ¡Queremos patria! Esta proclama años después fue recogida por José Carlos Mariátegui y Víctor Raúl Haya de la Torre para convertirla en pensamiento y acción. Esa reivindicación de la patria y de la Nación ha sido, como ahora, el sueño encendido de generaciones y de pueblos. Desde la fundación de la República, la patria constituye una aspiración al destino común, una esperanza, una promesa inacabada por la que el pueblo peruano ha estado siempre dispuesto a entregar su vida en la paz y en la guerra, para defenderla y para legarla a sus hijos con orgullo y con fe. La patria es nuestra Historia común, es el espacio donde todos los peruanos y peruanas queremos vivir en paz y en democracia. Por ello todos nos sentimos orgullosos de nuestra bandera, símbolo de nuestras luchas y de nuestro amor a la paz. Esta aspiración a una patria inclusiva es la que abre el camino al progreso social.

Para hacerla posible nos presentamos ante el pueblo y por eso estamos aquí. Asumo este reto como un desafío y como una promesa que expreso hoy ante todos los peruanos, y especialmente ante mis hijos y ante todos los niños y niñas del país, que serán el futuro de esta nación. Este es nuestro patrimonio más rico. Por ellos y para ellos voy a cumplir mi promesa de hacer de este país un lugar donde todos disfruten del mismo derecho a la plenitud y a la felicidad, a una vida digna y a una vejez protegida. Peruanas y peruanos: el cinco de junio una mayoría de ciudadanos expresó su deseo de que el crecimiento económico y la inclusión social marchen juntos para transformar nuestro país en una patria de oportunidades para todos. El incremento desmedido de los conflictos, muchos de ellos absurdamente violentos, nos demuestra, día a día, que es urgente reparar las injusticias, corregir el rumbo y restablecer el diálogo en nuestra sociedad.

El Perú es un país plurilingüe y multicultural. Esta múltiple diversidad constituye sin duda nuestra mayor riqueza. Sin embargo, durante mucho tiempo ha existido un discurso y una práctica de la exclusión, del rechazo a

la diferencia, un «tú no eres igual que yo», que cobijó la discriminación y la intolerancia. Esto resulta cuando menos extraño porque los comportamientos excluyentes provienen muchas veces de quienes elogian nuestra diversidad cultural. Pero precisamente esa diversidad cultural proviene de entender que nuestra nación es un crisol de razas y tradiciones. Son ellas las que sustentan, por ejemplo, nuestra extraordinaria gastronomía, hoy admirada y reconocida en el mundo. Porque somos diferentes, pero iguales en el fondo, labramos nuestra existencia en el trabajo y en el esfuerzo cotidiano. Somos mezcla y creatividad. Somos imaginación y trabajo. Y esta diversidad, que queremos integradora y no marginadora, constituye el fundamento de nuestra riqueza.

Nuestro país trabajador, honrado y diverso se encuentra fracturado y herido. Sufre el abandono de los políticos y de un Estado insensible, burocrático y centralista. Un Estado que le ha dado la espalda al interior del país, un Estado que sufre de «mal de altura» o «soroche» y que se niega a subir los Andes y extenderse en nuestra Amazonía. Un Estado acechado por la corrupción y el despilfarro, alejado de sus ciudadanos, incapaz de protegerlos de la violencia y la criminalidad. Sueño con un Perú donde la vida no sea un riesgo, donde las ciudades sean espacios seguros, donde el ciudadano se encuentre protegido. Para esto necesitamos más Estado, más patria, y que la corrupción sea sancionada. La democracia expresó en las urnas un mandato y ese mandato debe ser honrado. Restablecer el valor de la palabra empeñada ante el pueblo constituye el eje de la recuperación de un sistema de valores éticos inherentes a la República. Sin embargo, transformar el país no es tarea fácil. No solo porque el cambio es siempre una tarea de multitudes, sino porque también implica enfrentar y superar nuestros problemas, proponer una nueva manera de convivir.

La realidad exige transformaciones para que la igualdad, la tolerancia, el reconocimiento de nuestra diversidad y un desarrollo que nos incluya a todos sean, al mismo tiempo, un acto civilizatorio y un compromiso colectivo. La realidad nos interpela diariamente y nos exige un nuevo contrato social que haga posible la convivencia armoniosa de todos los peruanos. Evoco aquí nuevamente la figura de Haya de la Torre y su legado, plasmado singularmente en la Constitución de 1979, que constituye para mí una verdadera inspiración por su contenido nacional, democrático y de

libertad. El gran Nelson Mandela, en un célebre discurso pronunciado en el marco de la Unesco, afirmó, con la convicción que lo caracterizaba, que la igualdad, la equidad económica y la justicia social eran la base de toda democracia. Él dijo: «No hay democracia con miseria, no hay democracia con asimetrías sociales». Y porque creo en la justicia de esta frase, yo he jurado respetar y defender la democracia. Fortalecerla en sus valores igualitarios para hacerla legítima ante el pueblo y así será. El gobierno se propone reconocer y reivindicar a los tres millones de peruanos migrantes. Para empezar, mejoraremos la defensa consular de los peruanos en el exterior y rebajaremos el costo de las remesas mediante convenios que realizará el Banco de la Nación. Deseamos que quien desea regresar a su patria lo haga y para ello fomentaremos que sea sujeto de crédito. Los detalles de estas políticas aquí señaladas serán presentados por la Presidencia del Consejo de Ministros, como corresponde, ante el parlamento.

De la tradición militar, que no olvido y llevaré en el corazón hasta la muerte, conservo orgulloso la tenacidad, la austereidad y el amor por el Perú y sus intereses. En esa tradición, a diferencia de lo que piensan algunos, se sabe mandar, pero también obedecer, hay jerarquía, pero también fraternidad, hay disciplina, pero también intercambio de opiniones. Esa tradición se funde con el espíritu generoso del Perú, lejano al odio. No vengo en son de guerra sino en son de paz, sin venganza y sin rencor. Yo, que he sido acusado casi de todo, he aprendido a perdonar hace muchos años, antes incluso de hacer política. Por eso, a los que aún persisten en el encono, les pido que bajen sus espadas y sus lanzas. A los que demandan salarios y derechos, les digo que no bajen sus banderas, pero que sepan que todo cambio, para ser sostenible, debe ser gradual y racional. A mis partidarios, les pido consecuencia, lealtad, sacrificio, inteligencia y honradez. A la oposición, la llamo a la responsabilidad. Le pido vigilancia y que, desde su posición, respete también el mandato de las urnas, su mandato y el nuestro. Al terminar este mensaje, reitero que soy un soldado de la democracia. Y como dice la frase célebre de la Independencia: «Firmes y felices por la unión». ¡Muchas gracias a todo el país! ¡Viva el Perú!

Fuente: Congreso de la República del Perú (2023).

3.2. Técnica 1. Selección de citas llamativas

Para realizar una selección de citas llamativas mediante el uso de ChatGPT, se siguen los pasos señalados a continuación.

Paso 1

Iniciamos el entrenamiento CORI-F³ (figura 3.2).

 **Tú**

- En el marco de una investigación científica en sociología política, estoy usando datos cualitativos para analizar el caso peruano. Para aproximarme a este tema uso discursos de investidura de los últimos presidentes.
- El objetivo específico por trabajar es identificar la forma en que se plantean los principales problemas y la crisis política en el contenido de los discursos.
- En este proyecto te desempeñas como asistente de investigación experto en el análisis de datos cualitativos.
- Debes basar tus respuestas exclusivamente en la información que te proporcionaré.



 **ChatGPT**

Entendido, estaré encantado de ayudarte en tu proyecto de investigación sobre la sociología política en el contexto peruano, centrándome en el análisis de los discursos de investidura de los últimos presidentes. Por favor, proporciona la información que necesitas que utilice para avanzar en tu investigación.

 **Tú**

¿Has entendido las indicaciones?

 **ChatGPT**

Sí, he entendido las indicaciones. Estamos enfocándonos en analizar los discursos de investidura de los últimos presidentes en el contexto peruano, con el objetivo específico de identificar la forma en que se plantean los principales problemas y la crisis política en el contenido de estos discursos. Por favor, proporciona los datos específicos que tienes disponibles para que pueda ayudarte con el análisis de datos cualitativos.

Figura 3.2. Entrenamiento CORI-F para la exploración.

3 Ver detalles y explicación sobre CORI-F en la sección 2.2.

Paso 2

Una vez validado el primer paso, solicitamos a ChatGPT que realice una selección de citas del discurso presidencial que le hemos proporcionado, como ilustra la figura 3.2. Esta solicitud debe estar acompañada del discurso completo⁴ (por motivos de espacio, en la figura se ha recortado). En cuanto al formato (la F de CORI-F), solicitamos que ChatGPT seleccione cinco citas y que incluya una breve explicación sobre su relevancia con el OE.

 **TÚ**

Con el objetivo específico del análisis (identificar la forma en que se plantean los principales problemas y la crisis política en el país), selecciona cinco citas significativas, con un formato legible y con una breve explicación sobre su relevancia para el objetivo específico, del siguiente discurso:

"Recibo con humildad y profundo fervor patriótico el cargo de Presidente de la República. Declaro ante el Congreso, ante los presidentes amigos aquí reunidos y ante el pueblo peruano que, fiel al mandato de las urnas y en pleno respeto al Estado de Derecho, dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la

Figura 3.3. *Prompt* para realizar la selección de citas relevantes.⁵

4 ChatGPT puede recibir hasta 3000 palabras, aproximadamente.

5 El discurso ha sido recortado por motivos de espacio.

A continuación, la figura 3.4 presenta la respuesta de ChatGPT al *prompt*.

 ChatGPT

Aquí tienes cinco citas significativas del discurso, junto con breves explicaciones sobre su relevancia para el objetivo específico de identificar la forma en que se plantean los principales problemas y la crisis política:

- "Dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la exclusión y la pobreza construyendo un Perú para todos."**
 - * *Explicación:* La cita resalta el compromiso del presidente con abordar los problemas de exclusión y pobreza en el país, lo cual indica una conciencia de los desafíos sociales fundamentales.
- "El incremento desmedido de los conflictos, muchos de ellos absurdamente violentos, nos demuestran, día a día, que es urgente reparar las injusticias, corregir el rumbo y restablecer el diálogo en nuestra sociedad."**
 - * *Explicación:* Esta declaración señala la urgencia de abordar los conflictos y la violencia, indicando una percepción de crisis política y social que necesita ser atendida.
- "Sueño con un Perú donde la vida no sea un riesgo, donde las ciudades sean espacios seguros donde el ciudadano se encuentre protegido. Para esto necesitamos más estado, más patria y que la corrupción sea sancionada."**
 - * *Explicación:* La referencia a la seguridad ciudadana y la necesidad de combatir la corrupción sugiere la percepción de problemas estructurales y la importancia de fortalecer las instituciones.
- "La realidad nos interpela diariamente y nos exige un nuevo contrato social que haga posible la convivencia armoniosa de todos los peruanos."**
 - * *Explicación:* La mención de un "nuevo contrato social" indica la necesidad de cambios profundos en la forma en que la sociedad peruana aborda sus desafíos, sugiriendo una conciencia de la crisis política y la necesidad de una transformación integral.
- "A la oposición la llamo a la responsabilidad. Le pido vigilancia y que, desde su posición, respete también el mandato de las urnas, su mandato y el nuestro."**
 - * *Explicación:* Esta declaración destaca la importancia de la responsabilidad política y del respeto mutuo en un contexto de polarización política, sugiriendo un llamado a la unidad y la estabilidad institucional para superar la crisis política.

Estas citas proporcionan un punto de partida para analizar cómo el presidente aborda los problemas y la crisis política en su discurso de investidura.

Figura 3.4. Resultados de la selección de citas relevantes en ChatGPT.

La selección de citas obtenida por medio de ChatGPT puede ayudarnos a ubicar contenidos llamativos que podrían ser valiosos para nuestra exploración a la luz de nuestro OE de investigación. Resulta esencial que podamos verificar que las citas corresponden con el contenido de los datos provistos a la herramienta y que sus explicaciones sean coherentes. Asimismo, se recomienda enfáticamente que los resultados generados por ChatGPT sean insumos para complementar una selección inicial realizada previamente por el investigador (durante la etapa de trabajo de campo, la transcripción o la relectura). Así, estos insumos pueden servirnos para mejorar nuestra selección de citas llamativas o para evaluar la que hemos realizado manualmente en un primer momento.

Consejo 3. Recuerda frecuentemente el OE

Si bien ChatGPT tiene la capacidad de recordar, su memoria no es excelente. Es por ello por lo que se recomienda reiterar con frecuencia un aspecto tan importante para una investigación como lo es el OE. De esta manera, se orienta mejor la respuesta de la herramienta y se obtienen insumos de exploración más precisos.

 TÚ

Con el objetivo específico del análisis (identificar la forma en que se plantean los principales problemas y la crisis política en el país), selecciona cinco citas significativas, con un formato legible y con una breve explicación sobre su relevancia para el objetivo específico, del siguiente discurso:

"Recibo con humildad y profundo fervor patriótico el cargo de Presidente de la República. Declaro ante el Congreso, ante los presidentes amigos aquí reunidos y ante el pueblo peruano que, fiel al mandato de las urnas y en pleno respeto al Estado de Derecho, dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la

Figura 3.5. Señalamiento del OE en el *prompt* del procedimiento de selección de citas.⁶

6 El discurso ha sido recortado por motivos de espacio.

3.3. Técnica 2. Síntesis casuística

Para realizar una selección de citas llamativas con el uso de ChatGPT, se hace lo siguiente.

Solicitar a ChatGPT elaborar un resumen corto del discurso. Para mayor precisión, recordamos el OE1. En cuanto al formato (la F de CORI-F), indicamos que se genere un párrafo de máximo cien palabras, como se ilustra en la figura 3.6.

P

Tú

Con el objetivo específico del análisis (identificar la forma en que se plantean los principales problemas y la crisis política en el país), elabora un resumen de 100 palabras, en formato de párrafo, del siguiente discurso:

"Recibo con humildad y profundo fervor patriótico el cargo de Presidente de la República. Declaro ante el Congreso, ante los presidentes amigos aquí reunidos y ante el pueblo peruano que, fiel al mandato de las urnas y en pleno respeto al Estado de Derecho, dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la

Figura 3.6. *Prompt* para realizar la síntesis casuística en ChatGPT.⁷

A continuación, la figura 3.7 presenta la respuesta de ChatGPT al *prompt*.

C

ChatGPT

En su discurso de investidura, el presidente expresó un firme compromiso con la erradicación de la exclusión y la pobreza en Perú, prometiendo construir un país para todos. Destacó la importancia de la democracia plena, la justicia y la paz social como fundamentos nacionales. Reconoció la diversidad cultural como riqueza y denunció la exclusión y la discriminación históricas. Señaló la urgencia de abordar la violencia y los conflictos excesivos como un llamado a reparar injusticias. Abogó por un Estado más presente, la lucha contra la corrupción y un nuevo contrato social para la convivencia armoniosa. Finalmente, llamó a la unidad, el perdón y la responsabilidad para enfrentar los desafíos del país.

Figura 3.7. Resultados de la síntesis del caso por parte de ChatGPT.

7 El discurso ha sido recortado por motivos de espacio.

La síntesis del caso obtenida por medio de ChatGPT puede ayudarnos a identificar elementos importantes que serían valiosos para nuestra exploración a la luz de nuestro OE de investigación. Es muy importante que leamos detenidamente y podamos confirmar que el resumen es coherente. Asimismo, se recomienda enfáticamente que los resultados generados por ChatGPT sean utilizados como insumos para complementar o mejorar la síntesis elaborada previamente por el investigador de forma manual. Así, los insumos brindados por ChatGPT pueden servir para evaluar y/o mejorar nuestra mirada del caso tratado.

3.4. Técnica 3. Recurrencia de términos

Para realizar una identificación de recurrencia de términos mediante el uso de ChatGPT, se realiza lo siguiente.

Pedir a ChatGPT hacer un recuento de las palabras más frecuentes en el discurso. En cuanto al formato (la F de CORI-F), indicamos que se genere una lista de las palabras más recurrentes, como se ilustra en la figura 3.8.

P **Tú**
Con el objetivo específico del análisis (identificar la forma en que se plantean los principales problemas y la crisis política en el país), realiza una lista de las palabras más frecuentes, que sean relevantes para nuestro objetivo específico, e incluye el número de veces que se repiten, en el siguiente discurso:

"Recibo con humildad y profundo fervor patriótico el cargo de Presidente de la República. Declaro ante el Congreso, ante los presidentes amigos aquí reunidos y ante el pueblo peruano que, fiel al mandato de las urnas y en pleno respeto al Estado de Derecho, dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la

Figura 3.8. Prompt para realizar un recuento de las palabras más frecuentes en ChatGPT.⁸

8 El discurso ha sido recortado por motivos de espacio.

La figura 3.9 presenta la respuesta de ChatGPT a dicho *prompt*.

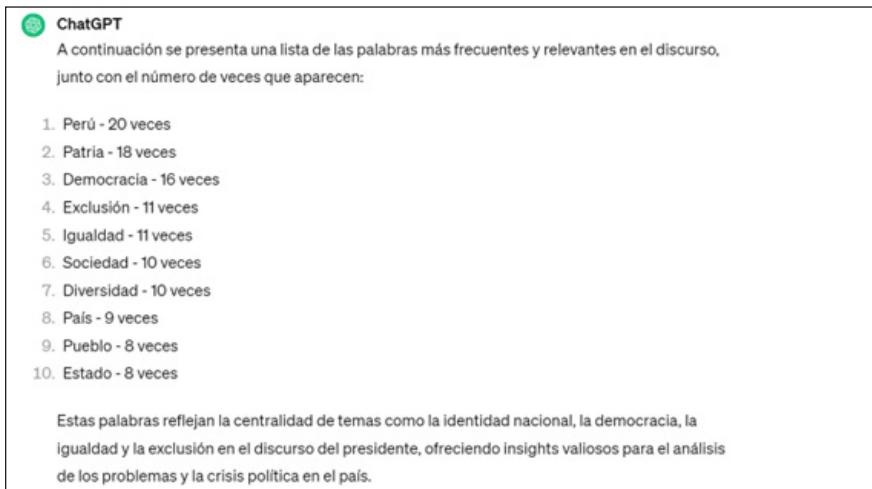


Figura 3.9. Resultados del recuento de términos por parte de ChatGPT.

El recuento de palabras obtenido por medio de ChatGPT ayuda a identificar la recurrencia de términos. Tal recuento podría ser aplicado a partes específicas del contenido o a su totalidad, como en el ejemplo previo. Es muy importante que leamos detenidamente y podamos confirmar que el resultado es coherente, por ejemplo, a través de búsquedas de palabras en otros programas. Estos insumos pueden servirnos para evaluar y/o mejorar nuestra mirada del caso tratado y comprender su singularidad a la luz de cada uno de nuestros OE.

3.5. Técnica 4. Constatación de inconsistencias

Para realizar una constatación de inconsistencias mediante ChatGPT, se hace lo señalado a continuación.

Como muestra la figura 3.10, se solicita a ChatGPT que identifique las dos más grandes aparentes inconsistencias en el discurso en relación con el OE 1.

 Tú
Identifica las dos más grandes aparentes inconsistencias en el contenido de este discurso en relación con el objetivo específico:

"Recibo con humildad y profundo fervor patriótico el cargo de Presidente de la República. Declaro ante el Congreso, ante los presidentes amigos aquí reunidos y ante el pueblo peruano que, fiel al mandato de las urnas y en pleno respeto al Estado de Derecho, dedicaré toda mi energía a sentar las bases para que borremos definitivamente de nuestra historia el lacerante rostro de la

Figura 3.10. *Prompt* para identificar inconsistencias en ChatGPT.⁹

Como resultado, la figura 3.11 presenta la respuesta de ChatGPT al *prompt*.

 ChatGPT

- 1. Inconsistencia en la promesa de eliminación de exclusión social:**
El discurso promete eliminar la exclusión social de la historia del país, pero no proporciona detalles específicos sobre cómo se abordará este desafío ni presenta medidas concretas para lograr la inclusión social. La falta de estrategias claras puede plantear interrogantes sobre la viabilidad de la promesa y la efectividad del plan propuesto.
- 2. Contradicción en la visión de la diversidad cultural:**
Aunque el discurso destaca la riqueza de la diversidad cultural del país y aboga por la integración en lugar de la marginación, también reconoce la existencia de un discurso y práctica de exclusión en el pasado. Esta contradicción plantea la pregunta de cómo se reconciliarán estas perspectivas divergentes y cómo se abordarán los desafíos persistentes relacionados con la discriminación y la intolerancia.

Figura 3.11. Resultados de la identificación de inconsistencias por parte de ChatGPT.

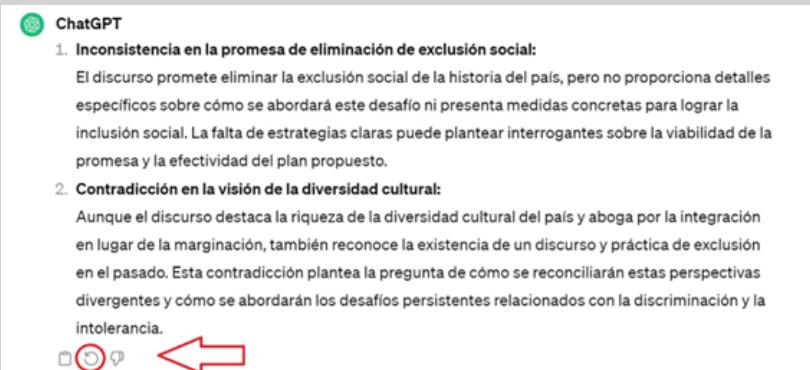
La identificación de incongruencias con el uso de ChatGPT es útil para mirar en perspectiva diferentes segmentos del caso explorado. Ello puede aplicarse a la totalidad del contenido de un

9 El discurso ha sido recortado por motivos de espacio.

caso, como en el ejemplo previo, o a secciones específicas. Es muy importante leer detenidamente el resultado y evaluar su calidad, sobre la base de nuestro buen conocimiento de los contenidos tratados. Estos insumos podrían servirnos para evaluar y/o desarrollar nuestra aproximación acerca de los vínculos entre diferentes secciones de un mismo caso, lo que nos permite enriquecer nuestra triangulación de informaciones al interior de tal caso y descubrir nuevas pistas de comparación a su interior.

Consejo 4. Usa regenerate

ChatGPT provee una opción que nos permite solicitar una nueva elaboración de la última respuesta que ha provisto: la opción *regenerate*, o regenerar. Como muestra la figura 3.12, al final de la respuesta se presenta esta opción que se activa con un solo clic en el botón correspondiente. El programa formulará de inmediato una respuesta alternativa a la primera dada. Esto nos puede llevar a obtener una mejor respuesta, una que se acerque más a la expectativa que se tenía al momento de brindar la solicitud. Una vez creada esta nueva respuesta, se mantendrá disponible la versión anterior, a la que se podrá acceder en el espacio de la respuesta.



The screenshot shows a portion of the ChatGPT interface. At the top left is the ChatGPT logo. Below it is a list of two numbered items under the heading "Inconsistencia en la promesa de eliminación de exclusión social". The first item discusses the lack of details on how to address the challenge of exclusion. The second item discusses the contradiction between diversity and past exclusion. At the bottom of the list is a red arrow pointing to a small circular button containing a circular arrow icon, which is the "regenerate" button. To the left of the arrow are three other small circular icons: a document, a person, and a gear.

1. **Inconsistencia en la promesa de eliminación de exclusión social:**
El discurso promete eliminar la exclusión social de la historia del país, pero no proporciona detalles específicos sobre cómo se abordará este desafío ni presenta medidas concretas para lograr la inclusión social. La falta de estrategias claras puede plantear interrogantes sobre la viabilidad de la promesa y la efectividad del plan propuesto.
2. **Contradicción en la visión de la diversidad cultural:**
Aunque el discurso destaca la riqueza de la diversidad cultural del país y aboga por la integración en lugar de la marginación, también reconoce la existencia de un discurso y práctica de exclusión en el pasado. Esta contradicción plantea la pregunta de cómo se reconciliarán estas perspectivas divergentes y cómo se abordarán los desafíos persistentes relacionados con la discriminación y la intolerancia.

Figura 3.12. Opción *regenerate* en ChatGPT.

3.6. Evaluación de los resultados

En esta última sección del capítulo 3, se presentan una evaluación de los resultados que nos ha proporcionado ChatGPT a partir del proyecto «Discursos presidenciales». Se trata de un ejemplo ilustrativo de los elementos a tomar en consideración en una evaluación crítica de esta herramienta.

Las capacidades de este programa se encuentran en constante evolución, y sin duda seguirán en crecimiento en cuanto a desempeño. Sin embargo, es preciso rendir cuenta de su estado actual. La tabla 3.2 presenta, para cada una de las técnicas usadas, algunos elementos de la evaluación realizada a los insumos generados en términos de fortalezas, debilidades y eficacia. Esta evaluación no es exhaustiva y busca principalmente llamar la atención acerca de la importancia de mirar de forma crítica cada procedimiento en el trabajo con IA.

En la tabla 3.2, el primer aspecto, las fortalezas, subraya el valor que aporta la herramienta en función del tipo de procedimiento que se le pidió. Por ejemplo, podemos constatar que ChatGPT tiene un mejor desempeño en la síntesis de textos que en la identificación de inconsistencias. El segundo aspecto, las debilidades, hace explícitas las limitaciones encontradas para cada técnica. Por ejemplo, hallamos que la recurrencia de términos tiene mayores limitaciones, como la identificación de sinónimos, que la selección de citas llamativas. Por último, la eficacia es un puntaje que valora los resultados de las técnicas de exploración sobre la base de los resultados. Notamos que, a partir de los *prompts* presentados aquí, ChatGPT es más eficaz en la selección de citas llamativas y en la síntesis casuística.

Tabla 3.2. Evaluación del desempeño de ChatGPT en la exploración de datos cualitativos

Técnica	Fortalezas	Debilidades	Eficacia
Selección de citas llamativas	- Presentación de variedad de posibles citas llamativas. - Elaboración de explicaciones vinculadas con los OE de ayuda.	- La falta de contexto específico puede generar ciertas citas con poca relevancia y precisión.	4/5
Síntesis casuística	- Muy rápido procesamiento de grandes volúmenes de texto. - Resúmenes adecuados y pertinentes.	- Algunos resúmenes tienden a ser muy generales y no siempre se ajustan al OE.	4/5
Recurrencia de términos	- Realización de un recuento preciso de las palabras.	- Dificultad para contabilizar correctamente sinónimos. - Limitaciones por el idioma y las expresiones del contexto específico.	3/5
Identificación de inconsistencias	- Identificación de patrones iniciales y detalles específicos que podrían pasar desapercibidos en una revisión superficial.	- La identificación de inconsistencias puede variar significativamente en calidad. - La presencia de sesgos inherentes al modelo puede afectar la objetividad en la identificación de patrones.	3/5

Consejo 5. Haz una retroalimentación a ChatGPT

Hacer una retroalimentación por escrito a ChatGPT, con el *prompt* de mejora, ayudará a que se elabore una nueva versión ajustada de la exploración específica. Para ello, puedes detenerte un momento y considerar qué fortalezas y debilidades has encontrado en ChatGPT.

En relación a lo anterior, y como se ha visto en los consejos 2 y 4, es oportuno utilizar los pulgares y la opción *regenerate* para evaluar los resultados y mejorar futuras elaboraciones; asimismo, respecto a los consejos 1 y 5, que ayudan a obtener mejores resultados, es importante recordar el OE y realizar *prompts* que tengan una estructura ordenada.

CAPÍTULO 4

Codificar contenidos paso a paso con ayuda de ChatGPT

4.1. Primeros pasos: orientaciones básicas

Objetivo

La codificación desempeña un papel crucial en el análisis de datos cualitativos. Tras la exploración inicial, ella se convierte en el medio para categorizar, comparar e interpretar nuestros datos (Char-maz, 2014). Este proceso implica clasificar segmentos de texto mediante el uso de palabras clave o «códigos». Estos pueden ser muy descriptivos o más abstractos –como nociones o concep-tos–. El uso de ambos tipos de códigos, y su posterior estudio, revisión y relacionamiento, permite una comprensión más integral y profunda de los datos.

El proceso de codificación se divide en dos ciclos. En el pri-mero, se lee rápidamente los contenidos ya conocidos y se com-para constantemente los contenidos intra- e inter-caso (Glaser y Strauss, 2006) con el objetivo de asignar códigos a los seg-mentos de texto. El movimiento rápido a través de los datos es importante porque ello fomenta el encuentro de similitudes en los contenidos de diversos casos, es decir, impulsa la transversali-

dad (Charmaz, 2014). En este sentido, la codificación nos brinda la posibilidad de trascender la mirada casuística. Posteriormente, en el segundo ciclo de codificación, se procede a releer los contenidos codificados, por caso y por código, para armonizar los códigos generados. Así, en esta segunda etapa se realizan los cambios necesarios –como fusión, eliminación, agrupación o creación de familias de códigos– para alcanzar un entendimiento mayor de lo que nos dicen los datos (Saldaña, 2021). Del mismo modo, en esta etapa se crean categorías más amplias y se busca comprender sus relaciones subyacentes. En esta guía nos concentraremos en la codificación durante el primer ciclo, para la cual se presentan dos técnicas concretas.

Técnicas de codificación

La forma que toma el proceso de codificación depende del tipo de enfoque en la elaboración de códigos. Para ello, existen dos técnicas principales, cada una de las cuales responde a tipos de razonamiento diferentes: la deducción y la inducción. Mientras que el razonamiento deductivo parte de hipótesis basadas en la teoría y en conceptos escogidos cuidadosamente para analizar una realidad, el razonamiento inductivo parte de temas de interés y se concentra en las particularidades del contexto estudiado para, posteriormente, interpretar por medio de conceptos y construcción de hipótesis. Lo importante a tener en cuenta es que la elección de la técnica idónea debe basarse en la naturaleza del OE al que se apunta. A continuación, se presentan las dos técnicas de codificación mencionadas.

Codificación predefinida

Sigue un razonamiento deductivo, por lo que se establece un repertorio de códigos en función de las hipótesis planteadas durante la formulación del proyecto de investigación. Dichas

hipótesis están respaldadas por un desarrollo teórico que precede a la fase de análisis de datos. En el contexto de una investigación social o educativa, para probar la validez de las hipótesis de trabajo, resulta necesario plasmar sus elementos constitutivos en el instrumento de recolección de datos (guía de entrevista o de observación, matriz de sistematización de documentos o informaciones, guion de grupo focal, entre otros). En este sentido, para el caso de una guía de entrevista, por ejemplo, las preguntas contendrán información valiosa para la elaboración de los códigos.

En esta técnica de codificación, se busca clasificar los contenidos en función de los códigos que hemos creado, los cuales, a su vez, provienen de los elementos constitutivos de las hipótesis de la investigación.

Codificación emergente

La codificación emergente sigue un razonamiento inductivo cuya orientación es adaptarse a lo que dicen los contenidos y crear códigos en función a lo que dicen los datos. No obstante, esta «libertad» para crear códigos según lo que emerge de los textos debe enmarcarse en los temas de interés delimitados en los objetivos de investigación. En investigaciones sociales y educativas, esta técnica sirve para rendir cuenta de la realidad específica de cada contexto, persona o grupo analizado, lo que permite que aspectos propios que caracterizan a cada uno sean fuentes de abstracción. En este sentido, la principal fuente de generación de códigos no son los planteamientos de la investigación, sino los significados e informaciones de los datos mismos. Esta direccionalidad tiene mayor apertura al descubrimiento de procesos y relaciones no contempladas por el investigador.

Codificación e inteligencia artificial

Antes de empezar a codificar, resulta esencial que estemos familiarizados con los contenidos. Por lo general, se empieza a codificar después de haber realizado exploraciones de nuestros datos y de tener insumos acerca de las características de nuestros casos. De hecho, si desarrollamos una codificación emergente, tras la exploración tendremos muchas ideas de posibles códigos porque ya conocemos bien el material estudiado. Por otra parte, si realizamos una codificación predefinida, habremos identificado bastante contenido que es susceptible de ser codificado con este o aquel código de nuestro repertorio. En este sentido, el potencial analítico de la primera fase de codificación es mucho mayor cuando se ha explorado previamente los contenidos, más aun si se ha elegido utilizar responsablemente insumos generados por la IA.

La IA puede ser una herramienta de gran utilidad para complementar el trabajo de codificación por su capacidad de añadir valor analítico al proceso. Así, lo beneficioso de trabajar con ChatGPT es generar códigos que enriquezcan y complementen las elaboraciones que el investigador ha realizado previamente. En este proceso se recomienda también realizar una primera ronda manual, y posteriormente pedir a ChatGPT insumos para que nuestra codificación sea más creativa.

Caso ilustrativo

En este capítulo, trabajaremos con el proyecto «Trayectorias laborales y educativas de personas con discapacidad», notablemente desarrollado por las estudiantes Andrea Chávez, Daniela Neyra, Lorena Pacora y Nicole Vilca, como caso ilustrativo para la explicación de las técnicas de codificación predefinida y codificación emergente. Como se detalla en la tabla 4.1, se trata de un proyecto cuyo objetivo es comprender la visión que tienen las personas con

discapacidad acerca de sus trayectorias educativas y laborales en diversas regiones del país. Para este capítulo, trabajaremos únicamente con el OE1 de ese proyecto para realizar la codificación de entrevistas; para ilustrar la codificación predefinida, utilizaremos la guía de entrevista con la que se realizó la recolección de información, que puede ser consultada en la tabla 4.1; y, para ilustrar los procedimientos de codificación emergente, se utilizará un extracto de entrevista, que se presenta en el recuadro 4.2.

Tabla 4.1. Términos del proyecto de investigación «Trayectorias laborales y educativas de personas con discapacidad»

Pregunta general	¿Cómo han sido las experiencias educativas y laborales de las personas con discapacidad entrevistadas?
Objetivo general	Comprender la manera en que las personas con discapacidad entrevistadas ven su trayectoria educativa y laboral.
Pregunta específica 1	Desde su experiencia como personas con discapacidad, ¿cómo han vivido su formación y su trayectoria educativa?
OE1	Comprender cómo recuerdan su experiencia educativa las personas con discapacidad entrevistadas.

Recuadro 4.1. Guía de entrevista del proyecto «Trayectorias laborales y educativas de personas con discapacidad»

Buenas tardes señor / señora _____

Antes que nada, muchas gracias por su tiempo y disposición para realizar la entrevista.

Voy a grabar la entrevista con el fin de tener registrada la información, pero esta solo se utilizará para fines académicos y todo se mantendrá de manera confidencial. ¿Está bien?

Me llamo _____. Esta entrevista tiene como finalidad conocer un poco sobre su historia de vida en relación a su experiencia educativa y su experiencia en el mundo laboral. La entrevista no durará más de media hora. Antes de empezar, quisiera hacerle unas preguntas introductorias.

Información personal

¿Cuál es su nombre completo?

¿Cuántos años tiene?

¿En dónde nació?

¿Cuál es su lugar de residencia?

¿Cuál es su profesión?

¿Cómo se compone la jefatura de su hogar?

¿Qué limitación o discapacidad física tiene?

Experiencia educativa

¿Cómo ha sido su experiencia educativa?

Durante su educación primaria e inicial, ¿el entorno era accesible y se acomodaba a su discapacidad?

Durante su educación secundaria, ¿el entorno era accesible y se acomodaba a su discapacidad?

¿Me podría contar cómo fue su experiencia educativa después del colegio? ¿El entorno fue accesible y se acomodaba a su discapacidad?

¿Tuvo usted formación en materias no académicas? ¿Cómo fue su educación en estas materias? ¿El entorno era accesible y se acomodaba a su discapacidad?

¿Cómo era su relación con sus compañeros y compañeras? ¿Fue fácil hacer amigos?

Experiencia laboral

¿Cómo ha sido su experiencia laboral? ¿Qué dificultades ha tenido para insertarse al mundo laboral?

¿Cómo fue la elección de su lugar de trabajo? ¿Fue la única opción para usted o lo escogió porque era el que se adecuaba mejor a sus necesidades?

¿Cómo era trabajar en ese lugar? ¿Cómo era su relación con sus compañeros?

¿El entorno era accesible y se acomodaba a su discapacidad?

Recuadro 4.2. Fragmento de entrevista a Celia, proyecto «Trayectorias laborales y educativas de personas con discapacidad»

D3. Celia (74-111) «... ella tiene que aprender a levantarse...»

Yo, primaria, lo realicé en el colegio 463, que quedaba, bueno, que queda, ahora a dos cuadras y media, pero antes quedaba a una cuadra, nada más, de mi casa. Toda mi primaria. Mi secundaria lo realicé en el colegio San Lorenzo 7013, a dos cuadras de mi casa. O sea, prácticamente siempre..., mi mamá, como toda mamá sobreprotectora, siempre quiso buscar lo más cercano para que ella esté detrás mío, ¿no? Tratando de evitar de que pudiera suceder algo. Pero gracias a Dios... siempre he sido una persona muy independiente y esto, por ejemplo, se lo debo a mi profesora de primaria. Que ella

nunca dejó de que yo fuera una persona con discapacidad y me sentara... Por ejemplo, en transición, mi profesora Valentina me sentaba en un rinconcito, en una mesita, a la puerta, justo al lado de la puerta de salida, para que yo no caminara nada. Había participación para salir adelante y escribir en la pizarra, no me sacaba. Yo siempre decía «¿pero por qué no me saca?», ¿no? Cuando paso a primer grado, en esa época vendría a ser pues el primer año... y con la que llevé hasta quinto de primaria, fue con la profesora Julia Torres. Ella... cambió totalmente lo que yo estaba acostumbrada, ¿no? A sentarme en un rinconcito, a no participar, y todo. Buscando la comodidad, entre comillas. Ella no, ella, al contrario. Me hizo sentarme adelante, cerca a su escritorio, y que participara a salir a la pizarra. Inclusive los primeros días yo le reclamaba «¿Por qué me saca a la pizarra si mi profesora el año pasado no me sacaba?». Ya. «No, porque tú no tienes corona. Tú tienes que salir igual que tus compañeros salen a la pizarra». Ella fue la que fue motivándome, ¿no?, que yo saliera, participara. Es más, me hizo desfilar... para el aniversario de Surquillo, toda la avenida Angamos. Y me hizo que yo fuera detrás de la persona que llevaba la bandera. «Tú vas a ir a paso lento, pero vas a marchar con nosotros». Primera vez en mi vida que hice eso. Y allí todo era... era yo, ¿no? La profesora me decía «Hay que hacer esto», «Quiero que te aprendas esta poesía», «Quiero que cantes esta canción». O sea, siempre ella. «Pero profesora...», le digo, «¿me voy a tener que memorizar toda esa poesía?». «Sí». Y ella era la que me motivaba, me preparaba. Me tomaba la lección y me decía «No, acá tienes que decirlo de esta forma, acá tienes que levantar más la voz, acá tu pausa». Y fue por ella, porque mi mamá no quería ni dejarme que fuera a los paseos. Y ella venía, como estaba a una cuadra de mi casa, venía y le decía «Señora, ¿por qué no quiere que ella vaya de paseo?». «No», le decía, «porque no vaya a ser que se caiga». «¿Y, señora? Toda

la vida usted no va a estar para ella. Ella tiene que aprender a caminar por sí sola. Si se cae, ella tiene que aprender a levantarse. Usted tiene que pensar en el mañana. Ella va a estudiar, se va a superar, se va a casar, va a formar su hogar, y usted no va a estar pues pegada a ella para todo». Y mi mamá, yo me acuerdo, le escuchaba decir «No... mi hijita nunca se va a casar», «No va a tener hijos», decía.

«Señora», dice, «¡Por favor! Yo me la llevo y yo me voy a hacer cargo de ella». Y a partir de esa fecha mi mami, yo la veía ¿no? Porque ella iba y veía que subía al bus, y lloraba, ¿no? Un poco que yo también me iba llorando, ¿no? Pero después ya no... Secundaria, ya mi vida era distinta, ¿no? Ya con las chicas participaba más, ya para todo era, ¿no? Las mismas chicas «Vamos a organizar tal cosa, ya pues, tú habla con la profesora». Bueno. O tal cumpleaños se celebraba. «Habla con tu papá, que te dé chance en tu casa para que no... no te vayan a poner trabas y te digan "no vayas", ¿no?». Y para qué, mi papá, mi mamá, muy asequibles, entonces ellos dejaban que ellos vinieran a la casa.

Y todo eso, más... a un sacerdote, el padre Carlos, que en paz descance. Desde que yo comencé a participar en la Fraternidad Solidaria de Salud, él siempre también me motivaba a ir a los paseos, a organizar, me ponía a cargo a los más pequeños, ¿no? Y luego empecé a tener cargos dentro de este grupo y esto es a nivel internacional. En cada distrito hay, o bueno, había, porque ahora también se ha desintegrado prácticamente, ¿no? Hay pequeños núcleos en cada distrito. Y en esa época, pues sí había en todos los distritos. Yo siempre tuve un cargo dentro del distrito de Surquillo, y poco a poco fue escalando, ¿no? Luego fue a nivel zonal, después a nivel de Lima, y en el de Lima comencé a trabajar con los niños porque gané un concurso en el mes de septiembre por el día de la

primavera. Salí elegida reina, y la labor de la reina es trabajar todo un año con los niños. Todos los niños a nivel de Lima Metropolitana.

Y de allí yo siempre fui escalando. Me presenté a la universidad, a la nacional porque en esa época pues todo era la nacional, ¿no? Y yo quería ser una sanmarquina. No ingresé. Eh, para esto, mi... en esa época, era enamorado de mi hermana, de mi segunda hermana, porque yo soy la mayor. En esa época eran solamente, bueno, eran enamorados, ahora ya son esposos, ¿no? Él me decía «Si no ingresas acá, preséntate a la [universidad] Villaflor. Ingresas a la Villaflor, espera al otro año, y vuelves», me dice. «En el intento está la vencida». Dicho y hecho. Me presenté a la Villaflor. No ingreso. Esperé al siguiente año. San Lucas, tampoco. Me presento a la Villaflor y yo me presenté por presentarme. Justo había habido una de mis operaciones de corrección del pie, estaba con yeso, y yo dije «No me voy a presentar». Pero mi cuñado fue, compró la carpeta de postulante, me lo trajo y me dijo «Tal día es tu examen, tienes que ir tal día, llevar tus papeles y todo. Yo te voy a venir a recoger, le voy a decir a mi papá que me preste su carro y te llevo». Me llevó enyesada y todo. Fui a dar el examen. Y al día siguiente... yo ni siquiera ya tenía ganas ni de recibir los resultados en el periódico. Y mi cuñado vino temprano y yo escuché que le dijo a mi mamá «Señora, ¿y ella?». «Ella todavía está durmiendo», le dice. «No, pero que se levante que hay que celebrar. ¡Ingresó!».

¡Había ingresado! Ingresé a la Villaflor a Psicología. Terminé de estudiar. Bueno actualmente soy egresada, porque justo cuando estaba por finalizar, ya había estado en internado, salí embarazada, ¿no? Salí embarazada de mi hija que actualmente ya tiene 26 años. Y entonces ya no pude seguir, pero sí terminé mi internado. Luego con eso... dije «Más adelante», pero eso del «más adelante» nunca se

dio, ¿no? Porque empezaba a juntar, «Voy a sacar esto, voy a sacar lo otro», ya se presentaba otra cosa, tenía que coger ese dinero. Ya había más necesidades para mi hija. Entonces, pensando en ella, fue lo que, ya luego, lo dejé de lado, ¿no? Lo dejé de lado. Y justo, mi hija cumplió un año, y tuvo una... un germen que le ingresó a ella a su organismo. Y un poquito se me va...

Fuente: extracto de entrevista realizada por las estudiantes Andrea Chávez, Daniela Neyra, Lorena Pacora y Nicole Vilca en el Taller de Procesamiento de Datos de la PUCP (2022-1).

Nota: tanto el nombre de la entrevistada como las referencias a lugar y personas fueron modificados.

4.2. Técnica 1. Codificación predefinida

Para generar códigos deductivos, podemos tener la asistencia de insumos de ChatGPT a partir de los contenidos de nuestra guía de entrevista. Así, en vistas a formular códigos con el uso de ChatGPT, seguiremos los siguientes pasos.

Paso 1

Se inicia con el entrenamiento CORI-F¹⁰ (figura 4.1).

 **TÚ**

- Me encuentro desarrollando una investigación científica en sociología, para la cual tengo datos cualitativos (entrevistas a personas con discapacidad) que analizar para entender las experiencias educativas y laborales de cinco entrevistados peruanos.
- El objetivo específico consiste en comprender cómo recuerdan su experiencia educativa las personas con discapacidad entrevistadas.
- Para este proyecto vas a ser un analista especializado en datos cualitativos para brindar asistencia en la investigación.
- Para el análisis de los datos, debes usar únicamente la información que te suministraré a lo largo de la investigación.

 **ChatGPT**

iEntendido! Estoy aquí para ayudarte con el análisis de los datos cualitativos sobre las experiencias educativas de las personas con discapacidad en Perú. Por favor, proporciona la información relevante y cualquier contexto adicional que consideres necesario para comenzar el análisis.

 **TÚ**

¿Has entendido las indicaciones?

 **ChatGPT**

Sí, he entendido las indicaciones. Estoy listo para ayudarte con el análisis de los datos cualitativos de las entrevistas a personas con discapacidad en el contexto educativo peruano. Adelante, proporciona la información y detalles necesarios para comenzar el análisis.

Figura 4.1. Entrenamiento CORI-F para la codificación.

Paso 2

Una vez validado el primer paso, solicitamos a ChatGPT que proponga diez códigos a partir del contenido de la guía de entrevista, y señalando el OE, como se muestra en la figura 4.2. En cuanto al formato (la F de CORI-F), pedimos que cinco de los códigos sean descriptivos y que cinco sean conceptuales, y que se presenten en una tabla. Por último, especificamos que sean códigos de una sola palabra.

10 Ver detalles y explicación sobre el entrenamiento CORI-F en la sección 2.2.

P Tú

Con el objetivo específico de análisis (Comprender cómo recuerdan su experiencia educativa las personas con discapacidad entrevistada), genera 10 códigos, de los cuales 5 sean códigos descriptivos y 5 sean códigos conceptuales (de extensión corta, de máximo una palabra), a partir de las siguientes preguntas.

Presenta los códigos en una tabla de forma ordenada y clara.

- ¿Cómo ha sido su experiencia educativa?
- ¿Durante su educación primaria e inicial? ¿El entorno era accesible y acomodaba para su discapacidad?
- ¿Durante la educación secundaria? ¿El entorno era accesible y acomodaba para su discapacidad?
- ¿Me podría contar cómo fue su experiencia educativa después del colegio? ¿El entorno era accesible y acomodaba para su discapacidad?
- ¿Cómo ha sido su educación en otras materias no académicas (tal vez)? ¿El entorno era accesible y acomodaba para su discapacidad?
- ¿Cómo era su relación con sus compañeros? ¿fue fácil hacer amigos?

Figura 4.2. *Prompt* para la creación de códigos predefinidos en ChatGPT.¹¹

La figura 4.3 presenta la respuesta de ChatGPT al *prompt*.

ChatGPT

Aquí tienes una tabla con los códigos descriptivos y conceptuales generados a partir de las preguntas proporcionadas:

Códigos Descriptivos	Códigos Conceptuales
Experiencia Educativa	Accesibilidad
Ambiente Escolar	Inclusión
Experiencia Post-Escolar	Adaptación
Educación No Académica	Integración
Relación con Compañeros	Socialización

Figura 4.3. Resultados de la codificación predefinida en ChatGPT.

La lista de diez códigos elaborada por ChatGPT puede ayudarnos mucho en nuestra codificación predefinida derivada del OE1 de investigación. Sin duda, como se ha dicho, es importante verificar la coherencia de los códigos, así como la pertinencia para su inclusión en nuestras listas de los mismos. Estos insumos generados

11 El discurso ha sido recortado por motivos de espacio.

por la herramienta pueden servir para enriquecer o evaluar nuestra formulación de códigos mediante un razonamiento deductivo.

4.3. Técnica 2. Codificación emergente

Para generar códigos inductivos, debemos brindar a ChatGPT los contenidos de nuestros casos de forma ordenada. Para ilustrar este procedimiento, usamos un extracto de la entrevista realizada a una persona con discapacidad, en la que ella narra sus vivencias durante su etapa escolar (recuadro 4.2.).

Tras el entrenamiento CORI-F, que debe ser realizado al inicio del trabajo, se hace lo siguiente.

Requerir a ChatGPT la creación de cinco códigos basados en los contenidos del fragmento de entrevista que concierne al OE1, como se muestra en la figura 4.4. En cuanto al formato (la F de CORI-F), pedimos adicionalmente que tres códigos sean descriptivos y que dos sean conceptuales, que tengan como máximo una palabra y que los resultados se presenten en una tabla.

P You

Con el objetivo específico de análisis (Comprender cómo recuerdan su experiencia educativa las personas con discapacidad entrevistada), genera 5 códigos, de los cuales 3 sean descriptivos y 2 sean códigos conceptuales (de extensión corta, máximo de una palabra), a partir del siguiente fragmento de entrevista.

Presenta los códigos en una tabla de forma ordenada y clara.

"Yo, primaria, lo realicé en el colegio 463 que quedaba, bueno que queda, ahora a dos cuadras y media, pero antes quedaba a una cuadra, nada más, de mi casa. Toda mi primaria. Mi secundaria lo realicé en el colegio San Lorenzo 7013 a dos cuadras de mi casa. O sea, prácticamente, siempre...mi mamá, como toda mamá sobreprotectora siempre quiso buscar, lo más cercano para

Figura 4.4. *Prompt* para realizar la codificación emergente.¹²

12 El discurso ha sido recortado por motivos de espacio.

A continuación, la figura 4.5 presenta la respuesta de ChatGPT al *prompt*. Observamos allí que los códigos descriptivos son «Experiencia escolar en primaria», «Influencia positiva de profesores» y «Apoyo familiar y resistencia»; y los códigos conceptuales, «Independencia» y «Superación».

 ChatGPT

Claro, a partir del fragmento proporcionado, he generado cinco códigos, tres descriptivos y dos conceptuales. Aquí están presentados en una tabla:

Nombre	Tipo	Descripción
Experiencia Escolar en Primaria	Descriptivo	La narrativa de la educación primaria, destacando cambios en la ubicación y enfoques de enseñanza.
Influencia Positiva de Profesores	Descriptivo	El papel de la profesora Julia Torres en cambiar la perspectiva y fomentar la participación activa.
Apoyo Familiar y Resistencia	Descriptivo	La sobreprotección inicial de la madre contrastada con el apoyo gradual a la independencia escolar.
Independencia	Conceptual	El énfasis en la independencia personal, aprendizaje para levantarse y superar obstáculos.
Superación	Conceptual	La historia de perseverancia y éxito en la educación superior, enfrentando desafíos personales.

Estos códigos resumen aspectos clave de la experiencia educativa de la entrevistada y pueden servir como puntos de partida para el análisis. Si necesitas ajustes o más detalles, estoy aquí para ayudarte.

Figura 4.5. Resultados de la codificación emergente en ChatGPT.

Los códigos emergentes obtenidos por medio de ChatGPT pueden ser útiles para encontrar palabras clave que no se habían considerado previamente y ser buenas alternativas para comprender lo que nos dicen los contenidos. Esto enriquecería el trabajo de codificación emergente manual que hayamos realizado. No obstante, es de suma importancia la pertinencia de los códigos propuestos por la herramienta y que sus explicaciones sean coherentes. Estos insumos también pueden ser de gran utilidad para mejorar nuestra codificación emergente descriptiva y conceptual con el propósito de establecer patrones en nuestros contenidos.

Consejo 6. El uso de formatos ordenados para la codificación

En el marco de la codificación asistida por ChatGPT, se recomienda que los *prompts* incluyan la generación de tablas y que haya un *prompt* acerca de la extensión de los códigos. Las ventajas de hacer esto son múltiples:

Un formato ordenado ayuda al modelo a entender mejor la estructura de la información y a generar respuestas más precisas.

En procedimientos de abstracción, como la codificación, las tablas evitan la generación de respuestas dispersas o ambiguas y mejoran la calidad general del resultado.

Delimitar el número de palabras a una cantidad acodada permite que se refuerce la necesidad de establecer una abstracción a partir de los contenidos. Para el caso de los códigos conceptuales, se recomienda solicitarlos con un máximo de dos palabras.

A continuación, la figura 4.6 ilustra la especificación para el formato.

 Tú

Con el objetivo específico de análisis (Comprender cómo recuerdan su experiencia educativa las personas con discapacidad entrevistada), genera 5 códigos, de los cuales 3 sean descriptivos y 2 sean códigos conceptuales (de extensión corta, máximo de una palabra), a partir del siguiente fragmento de entrevista.

Presenta los códigos en una tabla de forma ordenada y clara.

Figura 4.6. Indicar formatos para la codificación inicial en ChatGPT.

4.4. Evaluación crítica de los resultados

En esta última sección del capítulo, se presenta una evaluación acerca de los resultados que nos ha proporcionado ChatGPT a partir de un extracto de entrevista del proyecto «Trayectorias laborales y educativas de personas con discapacidad». Las capacidades de la IA se encuentran en constante evolución, por lo que resulta importante reservar un espacio para reflexionar críticamente acerca de la calidad de los resultados que brinda, lo que nos permite identificar patrones –de eficacia e ineficacia– en el trabajo con ChatGPT. Esta evaluación no es exhaustiva y busca principalmente mostrar una forma de mantener una mirada crítica a cada procedimiento realizado con IA.

La tabla 4.2 presenta, para las dos técnicas aplicadas, un ejemplo de la evaluación realizada a los insumos generados por ChatGPT en términos de fortalezas, debilidades y eficacia. Las fortalezas describen el valor que ha aportado ChatGPT. En este caso, obtuvimos códigos interesantes y pertinentes en una fracción de segundo, tanto con la codificación emergente como con la codificación predefinida. Asimismo, se lograron buenos resultados en ambos procedimientos al solicitar códigos descriptivos y códigos conceptuales. En cuanto a las debilidades, estas rinden cuenta de las limitaciones identificadas al realizar las dos técnicas de codificación, que se vinculan principalmente con la falta de entendimiento contextual y con el idioma, así como con propuestas muy generales de códigos en algunos casos. Por último, la eficacia es un puntaje que valora los resultados de las técnicas de exploración sobre la base de los resultados, puntaje que en ambos casos fue bastante aceptable.

Tabla 4.2. Fortalezas y debilidades identificadas en el proceso de codificación inicial con ChatGPT

Procedimiento	Fortalezas	Debilidades	Eficacia
Codificación predefinida	<ul style="list-style-type: none"> - Puede procesar preguntas de investigación, objetivos y/o guías de recolección de datos de manera rápida y consistente. - Puede proponer códigos útiles para la clasificación de contenidos y la identificación de patrones y conceptos previos al análisis de contenido. 	<ul style="list-style-type: none"> - Puede no tomar en consideración ciertas preguntas o repreguntas, debido a su formulación ambigua o excesivamente concisa. - Pueden generarse códigos muy generales, que no se ajustan completamente a la especificidad requerida para el análisis. 	4/5
Codificación emergente	<ul style="list-style-type: none"> - Puede procesar grandes volúmenes de texto⁽¹⁾ de manera rápida y consistente. - Puede identificar de forma eficaz patrones, temas o conceptos emergentes que podrían pasar desapercibidos en enfoques más estructurados. 	<ul style="list-style-type: none"> - La identificación de patrones puede no ser tan profunda en cuestión de significados subyacentes e interpretaciones contextuales, en particular en castellano del Perú. - Pueden generarse códigos conceptuales que no son lo suficientemente precisos. 	4/5

Nota: ⁽¹⁾ La versión gratuita de ChatGPT tiene un límite de palabras, por lo que se puede trabajar por bloques de contenido. Por su parte, la versión pagada no tiene límite de palabras y soporta documentos (no se requiere pegar los contenidos en la ventana de diálogo).

Consejo 7. Haz una retroalimentación a ChatGPT

Es importante hacer una retroalimentación por escrito a ChatGPT, con el *prompt* de mejora, lo que ayudará a que se elabore una nueva versión ajustada de los códigos propuestos. Para ello, puedes detenerte un momento y considerar qué fortalezas y debilidades has encontrado en ChatGPT.

En relación a lo anterior, y como se ha visto en los consejos 2 y 4, es oportuno utilizar los pulgares y la opción *regenerate* para evaluar los resultados y mejorar futuras elaboraciones; asimismo, respecto a los consejos 1 y 5, que ayudan a obtener mejores resultados, es importante recordar el OE y realizar *prompts* que tengan una estructura ordenada.

CIERRE

Cinco consejos para ser más creativos, eficaces y responsables al usar inteligencia artificial en la investigación social y educativa

1. Te estás aventurando en procedimientos novedosos, con una herramienta nueva. Por ello, asegúrate de hacerlo a partir de un diseño de investigación consistente y bien delimitado. Esto debe guiar la toma de decisiones en cada procedimiento y cada decisión metodológica. En particular, es muy útil que tengas permanentemente en consideración las preguntas y/o los objetivos específicos en el tratamiento y el análisis de los datos.
2. Para muchas personas, el uso de ChatGPT es poco conocido e incluso, en algunos círculos, polémico. Por ello, antes de empezar a usar esta herramienta, tómate el tiempo de conocer los lineamientos éticos para su uso en la institución o instituciones con las que tienes un compromiso laboral o de otro tipo. A partir de esto, sigue los lineamientos y políticas de tu organización para el uso de IA en la creación de conocimiento y/o en el análisis de datos. Del mismo modo, es conveniente transparentar el uso de IA cuando presentes las herramientas usadas en el análisis de datos cualitativos.
3. La meta de la investigación social y educativa es profundizar en el entendimiento de contextos específicos y construir conocimiento e interpretación difícil de obtener a simple vista. La es-

trategia Aexco, propuesta en esta guía, te permitirá vincular tus anotaciones, exploración de datos y codificación para generar sinergias y reflexiones novedosas sobre tus datos. En particular, es importante que apliques las seis técnicas de exploración y codificación, ya que ellas te permitirán aprovechar la riqueza de los datos en beneficio de tu proyecto de investigación.

4. Antes de empezar a trabajar con ChatGPT, asegúrate de conocer bien los contenidos de los casos que estás estudiando. Del mismo modo, antes de compartir cualquier tipo de información con esta herramienta, asegúrate de proteger la identidad de personas y entidades a través de un proceso minucioso de anonimización. Hecho esto, realiza una buena configuración de ChatGPT a través del entrenamiento CORI-F, lo que te permitirá establecer un entorno de trabajo adaptado a tu proyecto de investigación. Recuerda formular preguntas para confirmar el buen entendimiento y brinda retroalimentación constante a través de los consejos que hemos dado en los capítulos 2, 3 y 4.
5. Por último, considera a ChatGPT como un asistente de investigación que genera insumos útiles para ampliar tu creatividad y perspicacia con respecto a los datos. En este sentido, recuerda siempre supervisar y evaluar con detenimiento las respuestas y los insumos generados por esta herramienta, con el objetivo de evitar sesgos de programación, errores o interpretaciones descontextualizadas. De ninguna manera consideres que la IA puede sustituir tu rol de investigador o ser exclusivamente una máquina para ahorrar tiempo. Ten en cuenta, más bien, que el potencial de la IA en investigación social y educativa consiste principalmente en potenciar tus capacidades de reflexión, análisis e innovación.

REFERENCIAS BIBLIOGRÁFICAS

- Anis, S. y French, J. A. (2023). Efficient, explicatory, and equitable: Why qualitative researchers should embrace AI, but cautiously. *Business and Society*, 62(6), 1139-1144. <https://doi.org/10.1177/00076503231163286>
- Birks, M., Chapman, Y. y Francis, K. (2008). Memoing in qualitative research: Probing data and processes. *Journal of Research in Nursing*, 13(1), 68-75. <https://doi.org/10.1177/1744987107081254>
- Charmaz, K. (2014). *Constructing grounded theory*. Sage.
- Comisión Europea. (2020). *Libro blanco sobre la inteligencia artificial: un enfoque europeo orientado a la excelencia y la confianza*. Comisión Europea. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_es
- Committee on Publication Ethics, COPE. (2020). *Authorship and AI tools*. COPE. <https://publicationethics.org/cope-position-statements/ai-author>
- Congreso de la República del Perú. (2023). Discurso de investidura del presidente Ollanta Humala. Congreso de la República del Perú. <https://www.congreso.gob.pe/>
- Creswell, J. W. y Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.
- Elder, G. H., y Giele, J. Z. (2009). *The craft of life course research*. Guilford Press.

- Etesse, M. (2019). «Transiciones postescolares y desigualdades sociales en Lima, Perú. Continuidad educativa e inserción profesional de los jóvenes de la generación "Ochentas"» (doctorado). Université Paris Descartes.
- Firat, M. (2023). *How chat GPT can transform autodidactic experiences and open education*. Anadolu University-Department of Distance Education-Open Education Faculty.
- Flick, U. (2014). *An introduction to qualitative research* (5a ed.). Sage.
- Ghosh, S. y Caliskan, A. (2023). *ChatGPT perpetuates gender bias in machine translation and ignores non-gendered pronouns: Findings across Bengali and five other low-resource languages*. Cornell University. arXiv preprint arXiv:2305.10510
- Gibbs, G. (2008). *Analysing qualitative data*. Sage Publications.
- Glaser, B. y Strauss, A. (2006 [1967]). *The discovery of grounded theory*. Aldine Press. http://www.sxf.uevora.pt/wp-content/uploads/2013/03/Glaser_1967.pdf
- Hartmann, J., Schwenzow, J. y Witte, M. (2023). *The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation*. Cornell University. arXiv preprint arXiv:2301.01768
- Hoyl, M. (2023). *El ABC de la IA. Tu guía de ChatGPT para aprender y mantenerte al día*. Notion. <https://hoyl.notion.site/Ingenier-a-de-preguntas-d-ndome-a-entender-con-ChatGPT-ba99e04123c34fef8d436a05aad40961>
- Lopezosa, C., & Codina, L. (2023). *ChatGPT y software Caqdas para el análisis cualitativo de entrevistas: pasos para combinar la inteligencia artificial de OpenAI con Atlas.ti, Nvivo y Maxqda*. Universitat Pompeu Fabra y DigiDoc. https://repositori.upf.edu/bitstream/handle/10230/55477/Codina_dig_chat.pdf?sequence=1&isAllowed=y

- Lopezosa, C., Codina, L. y Boté-Vericad, J. J. (2023). *Testeando Atlas.TI con OpenAI: hacia un nuevo paradigma para el análisis cualitativo de entrevistas con inteligencia artificial*. Universitat Pompeu Fabra y DigiDoc. https://repositori.upf.edu/bitstream/handle/10230/56449/Lopezosa_digidoc_atlas.pdf?sequence=1&isAllowed=y
- OpenAI. (2023). *Presentamos ChatGPT*. OpenAI. <https://openai.com/blog/chatgpt/>
- Pontificia Universidad Católica del Perú, PUCP. (2016). *Comité de Ética de la Investigación de la Pontificia Universidad Católica del Perú. Reglamento*. PUCP-Vicerrectorado de Investigación. <https://cdn02.pucp.education/investigacion/2016/10/14160435/Reglamento-2.pdf>
- Rahman, M., Terano, H. J. R., Rahman, N., Salamzadeh, A. y Rahaman, S. (2023). ChatGPT and academic research: A review and recommendations based on practical examples. *Journal of Education, Management and Development Studies*, 3(1), 1-12. <https://doi.org/10.52631/jemds.v3i1.175>
- Richards, L. (2015). *Handling qualitative data. A practical guide*. Sage.
- Rutinowski, J., Franke, S., Endendyk, J., Dormuth, I. y Pauly, M. (2023). *The self perception and political biases of ChatGPT*. Cornell University. arXiv preprint arXiv:2304.07333
- Sabzalieva, E. y Valentini, A. (2023). *ChatGPT e inteligencia artificial en la educación superior: Guía de inicio rápido*. Unesco. https://unesdoc.unesco.org/ark:/48223/pf0000385146_spa
- Saldaña, J. (2021). *The coding manual for qualitative researchers*. Sage.
- Strauss, A. L. y Corbin, J. M. (2002). *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundada*. Editorial Universidad de Antioquia-Facultad de Enfermería.

Unesco. (2021). *Recomendación sobre la ética de la inteligencia artificial*.

Unesco. <https://www.unesco.org/es/legal-affairs/recommendation-ethics-artificial-intelligence>

Unesco. (2023). *Oportunidades y desafíos de la era de la inteligencia artificial para la educación superior: una introducción para los actores de la educación superior*. Unesco. https://unesdoc.unesco.org/ark:/48223/pf0000386670_spa

Vargas-Bianchi, L. (2023). *IA en el análisis cualitativo de datos: utilidades y limitaciones*. Thma. https://thma.substack.com/p/ia-en-el-analisis-cualitativo-de?r=1eogxo&utm_campaign=post&utm_medium=web

Xames, M. D. y Shefa, J. (2023). ChatGPT for research and publication: Opportunities and challenges. *Journal of Applied Learning and Teaching*, 6(1), 390-395. <https://doi.org/10.37074/jalt.2023.6.1.20>

Zhang, H., Wu, C., Xie, J., Kim, C. y Carroll, J. M. (2023). *QualiGPT: GPT as an easy-to-use tool for qualitative coding*. Cornell University. arXiv preprint arXiv: 2310.07061

SIGLAS

Aexco	Anotación, exploración y codificación
Cisepa	Centro de Investigaciones Sociológicas, Económicas, Políticas y Antropológicas
COPE	Committee on Publication Ethics
CORI-F	Contexto, rol, información y formato
IA	Inteligencia artificial
LLM	Lenguaje natural a gran escala (por sus siglas en inglés)
OE	Objetivo específico
PUCP	Pontificia Universidad Católica del Perú
Unesco	Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura

Manuel Etesse

El autor es doctor en Sociología por la Université Paris Cité Sorbonne (Francia), magíster en Población y Desarrollo por la misma casa de estudios y licenciado en Sociología por la Pontificia Universidad Católica del Perú. En la PUCP es docente a tiempo parcial de la Facultad de Ciencias Sociales desde el año 2016 y profesor a tiempo completo del Departamento de Educación desde 2024. Se ha desempeñado en múltiples organizaciones públicas y privadas vinculadas al desarrollo, la educación y la juventud en Perú y Francia. Sus temas de interés son educación y sociedad, juventudes y población, tecnología y métodos de investigación social y educativa.

Contacto: manuel.etesse@pucp.edu.pe

 <https://www.linkedin.com/in/manuel-etesse>

 <https://www.researchgate.net/profile/Manuel-Etesse>

 <https://orcid.org/0000-0002-1521-5405>

ISBN: 978-612-4355-17-2



9 786124 355172

DOI: <https://doi.org/10.18800/978-612-4355-17-2>