Recuperando infancias desaparecidas. Desafíos para analizar los avisos de fuga desde las humanidades digitales y la ciencia de datos[1]

Gabriela Arriagada Bruneau[2], María Catalina Cárdenas Muñoz[3], Nicole Larrondo[4], José Ragas[5], Sofía Roldán Bugueño[6], Diego Salazar Carmona[7], Matías Valenzuela Lucero[8] , Montserrat Santis Aliste[9]

Introducción

El avance de las humanidades digitales y la ciencia de datos —junto con instrumentos como la inteligencia artificial— ha significado un doble desafío para los archivos históricos. De un

[1] El análisis preliminar de este ensayo fue presentado inicialmente en el Primer Seminario de Humanidades Digitales (9–10 de abril de 2026), organizado por el Laboratorio de Humanidades Digitales de la Pontificia Universidad Católica de Chile. De igual manera, el ensayo forma parte del Proyecto Interdisciplina II202420 titulado «Se ha perdido un cholito: construcción de categorías biométricas a partir de los avisos de fuga. Un análisis desde la ciencia de datos y las humanidades digitales», patrocinado por la Dirección de Investigación de la Pontificia Universidad Católica de Chile.

[2] Profesora Asistente del Instituto de Éticas Aplicadas y el Instituto de Ingeniería Matemática y Computacional, Pontificia Universidad Católica de Chile. gcarriagada@uc.cl

[3] Estudiante de Estadística, Pontificia Universidad Católica de Chile. mccardenas1@uc.cl

[4] Coordinadora del Laboratorio de Patrimonio Documental y Humanidades Digitales y del Laboratorio de Historia Digital, Pontificia Universidad Católica de Chile. nilarron@uc.cl

[5] Profesor Asistente del Instituto de Historia, Pontificia Universidad Católica de Chile. jose.ragas@uc.cl

[6] Estudiante del Magíster en Historia, Pontificia Universidad Católica de Chile. smroldan@uc.cl

[7] Licenciado en Historia, Pontificia Universidad Católica de Chile. diegosalazar25@gmail.com

[8] Estudiante de Licenciatura de Ciencia de Datos, Pontificia Universidad Católica de Chile. mvalenzuelu@uc.cl

[9] Estudiante de Licenciatura de Historia, Pontificia Universidad Católica de Chile. mjsantis@uc.cl

lado, ha permitido su preservación —al permitir la digitalización de estos a partir de soportes analógicos en riesgo de daño o pérdida— así como su divulgación y accesibilidad al estar disponibles en plataformas de acceso abierto para consultar y descargar tanto por parte del público general como de investigadores/as. De otro lado, el procesamiento de estos archivos ha planteado una serie de temas éticos que en ocasiones han sido soslayados por los aspectos técnicos de la digitalización. Así, la enorme masa documental disponible en la nube y nuestras pantallas ha llevado a «aplanar» la información existente con propósitos de análisis, reduciendo con ello (acaso inadvertidamente) la densidad del proceso social y político de los actores que aparecen en los documentos originales o también reduciendo el proceso tan solo al aspecto técnico.

Este ensayo busca continuar esta discusión. Si bien reconocemos los importantes avances conseguidos en los últimos años desde las humanidades digitales y la ciencia de datos al preservar y ampliar el conocimiento en torno a la información histórica, también consideramos que es necesario enfatizar el componente humano y ético tanto en los documentos originales como en el proceso de digitalización y análisis. Recientemente, dos grupos de investigación han llamado la atención respecto de este componente ético al momento de digitalizar documentos provenientes de procesos de violencia como lo fue el tráfico transatlántico de esclavos (Bloch et al., 2025) y la Operación Cóndor en América del Sur (Smart et al., 2026). En ambos casos, la reflexión se orienta hacia la interacción que se produce entre herramientas digitales y narrativas históricas, donde los actores (víctimas de tráfico humano y de desapariciones forzadas) bien pueden ser empoderados por las nuevas tecnologías o también revictimizados.

Por ello, este ensayo plantea que un campo emergente como la analítica cultural (entendida como un espacio de encuentro entre métodos cualitativos y computacionales) debe ser un ante todo un espacio crítico, donde los mecanismos de poder en los que se encuentran insertos los actores de los documentos no pierdan relevancia durante y después del proceso de digitalización. Al igual que los ejemplos anteriores, nuestro ensayo estudia a un grupo de

actores víctimas de sistemas racializados y de segregación, como lo fueron los niños y niñas de origen indígena conocidos como *cholitos* y que formaron parte del servicio doméstico en la ciudad de Lima en el siglo XIX.

La primera parte de este ensayo explica el contexto de estos actores y su desafío al sistema de servidumbre al que fueron forzados a pertenecer, lo que generó un fenómeno tan fascinante como perturbador: los avisos de fuga. La segunda sección establece un análisis a partir de avisos entre 1839 y 1866, digitalizados, para establecer las variables y categorías que se forman. La tercera y última sección profundiza en este análisis, recordando la importancia de comprender los datos como información situada y como una invitación para plantear nuevas preguntas.

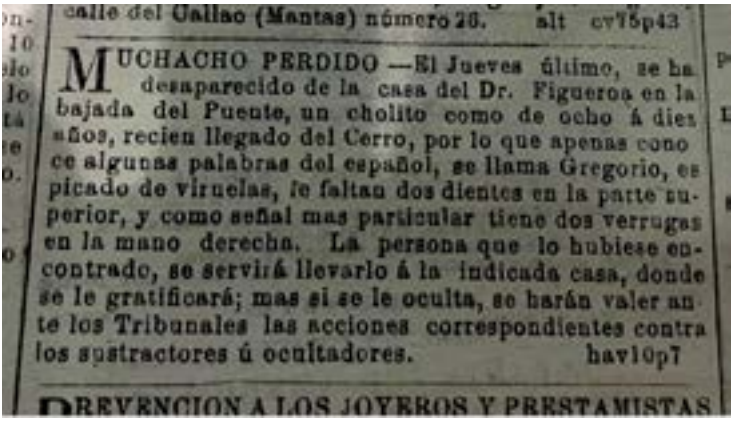
Niños fugitivos y avisos de fuga

A mediados del siglo XIX, Perú experimentó un proceso migratorio bastante inusual. Las medidas liberales instauradas como parte de una ola global de reformismo político y social significaron el fin de la esclavitud de origen africano y el fin del tributo sobre las comunidades indígenas. Esta última medida dio un impulso para que cientos y miles de personas se desplazaran lentamente desde el interior del país hacia la costa y, en especial, hacia la capital, Lima. Para muchos de ellos, se trataba de su primera experiencia fuera de su entorno local. Con el tiempo, la migración de los Andes a la capital se convirtió en una de las transformaciones más importantes del país, con consecuencias que llegan hasta el día de hoy (Ragas, 2025).

El perfil de los migrantes que se aventuraron a llegar a la capital en el siglo XIX era muy variado, pero se concentró principalmente en los sectores de mediana edad y que podían ofrecer fuerza laboral en la ciudad. La ausencia de un sector industrial impidió que esta mano de obra pudiese ser absorbida y convertirse en un sector proletario o de trabajadores. Como consecuencia, un segmento importante de los recién llegados pasó a convertirse en sirvientes domésticos al interior

de los domicilios de la clase media y alta de limeños. Dentro de ellos, destaca un sector en especial, conformado por menores de edad, que ha sido tema de investigación en los últimos años (Flores Galindo, 1994; Ragas, 2017; Peluffo, 2022; Velásquez, 2024; Roldán, 2026).

Los *cholitos*, como fueron conocidos, eran niños y niñas indígenas provenientes de la sierra peruana. De acuerdo con diversas fuentes y testimonios, su edad oscilaba entre los cinco y catorce años y formaban parte de un mercado informal y subterráneo en el que, con frecuencia, eran raptados de sus familias por efectivos militares y trasladados a la capital para trabajar como reemplazo de la mano de obra esclava (Roldán, 2026). Durante el periodo del guano (1845–1873) su presencia se hizo más visible. Lo trágico y paradójico de este caso es que conocemos sobre ellos especialmente a través de testimonios indirectos, como son los avisos de fuga. Cuando alguno de los *cholitos* tomaba finalmente la decisión de huir de la casa de sus amos, estos acudían a las oficinas del diario *El Comercio*, el más importante de entonces, para publicar un aviso, denunciando el hecho y pidiendo su captura y devolución lo antes posible (Imagen 1).



calle del Callao (Mantas) número 26. alt cv15p43

MUCHACHO PERDIDO.—El Jueves último, se ha desaparecido de la casa del Dr. Figueroa en la bajada del Puente, un cholito como de ocho á diez años, recien llegado del Cerro, por lo que apenas conoce algunas palabras del español, se llama Gregorio, es picado de viruelas, le faltan dos dientes en la parte superior, y como señal mas particular tiene dos verrugas en la mano derecha. La persona que lo hubiese encontrado, se servirá llevarlo á la indicada casa, donde se le gratificará; mas si se le oculta, se harán valer ante los Tribunales las acciones correspondientes contra los sustractores ú ocultadores. hav10p7

PREVENCION A LOS JOYEROS Y PRESTAMISTAS

Figura 1. Aviso de fuga de cholito publicado en *El Comercio*.

En los periódicos limeños del siglo XIX, como *El Comercio*, era común encontrar una sección de avisos del día, generalmente, ubicada en las últimas páginas. Los avisos de fuga formaban parte de un espacio escrito e impreso relacionado con la compra y venta de esclavos y la transición hacia un mercado de mano de obra. Estos anuncios informaban sobre objetos y animales perdidos, sin embargo, también incluían avisos de fuga de sirvientes esclavos, afrodescendientes y de culíes. Esto da cuenta de la amplia variedad de avisos publicados, que no solo referían a bienes materiales, sino también a personas que frecuentemente huían de los hogares de sus patrones, en muchos casos debido a sus condiciones de vida, en busca de nuevas oportunidades laborales y de una eventual movilidad social. La esclavitud y la servidumbre eran temas recurrentes en la prensa local, donde junto a los anuncios clasificados se publicaba la opinión de quienes legitimaban el uso de la violencia hacia los sirvientes domésticos, apoyado de reglamentos policiales vigentes de la época, junto con la de quienes buscaban eliminar dichos abusos.

Los avisos de fuga eran textos escritos, breves, sin imagen y cuya publicación y frecuencia de aparición dependían de factores diversos, desde el precio (ya que su costo variaba según la extensión y número de reimpresiones) hasta la valoración o interés por parte del patrón o patrona para recuperar al sirviente huido. Así, el nivel de detalle en la descripción, e incluso la reiteración del anuncio, permiten inferir el grado de interés en la recuperación del menor, lo que sugiere que el valor de los niños dentro del espacio doméstico no era homogéneo. De este modo, los avisos estaban destinados a que el público de Lima proporcionara información sobre el paradero de los cholitos. Para facilitar este propósito, los patrones describían a los sirvientes mediante rasgos físicos tales como la edad, el color de piel, la complexión, la estatura, entre otras. Estas descripciones constituyen, a su vez, la base de información a partir de la cual es posible analizar las características de estos menores indígenas.

Como veremos en la segunda sección, la información que los avisos ofrecen es muy rica. Más allá de la información potencialmente cuantificable que ofrecen, estos actúan como reflejo de la sociedad peruana y latinoamericana de mediados del siglo XIX, cuya herencia

del periodo colonial se mantenía aún vigente a través de lógicas de dominación y explotación. Así, la circulación forzada de niños indígenas hacia Lima no solo era conocida sino hasta aceptada y promovida como algo natural, que legitimaba un determinado orden social republicano (Ragas, 2017). Esto sugiere que no se trataba de prácticas excepcionales, sino de dinámicas profundas de la vida cotidiana, en una sociedad que tendía a ignorar el problema del tráfico infantil indígena. De igual manera, la exigencia de devolución de los menores sugiere que las lógicas laborales y domésticas los concebían como propiedad de los patrones, una percepción que continuó hasta bien entrado el siglo XX.

Un aspecto relevante del análisis de los avisos es la dificultad de conocer qué ocurría después de su publicación. La información que podemos conocer a partir de los avisos es compleja y muy sugerente para comprender la transición de estructuras sociales, económicas y raciales de la colonia a la república, y cómo estas permearon la manera actual en la que entendemos la igualdad de las personas y sus derechos, es decir, la democracia. No obstante, junto con los análisis que podamos aplicar a los avisos, es necesario considerar que estos no son la única forma de estudiar a estos actores ni a las estructuras que permitieron y sostuvieron su explotación por varias décadas. Las siguientes secciones constituyen un intento por comprender qué podemos aprender de los *cholitos*, sus decisiones de dejar atrás un régimen de abuso y los intentos desesperados de sus amos por traerlos de vuelta y sostener un régimen a todas luces injusto e inhumano.

Los avisos de fuga como infraestructura de datos

La formalidad estereotipada del aviso de fuga no es un accidente retórico, sino la marca de un género textual altamente codificado. El análisis de tópicos aplicado al corpus mediante LDA identificó cinco dimensiones narrativas que se articulan sistemáticamente en cada aviso: el marco legal e institucional (la obligación del patrón, el rol de la policía), la circulación editorial en *El Comercio*, la transacción económica de la gratificación, la

descripción facial y de la indumentaria. Las partes involucradas en la redacción de los avisos no escribían libremente. El contenido y extensión de los avisos dependía de factores tan variados como el precio por palabra, la información que los dueños de casa desearan proporcionar y la que recordaban. Así, los avisos seguían un protocolo implícito, donde una plantilla mental en función de la información mínima y necesaria determinaba qué información, qué orden y con qué propósito impresores, dueños de casa y público podrían responder al aviso. Que estas cinco dimensiones emerjan del modelado computacional con un valor de coherencia de 0,795[10] confirma lo que una lectura atenta de los avisos ya sugería.

Las cuarenta variables con las que el equipo codificó el corpus no fueron impuestas desde fuera sobre los avisos, sino que emergieron desde ellos. Los propios avisos organizaban la información en dimensiones recurrentes y reconocibles. En realidad, lo que dicha infraestructura producía no eran descripciones neutrales de personas, sino categorías de clasificación. Las variables registradas y omitidas revelan una jerarquía implícita sobre cuáles aspectos de la persona constituían datos relevantes para identificar a alguien en el espacio urbano de Lima. Que la edad aparezca en más del 90% de los registros y la frente en el 3–4% no es ruido estadístico. Es la huella cuantificable de una economía de la atención que el costo de la publicación, la urgencia del dueño de casa y las convenciones perceptuales históricas moldeaban, privilegiando ciertos rasgos del cuerpo sobre otros.

El color de piel es la segunda variable con mayor cobertura del corpus, entre el 68% y el 76% según la década y su valor más frecuente (*cholos/chinas*) concentra el 67,9% de los registros. Las

[10] El valor de coherencia mide qué tan interpretables y semánticamente cohesionados son los tópicos que produce el modelo: va de 0 a 1, donde valores más altos indican que las palabras de cada tópico tienden a aparecer juntas en los documentos y tienen sentido como conjunto. Para determinar el número óptimo de tópicos, el equipo probó configuraciones entre k=4 y k=15. El valor de 0,795 correspondiente a k=5 es el más alto de toda la serie: los modelos con 6 y 7 tópicos obtienen valores de aproximadamente 0,78, y a partir de k=8 la coherencia desciende con fluctuaciones hasta valores cercanos a 0,70. Esto indica que la división en cinco dimensiones narrativas no es una decisión arbitraria del equipo, sino la que mejor captura la estructura temática del corpus.

palabras *cholo* o *china* no necesariamente identifican a una persona, sino que la sitúan en un estrato social y racial. Lo que el aviso presenta como dato biométrico opera, asimismo, como marcador colectivo de clase-raza. Como argumentan D'Ignazio y Klein (2020), los datos nunca son neutrales, sino que reflejan las relaciones de poder del contexto en que fueron producidos y su análisis crítico exige examinar quién los recopila, sobre quién y con qué propósito. En esta línea, Benjamin (2019) muestra que las tecnologías de clasificación heredan y amplifican categorías raciales históricamente específicas, produciendo formas de discriminación que se presentan como objetivas y neutrales al estar codificadas. No son racistas por accidente, sino por diseño. De este modo, el corpus de avisos de fuga permite rastrear ese proceso en su origen, pues las categorías raciales que los patrones limeños usaban para identificar a un niño anticipan directamente las clasificaciones que décadas después serían incorporadas a sistemas de identificación formal.

La cobertura desigual de las variables no es solo un problema metodológico para el análisis computacional, sino, en sí misma, un hallazgo histórico. Que la frente aparezca en el 3–4% de los registros, el peinado en el 1–7% y la edad en más del 90% no refleja características objetivas de los niños fugados. Refleja decisiones, conscientes o no, de quienes tenían el poder de escribir el aviso. Lo que falta en el registro documenta qué aspectos del cuerpo de un *cholito* consideraban prescindibles los patrones limeños al momento de reconocerlo en la calle. Leer el corpus desde esta perspectiva implica tratar los valores nulos no como deficiencias de la fuente, sino como evidencia de jerarquías perceptuales. Burton (2005) propone precisamente este giro epistemológico, según el cual los archivos son construcciones activas y no depósitos pasivos de hechos. Así, lo que no entra en ellos constituye una decisión histórica tan cargada de poder como lo que sí lo hace. Stoler (2009) complementa ese argumento al proponer leer el archivo siguiendo su propia lógica interna, es decir, tratar las categorías perceptuales y las convenciones de quienes lo produjeron como objeto de análisis en sí mismo. Así, la mirada de los patrones limeños refleja qué variables se registran y cuáles no. Es lo que Johanna Drucker (2011) llama *capta* en lugar de

data. Es la información capturada, no simplemente recogida y que, por tanto, lleva inscrita la perspectiva de quien la capturó[11].

Esa perspectiva tenía un precio literal. El análisis de cobertura según la frecuencia de publicación revela una relación inversa entre cuántas veces se publicó un aviso y cuán detallada era su descripción; los avisos publicados más de tres veces tienen una cobertura de color de piel del 39%, frente al 58% de los publicados entre una y tres veces. La explicación más plausible es económica: repetir un aviso en *El Comercio* tenía un costo. Repetirlo implicaba no haber recuperado al niño y se priorizaba la brevedad por encima del detalle. Esto significa que la infraestructura de datos que producían los avisos estaba condicionada estructuralmente por la economía de la prensa. Por lo tanto, una conclusión plausible es que los registros más ricos en información corresponden a patrones con mayores recursos, mientras que los más pobres en detalle podrían sobrerrepresentar a sectores con menor capacidad de pago. Así, el corpus no es una muestra aleatoria de la servidumbre doméstica limeña, sino una muestra sesgada por el capital de quienes accedían a la plataforma.

Sin embargo, el efecto acumulado de cientos de avisos publicados durante casi tres décadas en el mismo diario produjo un resultado imprevisto: un archivo involuntario de información biométrica colectiva, con categorías recurrentes, convenciones descriptivas estables y una lógica clasificatoria implícita que prefigura los sistemas de identificación que vendrían décadas después. Comprender esta dimensión involuntaria es indispensable para enfrentar los desafíos éticos de su procesamiento computacional. Transformar los avisos en un dataset estructurado obliga al investigador a replicar, con plena conciencia de ello, el mismo gesto que los patrones realizaban al escribirlos. Seleccionar variables, normalizar categorías y asignar valores son operaciones técnicas

[11] Drucker toma el término del latín *capere* (capturar, tomar), en contraposición a *datum* (lo dado). La distinción busca cuestionar la idea de que los datos existen de forma objetiva y previa a cualquier interpretación. Para Drucker, toda información es el resultado de una decisión de alguien que eligió registrar algo de una manera específica, desde una perspectiva específica y con un propósito específico. No hay datos que simplemente «estén ahí»; hay datos que alguien decidió capturar y esa decisión lleva inscrita una posición.

que no resultan neutrales, sino que condensan relaciones de poder y compromisos éticos que deben ser explicitados (Smart et al., 2026). Se trata de decisiones sobre cómo representar a personas que no tuvieron voz en su representación original y que tampoco la tienen en su representación digital.

Como señala Trouillot (1995), el poder opera en el archivo no solo a través de lo que se registra, sino de lo que se silencia. Como señalábamos en el acápite anterior, en el caso de los avisos de fuga, los silencios se revelan al comprender que los únicos datos que existen sobre estos niños provienen de quienes los dominaban, incluidas las categorías biométricas con que sus cuerpos fueron descritos, lo que impide recuperar su perspectiva sin replicar esa misma violencia (Hartman, 2008; Rodrigo Gonzales, 2026). Reconocer esa asimetría no detiene ni ralentiza necesariamente el análisis, pero sí obliga a practicarlo con una reflexividad que los aspectos técnicos de la digitalización y del procesamiento algorítmico de los datos, por sí solos, no garantizan.



Figura 2. Cobertura de variables relevantes según la década de publicación del aviso (Lima, 1841–1869).

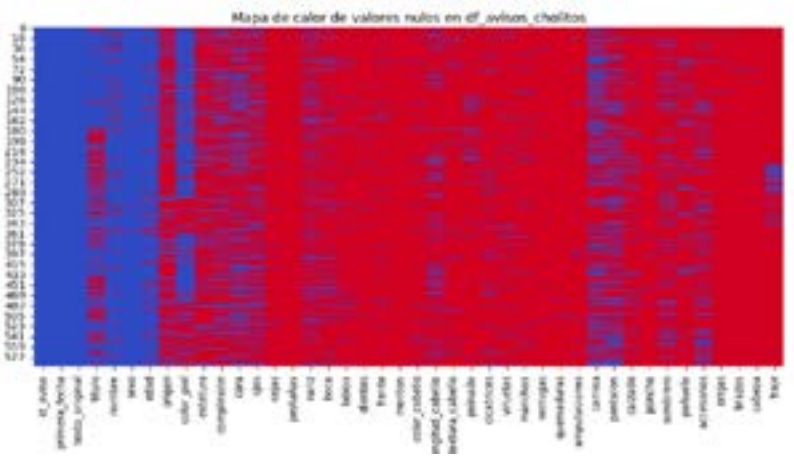


Figura 3. Mapa de calor de valores nulos en el corpus de avisos de fuga: el color azul indica la presencia de un valor; el rojo, su ausencia.

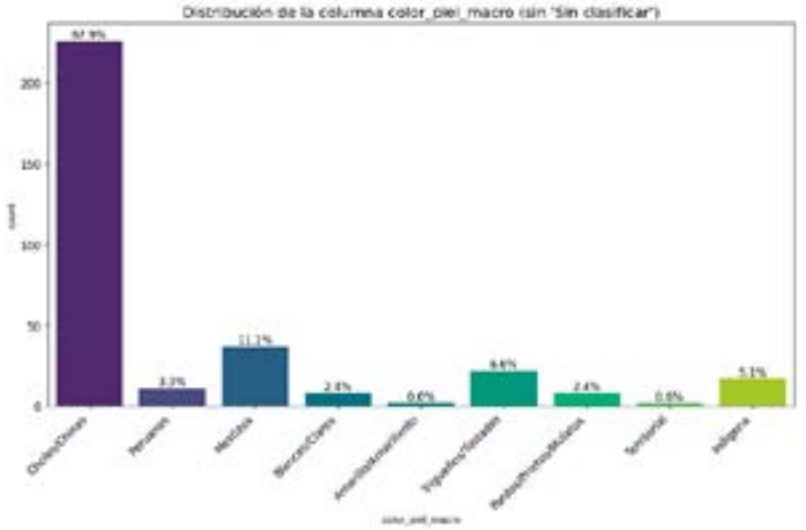


Figura 4. Distribución de la variable color_piel_macro en el corpus de avisos de fuga (Lima, 1841–1869).



Figura 5. Cobertura de variables relevantes según frecuencia de publicación del aviso. Se divide en dos grupos: de 1–3 versus 3–13 publicaciones del aviso.

Sesgos, interpretación y los riesgos de la cuantificación

El procesamiento de los avisos de fuga de *cholitos* en el Perú del siglo XIX permite observar, con notable claridad, tanto la riqueza como las limitaciones de la cuantificación histórica. Lejos de tratar estos registros como una ventana transparente hacia la realidad social de los menores indígenas, conviene asumir desde el inicio que se trata de fuentes producidas bajo condiciones específicas de enunciación, circulación y costo (Gitelman, 2013). En ese sentido, el dato no debe ser leído como una verdad inmediata, sino como el resultado de una operación histórica situada: alguien decidió qué escribir, qué omitir y qué rasgos considerar útiles para la identificación de una persona fugitiva. Precisamente allí reside uno de los problemas centrales de toda base de datos histórica: su aparente objetividad puede inducir a conclusiones demasiado firmes sobre fenómenos que, en realidad, fueron mediados por múltiples sesgos (Crutcher y Belton, 2025).

Uno de los riesgos más evidentes consiste en extrapolar las características de la base de datos como si fueran equivalentes a las de la población en general. Que en los avisos aparezca con gran frecuencia la edad, el sexo, el nombre, o determinadas marcas corporales, no significa necesariamente que esas variables hayan sido

las más relevantes en la vida cotidiana de niños y niñas involucrados. Más bien, su alta presencia responde a su utilidad para el formato del aviso: eran rasgos que facilitaban el reconocimiento y, por tanto, aumentaban las probabilidades de recuperación del menor. En otras palabras, la preponderancia de ciertos atributos en los anuncios no refleja de forma automática una realidad demográfica o cultural, sino una lógica práctica de selección. El medio condiciona el mensaje y el medio impreso, además, impone restricciones materiales que orientan la escritura hacia aquello que pueda ser rápidamente identificable.

Esta advertencia resulta especialmente importante cuando se trabaja con cifras. Los porcentajes, por sí mismos, no deben cerrar la interpretación, sino abrir preguntas. El hecho de que alrededor del 14% de los avisos que mencionan edad corresponda a niños de siete años o menos no autoriza, sin más, a concluir que existía una presencia masiva de fugas infantiles a esa edad; antes bien, obliga a preguntarse por la naturaleza de ese registro, por la posibilidad de secuestro, por la circulación de menores en hogares ajenos, o por la manera en que la infancia era comprendida y descrita en la época. Algo similar ocurre con el pequeño porcentaje de individuos reportados como de veinte años o más: esa cifra tensiona la propia categoría de *cholito* y obliga a cuestionar hasta qué punto el término designaba una edad estricta, una posición social, una relación doméstica o una forma de nombrar más amplia y difusa. En este punto, la estadística no clausura el problema, sino que lo vuelve más visible.

De ahí que la cuantificación deba ser entendida como una herramienta de orientación interpretativa y no como un mecanismo de validación automática. Cuando los datos ordenados muestran, por ejemplo, que ciertas descripciones físicas aparecen con mayor frecuencia, ello no significa necesariamente que esas características fueran las más abundantes en la población; puede significar, simplemente, que eran las más útiles para el propósito del aviso. Lo mismo puede decirse de rasgos como la complexión «gruesa», «gorda» o «robusta», o de la estatura «baja» o «pequeña». Que esas categorías aparezcan reiteradamente no prueba una realidad corporal

uniforme, sino un modo de mirar y clasificar el cuerpo ajeno. En este nivel, el dato cuantificado deja de ser un retrato y se convierte en un síntoma: revela qué tipo de diferencia era percibida como significativa por quien redactó el anuncio.

Esto obliga, a su vez, a una postura crítica frente a los valores estadísticos producidos por algoritmos o procedimientos de clasificación automática (Arriagada Bruneau, 2024). La modelación de datos puede ordenar, agrupar y visualizar regularidades, pero no puede sustituir el juicio histórico. Por el contrario, una lectura acrítica del resultado numérico corre el riesgo de naturalizar categorías que fueron históricamente construidas, inestables y, en algunos casos, ambiguas. El hecho de que un algoritmo marque correlaciones, como sucede con ciertos rasgos faciales, no implica que se haya descubierto una verdad definitiva sobre los *cholitos*; implica, más bien, que se ha encontrado una regularidad dentro de un corpus textual producido por criterios humanos específicos. Por eso, el valor del procesamiento no reside en su capacidad de dictar conclusiones, sino en su potencia para señalar zonas de tensión, vacíos y posibles distorsiones.

En esta línea, la colaboración entre el análisis cuantitativo y la perspectiva histórica resulta indispensable. Los datos ordenados permiten detectar patrones que de otro modo podrían pasar inadvertidos, pero esos patrones solo adquieren sentido si se los reconecta con el contexto social, económico y racial de la Lima decimonónica. La presencia de menores indígenas en hogares limeños, la coexistencia con otros grupos étnicos que ocupaban espacios laborales antes vinculados a la esclavitud, o la circulación de términos como *indio*, *mestizo*, *chino* o *cholito*, no pueden comprenderse adecuadamente sin una lectura histórica más amplia. El número, por sí solo, no sabe de jerarquías, de relaciones de servidumbre, de desigualdad estructural ni de representaciones sociales. Por eso, la historia no complementa de manera ornamental a los datos, sino que los vuelve legibles.

También resulta relevante observar los patrones temporales y espaciales que pueden emerger del corpus. Analizar años,

estaciones, horas del día o recurrencias editoriales permite evitar una interpretación excesivamente rígida de fenómenos que fueron variables y situados. Saber que ciertos avisos se publicaban más de una vez, y que los textos más reiterados tendían a ofrecer menos cobertura descriptiva, es un hallazgo que no debe leerse solo en clave estadística, sino también material: publicar costaba, repetir costaba, describir costaba. La forma del aviso estaba atravesada por restricciones económicas que podían empujar a la condensación del contenido. En consecuencia, la menor densidad descriptiva en algunos casos no expresa desinterés por el niño fugado, sino un equilibrio precario entre urgencia, presupuesto y utilidad práctica.

Algo semejante ocurre con el origen geográfico y con las categorías étnicas. La aparición de referencias a la sierra central o a lugares como Ayacucho, Huancayo o Jauja no puede ser tratada como una geografía cerrada de procedencias exactas. Los niños podían haberse fugado más de una vez, haber sido sustraídos o movidos entre hogares o descritos por terceros con información incompleta. Así, incluso aquello que parece más concreto requiere cautela. El dato geográfico, en este contexto, no es un punto de llegada sino un indicio parcial. La misma cautela debe aplicarse a la identificación étnica: una etiqueta como *cholo* puede coincidir con rasgos fenotípicos, pero también puede operar como una marca social que condensaba posición, servicio y percepción racializada. Por ello, no se trata solo de preguntar quiénes eran, sino quiénes fueron considerados como tales.

Conclusiones

Una gran lección de este tipo de procesamiento es que la cuantificación no elimina la interpretación; la vuelve más exigente. Los datos ordenados muestran tendencias, pero también exponen sus propias condiciones de producción. Frente a ello, la responsabilidad del investigador consiste en no absolutizar ni el algoritmo ni la cifra, sino en someterlos a contraste, revisión y diálogo con otras fuentes y con otros métodos. Solo así es posible construir una aproximación más holística, capaz de articular la información estadística con el

estudio histórico de los procesos que le dan sentido (Browne, 2015; Cole, 2001).

En vez de asumir los resultados como un reflejo directo de la realidad, conviene leerlos como una huella de las formas en que esa realidad fue percibida, jerarquizada y representada. Allí reside, finalmente, el valor historiográfico de los avisos de fuga: no en ofrecer una verdad total sobre los *cholitos*, sino en permitirnos pensar críticamente cómo se ordenaron sus cuerpos, sus nombres y sus ausencias.

Referencias bibliográficas.

Arriagada Bruneau, G. (2024). *Los sesgos del algoritmo. La importancia de diseñar una inteligencia artificial ética e inclusiva*. Santiago: La Pollera.

Bloch, A., Oms, G. y Santana, C. (2025). «Decolonizing Archival Narratives: Exploring Digital Bias in the Catalogs of Portuguese-Colonized African Territories». *The Journal of African History*, n. 66, 1–30. https://doi.org/10.1017/S0021853725100601

Benjamin, R (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge: Polity Press.

Browne, S. (2015). *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press.

Cole, S. (2001). *Suspect Identities: A History of Fingerprinting and Criminal Identification*. Cambridge: Harvard University Press.

Crutcher, M. y Belton, L. (2025). «'They Do Not Have Physiognomies of Their Own': Tracking Racialization of Maritime Professions in the Eighteenth-Century Portuguese Empire». *Journal of Social History*, n. 59, 215–237. https://doi.org/10.1093/jsh/shaf001

Drucker, J. (2011). «Humanities Approaches to Graphical Display». En: *Digital humanities quarterly*. Volume 5 Number 1. https://doi.org/10.63744/r4ysrh7ae534

Drucker, J. (2012). «Humanistic Theory and Digital Scholarship». En: *Debates in the Digital Humanities*, editado por Matthew K. Gold, 85–95. Minneapolis: University of Minnesota Press.

Flores Galindo, A. (1994). *Buscando un Inca. Identidad y utopía en los Andes*. Lima: Horizonte.

Gitelman, L. (2013). «*Raw Data*» *Is an Oxymoron*. Cambridge: MIT Press.

Hartman, S. (2008). «Venus in Two Acts». *Small Axe*, v. 12, n. 2, 1–14.

Peluffo, A. (2022). «Esclavitud infantil y muñecas vivientes en el Perú del siglo XIX». *Revista Chilena de Literatura*, n. 105, 511–530. http://dx.doi.org/10.4067/S0718-22952022000100511

Ragas, J. (2017). «Cholitos, militares y activistas. La Sociedad Amiga de los Indios y la campaña de rescate de niños indígenas (Lima, 1867–1868)». *En América en diásporas. Esclavitudes y migraciones forzadas en Chile y otras regiones americanas (siglos XVI–XIX)*, editado por Jaime Valenzuela, 511–531. Santiago: RIL Editores.

Ragas, J. (2025). *Lima Chola. Una historia de la Gran Migración Andina*. Lima: Instituto de Estudios Peruanos.

Rodrigo Gonzales, P. (2026). «Inscriptions and omissions: registers and the racialization of orphans in late-19th-century Lima, Peru». *Latin American and Caribbean Ethnic Studies*, 1–17. https://doi.org/10.1080/17442222.2026.2659508

Roldán, S. (2026). «Se fugó sin motivo alguno». El caso de los cholitos como servidumbre infantil indígena (Lima, 1839–1868). En *Historia del trabajo doméstico en América Latina. Siglos XVI–XX*, editado por Pedro Valenzuela Reyes y Natalia Baeza Contreras, 117–149. Santiago: RIL Editores.

Smart, S., Lessa, F., Barbano, R. y Risso, M. (2026) «Archives and Algorithms: The Plancondor.org Experience and the Human Rights Potential of New Technologies». *International Journal of Transitional Justice*. https://doi.org/10.1093/ijtj/ijaf023

Stoler, A. (2009). *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*. Princeton: Princeton University Press.

Trouillot, M. (1995). *Silencing the Past: Power and the Production of History*. Boston: Beacon Press.

Velázquez, M. (2024). *Cuerpos vulnerados. Servidumbre infantil y anticlericalismo en el Perú (1840-1920)*. Lima: Random House.