

Introducción

En 1991, el Gobierno de Alberto Fujimori introdujo una nueva moneda peruana —el nuevo sol— y billetes para las denominaciones de 10, 20, 50 y 100. Estos billetes mostraban en el anverso a reconocidas figuras nacionales: José Abelardo Quiñones, Raúl Porras Barrenechea, Abraham Valdelomar y Jorge Basadre, respectivamente¹. En 2021, una nueva serie incluyó otras figuras, como una forma de reflejar los cambios en la sociedad y la identidad peruana: Chabuca Granda, José María Arguedas, María Rostworowski, Pedro Paulet y Tilsa Tsuchiya.

El cambio más obvio es que se ha pasado de un imaginario masculino de la heroicidad nacional (4 hombres y la tardía inclusión de 1 mujer —por supuesto, no es casualidad que se trate de una figura religiosa—) a una distribución más equitativa con respecto al género (2 hombres y 3 mujeres). Es menos evidente que las diferencias entre las series de 1991 y 2021 también reflejan cambios en la imagen de la literatura nacional. En la primera serie, se encuentran 3 figuras relacionadas con la cultura y la literatura peruana: Abraham Valdelomar es autor de célebres cuentos y poemas, todavía leídos en las escuelas; Raúl Porras Barrenechea es uno de los más importantes estudiosos de las crónicas coloniales y autor de *El sentido tradicional en la literatura peruana* (escrito en 1945 y publicado en 1969); Jorge Basadre es un reconocido historiador, pero su primer libro es una compilación de crítica literaria: *Equivocaciones. Ensayos sobre literatura penúltima* (1928). Además de ser hombres, todos ellos tienen en común escribir en español y ser originarios de regiones de la Costa (Ica, Lima y Tacna, respectivamente). La segunda serie ofrece una imagen distinta con 2 figuras: José María Arguedas, escritor de origen andino (Andahuaylas), quien publicó en español y quechua, y Chabuca Granda, compositora y cantautora, quien revela una idea ampliada de la creatividad literaria. En general, se puede afirmar que los billetes actuales forman parte de un esfuerzo mayor por imaginar una literatura peruana más diversa e inclusiva en criterios como el género, la lengua o la región.

¹ El billete de 200, con la figura de Santa Rosa de Lima, empezó a circular en 1995.

El ejemplo anterior permite introducir dos ideas fundamentales en el desarrollo de esta investigación. Primero, este estudio se centra en los escritores, no en sus obras; especialmente, se investiga a los escritores que alcanzan una posición prestigiosa dentro del campo literario, en tanto que son relevantes en otros campos sociales. Estos escritores son parte del denominado «canon de la literatura peruana». Segundo, en este libro no se intenta analizar la producción literaria, sino comprender la recepción de esos autores —sus obras y sus ideas— en otras instancias discursivas (historias, antologías, enciclopedias) e incluso fuera del campo literario (el mundo digital). En otras palabras, se intenta responder a la siguiente pregunta: ¿cuáles son los procesos que permiten a un escritor ser tan relevante para figurar en un billete de circulación nacional? Ahora que los escritores son billetes, es también posible objetivarlos y acumularlos, así que no están muy lejos los días en que se escuchen expresiones como «Esto cuesta un Arguedas» o «Tengo 4 Arguedas». En efecto, en esta investigación también se utiliza un enfoque cuantitativo para analizar el valor literario. Para esto, se utilizan dos conceptos: capital y acumulación.

* * *

En tiempos en que el mercado influye tanto en las obras de ficción como en los textos académicos, no es difícil encontrar, usualmente en la contratapa o la página web que presenta el libro, afirmaciones de este tipo: «Con este libro de poemas, el autor se introduce inmediatamente en el canon de los escritores nacionales» o «El crítico revoluciona el canon literario con sus últimos ensayos». Sin lugar a dudas, el término se banaliza, pierde su eficacia y, ahora, que la palabra «canónico» se use para calificar a un escritor no significa mucho. En el pasado, el contexto no era muy diferente, ya que el canon ha carecido de un aparato metodológico que permita su análisis y su delimitación. Por eso, siempre es posible cuestionar el sustento de esas ideas: ¿cómo podemos saber si un autor es realmente canónico? ¿Es posible determinar el momento y la forma en que un autor medianamente reconocido se convierte en uno canónico? ¿Cómo la crítica influye en la canonización de un escritor? ¿Cuánta influencia puede tener un solo libro de crítica en la construcción del canon? Podemos imaginar un ejemplo concreto con tres narradores peruanos que publicaron alrededor de la misma época: Abraham Valdelomar, Enrique A. Carrillo y Augusto Aguirre Morales. En la actualidad, cualquier persona piensa en el primero como uno de los mayores representantes de nuestra literatura, mientras los otros son mucho menos

conocidos. ¿Qué procesos permitieron que todos ellos ocupen posiciones tan diferentes en el campo literario o en el imaginario nacional de la literatura peruana? La respuesta no se encuentra en la calidad de la escritura o la relevancia de los temas, por un presupuesto que tomamos de Bourdieu y es fundamental en esta investigación: el valor literario no es intrínseco a la obra, sino una construcción de entidades externas.

Para responder esas preguntas, hemos optado por utilizar un enfoque cuantitativo. Si bien la posibilidad de «reducir» un texto de ficción —un poema o una novela— a su equivalente numérico merece más discusión, la recepción y la construcción del valor son los objetos perfectos para ser cuantificados y medidos dentro de los estudios literarios. Con las herramientas adecuadas y el acceso a la información, podemos saber cuántos artículos críticos se han escrito sobre un autor, cuántas veces un libro fue prestado por una biblioteca, cuántos lectores compraron la obra en Internet o cuántas reediciones de un mismo libro se han publicado. Nuestra propuesta implica que utilizar esta información, estos datos, constituye la mejor forma de explicar qué es un autor canónico y cómo ocurre el proceso de canonización. El problema todavía radica en esas herramientas: las teorías, los métodos y los conceptos vinculados al análisis de las obras literarias no ayudan mucho cuando no se quiere examinar el texto, sino el contexto social que recepciona la obra.

En esta investigación, utilizamos herramientas teóricas y metodológicas de dos ámbitos académicos: la sociología de la literatura y las humanidades digitales —en específico, los estudios literarios computacionales—. Sapiro (2016) define la primera como una disciplina que estudia la representación literaria de los hechos sociales y explora la actividad literaria como un fenómeno social (pp. 13-14); esta es la perspectiva que guía nuestro análisis. La sociología literaria no tiene una tradición muy larga, ya que solo se puede rastrear su aparición hasta la segunda mitad del siglo XX, con la publicación de *Sociologie de la littérature* (1958), de Robert Escarpit. En la década de los setenta destacaron dos críticos: Itamar Even-Zohar y Pierre Bourdieu, cuyas propuestas resultan fundamentales para nuestra definición del canon literario. En general, tres planteamientos de la sociología de la literatura son las que influyen directamente en este proyecto: la objetivación de lo literario —o la convicción de que el arte no es un hecho inexplicable—, la importancia de las instituciones para el funcionamiento de la cultura y las formas en que los lectores —los críticos son también lectores— se apropián de los textos.

Sobre las humanidades digitales, todavía existe un debate acerca de su definición². Sin embargo, como ha explicado Kirschenbaum (2014), las humanidades digitales se encuentran bastante bien definidas en el ámbito académico, con ofertas de trabajo, financiamientos, departamentos e institutos, revistas, congresos internacionales y más (p. 55)³. En sentido amplio, las humanidades digitales se caracterizan por la confluencia de la tecnología, los métodos computacionales y el enfoque humanístico para analizar productos culturales. Su impacto en la investigación literaria se denomina «estudios literarios computacionales» (ELC)⁴: la minería de textos para identificar patrones y tendencias, la teoría de redes para analizar la relación entre los personajes de una narración o la estilometría para identificar a los autores de una obra. Todos ellos tienen en común la utilización de datos y el análisis cuantitativo; por eso, es lógico pensar que los ELC pueden funcionar como una extensión de la sociología de la literatura. En efecto, English (2010) dedica un párrafo a este tipo de investigaciones en su revisión de las nuevas formas del quehacer sociológico (p. ix).

En Perú, este campo recién se está construyendo con diversos proyectos que ya empiezan a autodenominarse «proyectos de humanidades digitales». En el ámbito institucional, hay dos ejemplos concretos: el Laboratorio de Humanidades Digitales de la Pontificia Universidad Católica del Perú y la carrera de Humanidades Digitales en la Universidad del Pacífico. Sin duda, existen investigaciones en las que se combinan la informática, el mundo digital y las humanidades desde hace varios años, pero no se consideran como parte de un mismo campo disciplinar, una tradición donde unos influyen en otros. Si en otros países latinoamericanos (México, Argentina, Colombia y Brasil) ya se reflexiona sobre las posibilidades y peligros de los métodos computacionales, esa es todavía una tarea pendiente en el ámbito peruano⁵. Sin lugar a dudas,

2 La página *What Is Digital Humanities?*⁶ contiene 817 definiciones de humanidades digitales propuestas por especialistas. Al hacer clic en un botón, la página ofrece una de esas definiciones; estos son dos ejemplos: «La intersección entre tecnología computacional y preguntas acerca del significado de la condición humana» (Andy Shocket) y «Las humanidades digitales son 1) el uso de la tecnología digital para facilitar y mejorar todas las etapas de la investigación en humanidades y 2) el estudio de la tecnología digital con un enfoque humanístico» (Jason Boyd); la traducción es nuestra.

3 Esto ocurre principalmente en la academia estadounidense y europea; la realidad en Latinoamérica es bastante diferente.

4 Preferimos este término a «estudios literarios digitales».

5 Estos países cuentan con carreras profesionales, revistas especializadas, oportunidades de financiamiento y conferencias. Las humanidades digitales también se han institucionalizado con diferentes asociaciones que todavía se encuentran activas: Red de Humanidades Digitales (Méjico, 2011), Asociación Argentina de Humanidades Digitales (Argentina, 2013), Red Colombiana de Humanidades Digitales (Colombia, 2016) y Laboratório em Rede de Humanidades Digitais (Brasil, 2018).

lo digital ha modificado lo literario, desde su construcción (la literatura electrónica) hasta su recepción (los *booktubers* y la difusión en redes sociales), por lo que no será una sorpresa que también influya en los estudios literarios peruanos para convertirse en una línea de investigación: «Tanto en lo que prometen como en la amenaza que suponen, las humanidades digitales sirven como simulacro de una manera futura de trabajar y estudiar las humanidades integrando algunos valores de la sociedad contemporánea sobre lo digital» (Liu, 2020, p. 125). En esta investigación, los métodos de las humanidades digitales y los ELC han servido para tres propósitos: recolectar información, analizar los datos y crear visualizaciones.

Como los ELC son nuevos en Perú, no existen investigaciones ni instituciones que se vinculen con esta tendencia. Como afirma Eve (2022), la diferencia entre los ELC y los estudios literarios no computacionales no radica en el tipo de investigación ni en los temas, sino en la escala, el grado y la velocidad (p. 8). Analizar la representación de la escuela no solo en *Todas las sangres* (1964) de Arguedas, sino en todas las novelas indigenistas desde la fundacional *La trinidad del indio* (1885) de José Torres Lara; demostrar las jerarquías entre personajes visualizando las relaciones entre ellos (análisis de redes) en lugar de analizar el tema de la subordinación en *Monólogo desde las tinieblas* (1975), de Antonio Gálvez Ronceros; no buscar cada antología de literatura quechua visitando diferentes bibliotecas y solicitándolas en préstamo, pero estudiarlas usando las descripciones que se hacen de ellas y de su contenido en repositorios digitales (metadatos), o acceder a su versión digital⁶.

Esos ejemplos demuestran que los ELC requieren un doble proceso de transformación. El primero consiste en la transformación del texto literario en datos. Este proceso es necesario porque los programas de computadora requieren que la información tenga una determinada estructura para poder ser analizada; así, se necesita convertir los textos sin estructura —novelas, cuentos, poemas, artículos académicos, entre otros— en datos estructurados —las bases de datos, por ejemplo—⁷. Esto, usualmente, se logra escaneando el libro y guardándolo en un formato legible por las computadoras (digitalización). Esto no significaría ningún problema para el segundo ejemplo del párrafo anterior (el análisis de *Monólogo desde las tinieblas* [1975]), ya que se

⁶ Todos estos son ejemplos de investigaciones que publicamos en años anteriores: nuestro propósito es mostrar la potencialidad de los ELC con temas que conocemos.

⁷ La idea de estructura es influencia de la ciencia de datos y la programación: en los estudios literarios, no tiene ningún sentido decir que un poema carece de estructura. Entendemos que esta primera transformación es uno de los tópicos más polémicos de los ELC, por lo que volvemos a este tema en el capítulo III.

trata de un solo texto, y un solo investigador podría realizar la digitalización muy rápido. En cambio, el primer ejemplo —el estudio sobre la novela indigenista— resulta más complicado por el acceso a las fuentes primarias y el número de libros. La labor de digitalización probablemente estaría a cargo de una institución (una biblioteca o una universidad): esto resalta la importancia del compromiso institucional con el acceso abierto a estos recursos digitales para que los investigadores puedan continuar con sus proyectos. La segunda transformación consiste en utilizar programas de computadora para visualizar esos datos literarios. En un primer nivel, diferentes programas pueden crear visualizaciones específicas: Voyant crea un gráfico con la tendencia del uso de las palabras en un grupo de textos; Gephi permite visualizar la relación entre los personajes en forma de redes. También es posible utilizar lenguajes de programación para realizar las visualizaciones: R y Python son los más populares actualmente. Cada opción presenta sus propios beneficios: la facilidad de utilizar programas que son amigables para el usuario en el primero; la mayor libertad para crear las visualizaciones en el segundo. En todo caso, este proceso —las dos transformaciones— requiere conocimientos y experiencia bastante alejados del campo de acción del estudioso de la literatura.

Un factor, probablemente, no quedó claro en esta breve explicación de las dos transformaciones: cada una de ellas implica un acercamiento específico al texto literario, una interpretación, una lectura. Como afirma Bode (2023), «data and models do not reveal literary phenomena as they really are but re-present or externalize an understanding of them»⁸ (p. 525) y, como cualquier representación, los datos y las visualizaciones son una refracción producto de una serie de elecciones voluntarias o involuntarias. Imaginemos que queremos determinar el léxico de un conjunto de textos —establecer, por ejemplo, un diccionario de la narrativa del conflicto interno—: para esos casos, la eliminación de las palabras vacías⁹ es usual para evitar que las más usadas de ese corpus sean «a», «de» y «el». Esto es, por supuesto, una elección del investigador que se alinea con el objetivo del proyecto. En cambio, los datos literarios tendrán una forma totalmente diferente si nuestro propósito es identificar al autor de un texto en específico: un escritor con un particular uso de las preposiciones y los signos de puntuación. Para comparar esa obra con las de otros escritores y tener un indicio de un posible autor, conviene dejar las palabras vacías e idear una forma de mantener también los signos y su

⁸ «Datos y modelos no revelan los fenómenos literarios tal y como son en realidad, sino que re-presentan o exteriorizan una comprensión de estos» (traducción propia).

⁹ Palabras muy usadas y sin un significado: las preposiciones «a», «ante», «bajo», «con», etc.

posición en las oraciones. Por otro lado, la manipulación de datos para crear visualizaciones constituye un segundo proceso de interpretación. Además de los diferentes tipos de gráficos que pueden usarse para visualizar la misma información —barras, nube de palabras, gráficos circulares, diagrama de árbol, entre otros—, cada uno de ellos incluye un repertorio de posibilidades que van desde el tipo de letras hasta la organización espacial de las variables. Por eso, Drucker prefiere denominarlos «expresiones gráficas de interpretación» en lugar de visualizaciones. En los siguientes capítulos, volveremos constantemente a esta idea de datos y visualización como forma de interpretación.

¿Vale la pena todo este esfuerzo —digitalización, organización de datos, visualización, tiempo para realizar estos procesos, además de la lectura y la revisión la bibliografía— para el estudio de la literatura? Volvamos a los ejemplos anteriores: si se trata de examinar un solo libro (*Monólogo desde las tinieblas* [1975]), la respuesta es probablemente no; pero si el objetivo es acercarnos al corpus de la novela indigenista, los ELC ofrecen herramientas útiles para el análisis y la argumentación.

Primero, detengámonos en nuestro objeto de estudio: el corpus. Este término alude a todas las novelas indigenistas, no solo las 5 o 6 que siempre son analizadas y leídas en las universidades; se refiere a todas aquellas que solamente aparecen mencionadas en las historias de literatura más minuciosas, incluso esas que ni siquiera llegaron a aparecer en esas historias y son simples referencias en el catálogo de las bibliotecas. En ese sentido, el corpus es lo opuesto al canon. Si nos enfrentamos a un grupo de 60 obras, será mucho más fácil y rápido analizarlas (leerlas, encontrar patrones, identificar similitudes y diferencias) utilizando las herramientas de los ELC. Moretti (2000) ha denominado a este enfoque «lectura distante» —*distant reading*, en inglés—, el cual se ha convertido en un componente fundamental de los ELC y se opone al de lectura atenta —*close reading*, en inglés—, una tradición interpretativa en la academia angloamericana basada en la lectura minuciosa de fragmentos. Para este crítico, la lectura atenta es canonizadora al enfocarse en unos pocos textos considerados relevantes; la lectura distante permite salir del canon para enfocar elementos literarios que trascienden la unidad de la obra literaria. Moretti (2000) también afirma que la lectura distante es la única forma de comprender el sistema literario en su totalidad, y nosotros agregamos: se necesita esa lectura en conjugación con los métodos computacionales.

En efecto, en la academia estadounidense, los proyectos de ELC suelen trabajar con cantidades inmensas de datos. Incluso, los títulos revelan esa tendencia: «Style, Inc. Reflections on Seven Thousand Titles (British Novels,

1740-1850» (Moretti, 2009) y «The Quiet Transformations of Literary Studies: What Thirteen Thousand Scholars Could Tell Us» (Goldstone y Underwood, 2014). La razón es la mayor capacidad tecnológica y económica de los centros de investigación: estos vienen trabajando en la creación de bases de datos y archivos digitales desde hace décadas, por lo que los investigadores pueden acceder fácilmente a esos materiales. En Perú, recién nos encontramos en esa primera fase de establecimiento de la infraestructura, que es fundamental para un análisis de ELC. Para colaborar en su consolidación, dirigimos el proyecto «De desastres a celebraciones. Archivo digital de novelas peruanas (1885-1921)»¹⁰, el cual es administrado por el Centro de Estudios Literarios Antonio Cornejo Polar (CELACP). Este archivo ofrece, en acceso abierto, las novelas en dos formatos: archivo PDF, para su lectura, y texto plano, archivo con extensión TXT, para el análisis computacional. No conocemos otro proyecto de preservación del patrimonio literario con las mismas características en el país, aunque ya algunas instituciones cuentan con importantes plataformas de preservación digital, como es el caso de la Biblioteca Nacional del Perú, el Instituto Riva-Agüero y el Centro de Documentación del Perú Contemporáneo. Es esta infraestructura la que permite —o permitirá— el análisis computacional de la literatura peruana a mayor escala.

En segundo lugar, un proyecto que analice 60 novelas o más debe resolver el problema de la argumentación y la inclusión de la evidencia. Por ejemplo, si queremos probar que todas esas novelas indigenistas comparten un rasgo formal, tenemos dos opciones para mostrar la evidencia que sustente este argumento: incluir citas de unas pocas novelas representativas —una estrategia que vuelve a alejarse del corpus para enfocarse en unas pocas obras canonizadas— o incluir citas de todos los textos —opción poco práctica e imposible de realizar en el espacio limitado de un artículo académico o un libro—. Una tercera opción consiste en explicar ese rasgo sin incluir ninguna evidencia. Sin importar que alternativa se escoja, debemos reconocer que la utilización de citas textuales que pueden ser cotejadas con el texto analizado es la estrategia predominante en la investigación literaria: «Literary scholars have continued to overwhelmingly rely on methods that depend on an authoritative observer choosing the “best” observation(s) to prove some larger point»¹¹ (Piper, 2020, p. 9). Sin embargo, los ELC ofrecen una nueva forma de proveer evidencia en los estudios literarios.

¹⁰ El archivo está disponible en <https://celacp.org/proyectos/de-desastres-a-celebraciones/>

¹¹ «Los estudiosos de la literatura han seguido confiando abrumadoramente en métodos que dependen de un observador autorizado que elige la mejor observación (u observaciones) para demostrar alguna idea más amplia» (traducción propia).

Es algo ya aceptado que las ciencias humanas —los estudios literarios— y las ciencias duras —las ciencias naturales y la física, por ejemplo— difieren en varios aspectos. Schaeffer (2013) utiliza el concepto de ecología para explicar esas diferencias: básicamente, cada una de estas ciencias posee una ecología con características propias con respecto al territorio, la densidad demográfica y la competencia (pp. 22-23). La disimilitud que consideramos más importante se refiere al método de validación: mientras las ciencias duras se vinculan al cosmopolitismo, las ciencias humanas surgieron al mismo tiempo que las naciones y los nacionalismos. En otras palabras, las primeras ya se caracterizaban por la investigación transnacional y los criterios de comprobación, cuando las segundas enfocaron temas y discusiones nacionales¹². En ese sentido, sobre las ciencias duras, es posible afirmar que se fundan con un acuerdo sobre el tipo de relaciones que el discurso debe mantener con el objeto de estudio, relaciones que han de poder reproducirse por fuera de la esfera subjetiva del investigador, y que admiten la evaluación de los resultados obtenidos a partir de la interacción con otros investigadores, así como con el mundo (Schaeffer, 2013, pp. 23-24). Ese acuerdo implica que la objetividad de la investigación se vincula estrechamente con su capacidad de ser reproducible: el responsable del análisis debe ofrecer todos los materiales que permitan la reproducción del experimento, y otro investigador debe obtener los mismos resultados si sigue el mismo proceso. Esta reproducibilidad no es común en las ciencias humanas. En palabras de Bode (2022 y 2023), el rango de posibilidades de aquello que puede ser verdad en los estudios literarios es mayor y está menos codificado que en otras ciencias (p. 531).

El paradigma de los ELC parece haber acercado ambas metodologías al utilizar programas de computadoras para la investigación y compartir el «paso a paso» del análisis para la publicación. Los defensores de esos métodos son bastante optimistas y afirman que esto significa un cambio radical en los estudios literarios¹³. Aunque no compartimos esa opinión, lo cierto es que el análisis basado en algoritmos, datos literarios, códigos de programación y otros es un terreno propicio para este nuevo método de validación en las ciencias humanas: compartir los fundamentos del proceso de interpretación para que otros investigadores puedan comprobar las conclusiones. En palabras de Piper

12 Esto explica porque es posible hablar de estudios literarios peruanos, pero no de química peruana, por ejemplo.

13 Para una discusión de los beneficios de la reproducibilidad o «investigación repetititva» en los ELC y las humanidades digitales, ver Schöch (2023). Otro ejemplo de las potencialidades de este factor es el artículo de Da (2019), quien replica varios proyectos para evidenciar las incongruencias y falencias en la metodología de los ELC.

(2018), el estudio de la cultura (y la literatura) se vuelve entonces más arquitectónico, social y colectivo (p. 7). Es ese carácter arquitectónico —en el sentido de etapas o procesos que se ensamblan entre sí cuando se utiliza una serie de instrucciones— o «descomposicional» —un problema que puede ser dividido en unidades más pequeñas de resolución (Eve, 2022, pp. 6-7)— de la investigación en ELC lo que explicamos en cada uno de los capítulos, con la intención de brindar las herramientas necesarias para reproducir los planteamientos y comprobar las conclusiones. Asimismo, compartimos todos los datos (en forma de archivos Excel y PDF) y el código utilizado para el análisis (en este caso, es necesario usar RStudio); de esa forma, cualquier interesado en este tema puede replicar la investigación¹⁴.

Esto quiere decir que los ELC también funcionan como una nueva forma de proponer evidencias en la investigación literaria. En el caso de la novela indigenista, la interpretación se puede sustentar con una cita o un grupo de citas, pero también es válido presentar un gráfico con la cantidad de veces que aparece una palabra específica o un elemento narrativo en todo el corpus analizado —es decir, una visualización de datos literarios—. Las dos formas de argumentar no son excluyentes y pueden aparecer juntas en el mismo análisis (citas de los textos y visualizaciones de datos); sin embargo, debemos admitir que la segunda es poco común en los estudios literarios. Es necesario advertir que los datos literarios no son, en sí mismos, evidencia; lo que puede servir para sustentar un argumento es el uso analítico de esos datos —esto se logra usualmente con una visualización y su contextualización (Owens, 2011). Como afirma Piper (2020), se trata de escoger el proceso discursivo que hace que nuestra evidencia sea más visible —y comprensible— para otros (p. 8). En esta investigación, hemos usado ambas de formas de evidencia, citando a críticos y especialistas sobre determinados temas para sustentar nuestras ideas e incluyendo visualizaciones que evidencian nuestros argumentos.

Otro elemento fundamental en los ELC es el enfoque cuantitativo, también denominado «análisis textual cuantitativo»¹⁵. Este enfoque se basa en dos aspectos: el empleo de métodos computacionales para convertir el texto en

¹⁴ Ver Carrillo Jara (2024).

¹⁵ El desarrollo de esta metodología ha conducido a la aparición de generadores de texto con inteligencia artificial (IA), como ChatGPT. La IA es uno de los temas más populares en la actualidad, tanto en el ámbito académico como en la vida cotidiana. Creemos que son dos los aspectos más preocupantes de esta nueva tecnología: la forma en que nos acercamos a los textos va a cambiar para enfocar más el autor que el contenido —ahora cada lectura puede empezar con la pregunta: ¿este texto lo escribió un ser humano?—, y la tecnología necesaria para desarrollar la inteligencia artificial involucra explotación laboral, extractivismo, contaminación ambiental, entre otros (ver Crawford, 2021).

datos numéricos y la utilización de matemáticas o estadística como método de análisis. En los ELC, es importante reflexionar cuidadosamente acerca de qué componente del texto literario necesita ser cuantificado, ya que, como advierte Hoover (2008), un diseño equivocado al inicio de la investigación ocasiona resultados improductivos y pérdida de tiempo. El mismo autor advierte sobre tres razones por las que este enfoque no ha tenido mayor impacto en los estudios literarios: no enfocarse en tópicos de mayor transcendencia, ignorar los antecedentes y el contexto del tópico analizado, y enfatizar en exceso el aspecto tecnológico. Nosotros agregaríamos otro motivo: el análisis cuantitativo es usualmente entendido como una suerte de reduccionismo del carácter estético de la obra literaria¹⁶.

No es común que un libro en los estudios literarios dedique sendos apartados a explicar la metodología. Sin embargo, creemos que el uso de un método nuevo requiere la explicación de los presupuestos de ese método. Debido a que esta es la primera investigación que utiliza las herramientas de las humanidades digitales y los ELC para analizar el sistema literario en Perú, es necesario dedicar algunos párrafos en cada uno de sus capítulos a hacer explícito cómo se han utilizado esas herramientas en el proceso de investigación.

* * *

La construcción del canon no es un tema nuevo en la crítica literaria peruana. Sin los aportes fundamentales de Cornejo Polar (1989), García-Bedoya (2012), González-Stephan (2002), Lergo Martín (2008) y Rodríguez Rea (2002), no hubiéramos podido construir nuestra exploración sobre el tema. No obstante, los estudios críticos han obviado dos componentes importantes: una reflexión teórica sobre el concepto y un análisis sistémico del proceso de canonización. Esto ha significado que los acercamientos siempre han sido parciales al problema del canon. Usualmente, esto se hace evidente cuando los críticos presentan la lista de escritores propuesta por un solo autor o una sola obra como el canon de la literatura peruana en un momento determinado. Es decir, analizan la opinión de un participante en el campo literario como si fuera el representante de todo un conjunto de actores e instituciones. Por esa

¹⁶ Este no es lugar para defender los ELC, pero nos interesa señalar sus principales críticas. Bode (2023) identifica tres líneas principales: los ELC son un instrumento del liberalismo y el capitalismo, debido a su énfasis en la innovación tecnológica; los ELC reducen la complejidad de la obra literaria en números y datos, y los ELC implican la negación de la principal forma de acercarse a los textos literarios: la lectura. Esperamos sinceramente que este libro genere la discusión necesaria para continuar esta conversación en otros espacios.

razón, el objetivo de este libro es proponer un concepto y una metodología del canon literario que puedan ser utilizados por otros investigadores al analizar diferentes literaturas nacionales.

El primer capítulo constituye el marco teórico y metodológico de la investigación. La revisión de los planteamientos de dos destacados sociólogos de la literatura —la teoría de los polisistemas de Itamar Even-Zohar y la de los campos sociales de Pierre Bourdieu— establece la base para proponer una definición: el canon es el conjunto de productores y productos que acumulan la mayor cantidad de capital literario otorgado por las instituciones de consagración. En este concepto, dos categorías son fundamentales: el capital literario y las instituciones. Con el primero, nos referimos a las objetivaciones del valor literario: manifestaciones del acuerdo que existe sobre la importancia de un autor y su obra —inclusiones en sílabos universitarios, premios y reconocimientos, textos antologados, menciones en historias literarias, textos críticos, entre otros—. Por otro lado, las instituciones constituyen las diferentes comunidades que participan en la actividad literaria, las cuales se rigen por normas y tienen cierto poder sobre otros agentes. Una versión inicial y mucho más breve de algunas secciones de este capítulo se publicó en línea en la revista *Molok* (Carrillo Jara, 2017).

En el segundo capítulo, presentamos una aplicación de nuestra metodología en el análisis de 4 textos críticos de principios del siglo xx: Riva-Agüero (1905), García Calderón (1914), Gálvez (1915) y Prado (1918). El enfoque cuantitativo resulta fundamental, porque defendemos que el capital literario es el resultado de acuerdo o desacuerdo entre esos críticos sobre el valor de los escritores de literatura peruana. Es decir, un autor tiene mayor capital si los cuatro textos lo consideran importante en la tradición literaria. Utilizamos esta premisa para organizar el campo literario peruano de acuerdo con tres posiciones: escritores consagrados, escritores legitimados y escritores aspirantes. Como fundadores de la historiografía y la crítica en Perú, estos intelectuales también incluyeron gran cantidad de referencias a autores extranjeros. Exploramos esta característica con la necesidad que tienen los campos en construcción de trasladar capital de literaturas nacionales consagradas o consolidadas.

Las antologías literarias son el principal objeto de estudio en el tercer capítulo. Luego de definir el género y analizar 27 antologías, nuestra metodología confirma la posibilidad de sistematizar las posiciones de los escritores: poetas consagrados, poetas legitimados y poetas aspirantes —desde escritores que aparecen en la mayoría de las antologías hasta los que son incluidos en solo una—. El análisis de la acumulación de capital conduce a proponer un

concepto que explica las modificaciones del valor literario: las trayectorias del capital. Una trayectoria ascendente o descendente muestra cuando el prestigio de un escritor ha aumentado o disminuido, respectivamente, con el transcurso del tiempo. Asimismo, ampliamos el enfoque de nuestra propuesta: de explorar la canonización en una sola institución pasamos a examinar el canon como el resultado del diálogo entre cuatro instituciones —antologías, tesis universitarias, ediciones y escuela—.

Finalmente, en el cuarto capítulo, consideramos pertinente analizar una institución no literaria: Wikipedia, una enciclopedia del mundo digital. La diferencia entre instituciones que funcionan dentro y fuera del campo literario es el fundamento para plantear la diferencia entre el prestigio y la popularidad: ambos conceptos tienen un funcionamiento similar, pero los agentes literarios tienen poco o nula participación en el segundo. En ese sentido, el capital que está en juego en Wikipedia es la popularidad. En nuestro análisis de dos versiones de la enciclopedia —español y quechua—, identificamos que los idiomas, así como los países y los escritores, pueden acumular capital literario. Por eso, la comunidad quechuahablante, al encontrarse en una posición de subordinación, requiere la utilización de diferentes estrategias para proponer su propia idea de literatura peruana en Wikipedia. Con respecto a la diversidad de las literaturas regionales, identificamos que el mundo digital extiende las simplificaciones de los estudios literarios: en ambos casos, se ha invisibilizado el aporte de las expresiones culturales de la Amazonía peruana. Algunas secciones de este capítulo se publicaron en inglés en *Journal of Cultural Analytics* (Carrillo Jara, 2023).

* * *

Esta última sección requiere un cambio del plural de la investigación y la academia al singular de la gratitud. Este proyecto se inició en 2013 cuando llevé un curso con Carlos García-Bedoya en la maestría de Estudios Culturales de la Universidad Nacional Mayor de San Marcos; él me introdujo a las propuestas de Even-Zohar y Bourdieu¹⁷. Este programa de maestría ya no existe y tampoco sustenté la tesis para obtener el diploma, pero esos dos años resultaron fundamentales para mis proyectos posteriores. Recuerdo con cariño y respeto no solo ese curso, sino también aquellos dictados por Mauro Mamani

¹⁷ García-Bedoya ha publicado un libro, *Hacia una historia literaria integral* (2021), donde explica la importancia de esas teorías para el estudio de los procesos literarios. Cotejar esa publicación demuestra la influencia de sus ideas en esta investigación.

y Ladislao Landa. Estoy convencido de que todos los que pasamos por San Marcos reconocemos la influencia de la universidad en nuestra investigación, nuestros proyectos personales y nuestras ideas. Por eso, me alegra que este libro se publique con el fondo editorial de la universidad: sin importar donde esté, siempre seré un orgulloso sanmarquino.

La mística de San Marcos se debe tanto a las clases, los profesores y los estudiantes. El tercer capítulo es una reescritura de un proyecto presentado en esa universidad junto a Alex Morillo y Eduardo Lino, especialistas en poesía peruana, compañeros en el pregrado de Literatura y amigos hasta hoy. Ya sea discutiendo algunos aspectos de esta investigación, escuchando algunas ideas que parecían imposibles en ese momento o empujándome a publicar este libro, los amigos sanmarquinos siempre estuvieron y están ahí: Eduardo Huaytán Martínez, Juan Cuya, Rafael Gallardo, Jannet Torres y Néstor Saavedra.

Llegar a Purdue University para hacer la maestría y el doctorado significó conocer el mundo de los ELC y afinar el rudimentario enfoque cuantitativo que ya había estado practicando en Lima. Igual de importante fue conocer amistades que hicieron olvidar el frío de Indiana y la sociedad estadounidense: Débora Borba, Megi Papiashvili, Manuel Cárdenas y Lorena Piña. El segundo y el cuarto capítulo empezaron como trabajos finales en los cursos de Matthew Hannah (uno sobre introducción a las humanidades digitales y otro sobre análisis computacional de textos), con quien aprendí desde limpiar datos hasta crear visualizaciones, y quien nunca dejó de apoyarme en esta investigación. Marcia Stephenson, Yonsoo Kim, Song No y Paul Dixon leyeron y comentaron una primera versión de este libro cuando se presentó como tesis de doctorado: sus sugerencias mejoraron este proyecto en todos sus aspectos. En realidad, Song No hizo más que leer la tesis: me animó a iniciarme en las humanidades digitales —aunque ahora yo prefiera hablar de ELC— y este proyecto es el resultado de su constante consejo. Este libro no se hubiera escrito sin todas estas personas; por supuesto, las carencias y los excesos en sus planteamientos son mi entera responsabilidad.

Leer, investigar y escribir implican robar tiempo a algo más importante que la literatura o la academia: mi familia. Gracias, Miluska y Paz, por su paciencia y su cariño.