

La Cadena de Valor de Datos Multimodales (M-DVC): Una Herramienta Conceptual para el Apoyo al Desarrollo de Soluciones de Analítica Multimodal del Aprendizaje

Shashi Kant Shankar, María Jesús Rodríguez-Triana, Adolfo Ruiz-Calleja, Luis P. Prieto, Pankaj Chejara, and Alejandra Martínez-Monés

Resumen—Los sistemas de analítica multimodal del aprendizaje (MMLA), entendidos como aquellos que emplean evidencia multimodal del aprendizaje para modelar mejor la situación de aprendizaje, aún no están muy extendidos en la práctica. Entre las posibles razones, cabe mencionar su complejidad técnica y la baja participación de las personas interesadas (e.g., docentes o alumnos) en el diseño de dichas soluciones. Este artículo propone una cadena de valor de datos multimodales (M-DVC) para dar apoyo al proceso de sistematización y a la comunicación entre las personas involucradas durante la especificación de sistemas de MMLA. Esta herramienta conceptual, fruto de las lecciones aprendidas del campo de los macrodatos (Big Data) y las necesidades emergentes de escenarios MMLA, ha sido evaluada en tres estudios de caso auténticos con sistemas MMLA en desarrollo. Los resultados de nuestra evaluación con métodos mixtos enfatizan la utilidad de la M-DVC para detectar presunciones no consensuadas o pasos dudosos del procesamiento de datos en las primeras etapas del desarrollo. La evaluación también reveló limitaciones de la M-DVC en cuanto a la terminología técnica utilizada, así como la necesidad de incluir información contextual más detallada. Dichas limitaciones apuntan a cómo, proporcionando una descripción más clara y facilitando la comunicación en grupos multidisciplinares, la M-DVC puede contribuir a la creación de soluciones MMLA significativas para el contexto educativo.

Index Terms—Cadena de Valor de los Datos, Analítica Multimodal del Aprendizaje, Utilidad, Evaluación, Comunicación entre las Partes Interesadas, Herramientas Conceptuales.

I. INTRODUCCIÓN

El campo de las analíticas del aprendizaje (LA, del inglés *Learning Analytics*) tiene como objetivo comprender y optimizar el aprendizaje, utilizando datos en formato computacional cuyo procesamiento manual no es factible [1]. Normalmente, esto se hace mediante el análisis de trazas digitales (i.e., logs) de plataformas educativas. El análisis de aprendizaje basado en trazas, sin embargo, ofrece solo una vista parcial de los procesos, resultados y entornos de aprendizaje [2]. El análisis

multimodal del aprendizaje (MMLA, del inglés *Multimodal Learning Analytics*), por otro lado, tiene como objetivo presentar una visión más holística a través de la recopilación, el procesamiento y el análisis de múltiples fuentes de datos sobre el aprendizaje [3], provenientes frecuentemente tanto de espacios digitales como físicos [4]. Las modalidades, en este contexto, se refieren a los diferentes canales de comunicación que el sistema MMLA utiliza para recopilar información sobre el escenario de aprendizaje [5] tales como los logs del sistema que registran las interacciones de los estudiantes con la interfaz de usuario, un audio con la voz de los participantes en una actividad colaborativa presencial o un vídeo registrando sus movimientos. A su vez, estas fuentes de datos reflejan los múltiples modos en que ocurren dichos procesos de aprendizaje (e.g., visual, auditivo, táctil o kinestésico) [6], [7]. Gracias a los recientes avances tecnológicos, es posible recopilar rastros de múltiples modalidades, lo que lleva a trabajar con conjuntos de datos heterogéneos [8].

Inspirado por las prácticas existentes en la minería de datos y los datos masivos (*Big Data*), MMLA hace uso intensivo de datos [9]. En el proceso de análisis multimodal, son necesarias múltiples actividades de procesamiento de datos, como preparar, organizar y fusionar los datos de diferentes fuentes [10]. Estas actividades se pueden organizar en lo que se denomina una cadena de valor de datos (DVC, del inglés *Data Value Chain*): un conjunto de actividades que realiza una empresa que opera en una industria específica para entregar un producto (ya sea un bien o un servicio) valioso para el mercado [11]. En el contexto de la minería de datos y los datos masivos, se han definido DVCs que incluyen siete actividades de procesamiento de datos: 1. Recopilación y anotación; 2. Preparación; 3. Organización; 4. Integración; 5. Análisis; 6. Visualización y 7. Toma de decisiones [12]. Cada una de estas siete actividades de procesamiento de datos, a su vez, conlleva múltiples pasos de procesamiento de datos que pueden variar según los requisitos de cada caso (e.g., la eliminación de valores ruidosos o incorrectos, los procedimientos en caso de valores faltantes o los pasos para la agregación)

A fin de ayudar a las partes interesadas en las prácticas de enseñanza-aprendizaje (e.g., docentes y alumnos) en la toma de decisiones basada en evidencia [13], MMLA analiza conjuntos de datos heterogéneos provenientes del escenario educativo. Sin embargo, a la hora de diseñar las soluciones,

S. Shankar, Escuela de Tecnologías Digitales, Universidad de Tallín, Estonia (e-mail: shashik@tlu.ee).

M. J. Rodríguez-Triana, and P. Chejara, Escuela de Tecnologías Digitales, Universidad de Tallín, Estonia.

L. Prieto, Facultad de Ciencias de la Educación, Universidad de Tallín, Estonia.

A. Ruiz-Calleja, and Alejandra Martínez-Monés, Grupo de Investigación GSIC-EMIC, Universidad de Valladolid, España.

se deben tomar múltiples decisiones para pasar de los datos multimodales brutos a una interpretación significativa para las partes interesadas. Tales decisiones van desde seleccionar qué actividades de procesamiento de datos son relevantes [14], [15], secuenciarlas [16] hasta organizar los pasos concretos de cada una de esas actividades [17]. Además de tomar y documentar sistemáticamente estas decisiones, trabajos previos en el área de MMLA destacan la necesidad de incluir información contextual del escenario de aprendizaje en el análisis, para ayudar al procesamiento e interpretación de dichos análisis [18], [19]. Por ejemplo, conociendo el listado de alumnos registrados en un curso y su asistencia a clase permitiría distinguir un estudiante inactivo de un estudiante ausente; una nota del docente indicando que la conexión WiFi de la escuela no funcionó durante 15 minutos de la lección ayudaría explicar por qué no hubo actividad de registro durante ese período.

MMLA sigue siendo un campo emergente, dominado principalmente por prototipos de investigación que no llegan a ser adoptados en la práctica educativa común. Además de la complejidad inherente de tales sistemas [17], [16], la literatura también señala que la falta de participación de las partes interesadas y las dificultades en la comunicación entre las mismas (e.g., entre docentes, investigadores y desarrolladores de tecnología) pueden desempeñar un papel fundamental en la adopción [20]. Las diferentes partes interesadas a menudo necesitan comunicarse para poder generar la especificación de requisitos del *software*, uno de los principales documentos contractuales de cualquier producto de desarrollo de *software* [21]. Sin embargo, debido a la corta trayectoria del campo, actualmente las partes interesadas tienen una experiencia limitada en el diseño, implementación y/o uso de sistemas MMLA (lo que dificulta dicha comunicación, incluso más allá de las brechas terminológicas clásicas entre la investigación, la práctica educativa y el desarrollo de *software*).

En nuestro trabajo previo [22], aplicamos la cadena de valor de datos genérica de Big Data [12] a cuatro escenarios MMLA para comprender las especificidades de las actividades de procesamiento de datos en MMLA. Extendiendo ese trabajo previo, en este artículo incorporamos las lecciones aprendidas que modelan la evidencia multimodal del aprendizaje en esos cuatro escenarios usando el DVC, y construimos un DVC más específico para MMLA: la cadena de valor de datos multimodales (M-DVC, del inglés Multimodal Data Value Chain). Para estudiar la efectividad de la M-DVC en el apoyo a la comunicación entorno al desarrollo de soluciones MMLA, aplicamos la M-DVC en tres estudios de caso. M-DVC se utilizó para apoyar la conversación entre las partes interesadas en diferentes etapas durante el desarrollo de soluciones MMLA con diferentes niveles de complejidad. Para la recogida de datos, se siguió el protocolo de pensamiento manifestado (en inglés, *think-aloud*) y se realizaron entrevistas con un total de seis participantes (dos por proyecto, uno con un perfil de desarrollador/analista y otro con un perfil de investigador educativo/docente). Ambas fuentes de datos permitieron estudiar el apoyo que brindó el M-DVC en términos de comunicación, sistematización, identificación de información contextual y toma de decisiones en cuanto al diseño del sistema.

El resto de este documento está estructurado de la siguiente manera: la Sección II presenta brevemente el estado del arte, mientras que la Sección III resume brevemente los cuatro escenarios MMLA utilizados para construir la M-DVC y describe esta contribución en detalle. Luego, la Sección IV detalla la metodología del estudio de evaluación y sus resultados se presentan en la Sección V, seguidos de una discusión sobre las implicaciones y limitaciones del estudio en la Sección VI. Por último, resumimos las principales conclusiones y las líneas de trabajo futuro en la Sección VII.

II. TRABAJO RELACIONADO

Los primeros investigadores en el área de la analítica del aprendizaje (LA) propusieron herramientas conceptuales (e.g., marcos conceptuales [23] y modelos de referencia [24]) que sintetizaban la experiencia obtenida en los primeros proyectos piloto. Tales trabajos fueron de gran ayuda para la comunidad de LA, ya que definieron los pasos a seguir (e.g., [24]) y los aspectos a tener en cuenta (e.g., [23]) en un proyecto de LA. Además, estas herramientas conceptuales permitieron identificar las dificultades más importantes encontradas para así proponer posibles soluciones. Tal fue el caso de los problemas éticos al administrar los datos personales de los procesos de aprendizaje, los cuales se plantearon como problemas importantes sin soluciones claras [23] y luego fueron progresivamente atajados por varias propuestas subsiguientes (e.g., [25]).

Estos trabajos de síntesis condujeron a proyectos de investigación en LA más avanzados y maduros, así como a las primeras muestras de adopción en instituciones educativas [26]. Actualmente, la comunidad investigadora sigue trabajando en añadir la experiencia obtenida en proyectos de LA y en su despliegue en escuelas y universidades. Muestra de ello son proyectos como LACE ¹, SHEILA ² o LALA ³. Gracias a estos esfuerzos, actualmente contamos con herramientas conceptuales mucho más maduras dirigidas a investigadores, profesorado, desarrolladores de software y responsables institucionales [27], [28] sobre cómo diseñar y poner en marcha proyectos de LA. Su impacto positivo puede apreciarse en el aumento de la adopción de LA por parte de las instituciones educativas durante los últimos años [28]. Sin embargo, estos trabajos se centran principalmente en procesos de LA que involucran una sola fuente de datos, quedando aún por cubrir la complejidad de las soluciones de LA que involucran varias fuentes.

Es posible encontrar cadenas de valor dirigidas a los datos masivos, las cuales proporcionan herramientas conceptuales que abstraen la complejidad de la heterogeneidad de los datos. Aunque en LA no está extendido el uso de DVCs, es posible encontrar algunas propuestas. Tal es el ejemplo del DVC propuesto por Ruiz-Calleja et al. [29], el cual, partiendo del análisis de varios procesos LA, fue usado para guiar a los actores interesados a la hora de trasladar los procesos de LA a una escala nacional. De forma similar, en nuestro

¹<http://www.laceproject.eu/>

²<http://sheilaproject.eu/>

³<https://www.lalaproject.org/>

trabajo previo usamos un DVC existente [12] para guiarnos en el análisis de la literatura acerca de infraestructuras para MMLA [17] y en el modelado del procesamiento de evidencia multimodal en cuatro escenarios MMLA [22]. De los trabajos mencionados se pueden extraer las siguientes conclusiones: en primer lugar, ninguna de las arquitecturas MMLA revisadas cubre todos los pasos definidos en el DVC [17]; y, segundo, que el DVC que empleamos (específico para datos masivos) no respondía a las necesidades de los procesos de MMLA [22]. De ahí la necesidad de definir un DVC especializado para el dominio MMLA que tenga en cuenta los siguientes aspectos:

1. Problemas en la toma de decisiones: la mayoría de los desarrolladores de tecnologías educativas no tienen experiencia en el procesamiento de evidencia multimodal. Sin embargo, deben decidir qué actividades y pasos del procesamiento de datos formarán parte del proceso de análisis [15]. Además, deben adaptar sus decisiones de diseño habituales, que se basan en su experiencia pasada en herramientas monomodales, al desarrollo de MMLA [14]. Finalmente, necesitan incluir información contextual sobre el proceso de aprendizaje en el análisis, ya que de lo contrario el análisis de los datos puede conducir a interpretaciones erróneas [18], [19].
2. Problemas en la sistematización del proceso: no es trivial integrar datos heterogéneos así como su información contextual en el procesamiento de datos [16]. Algunas soluciones existentes proponen dividir esta complejidad en módulos, cada cual encargado de una actividad de procesamiento de datos de la DVC [17]. Sin embargo, para beneficiarse de tales propuestas a la hora de desarrollar los módulos, sería necesario seguir el enfoque modular del desarrollo de software y clasificar cada módulo en función de la información de entrada, de salida y de procesamiento [22].
3. Problemas de comunicación: para satisfacer las necesidades de los usuarios es necesario que estos comuniquen los requisitos de la solución MMLA a los desarrolladores. Por un lado, la comunicación entre personas con diferente perfil (educativo *versus* técnico) no es siempre sencilla debido a la falta de experiencia en la especificación de requisitos de MMLA [30] y a la brecha terminológica entre ambos [20]. Por otro lado, el tiempo y recursos disponibles para desarrollar una solución MMLA es limitado [16]. Por ello, es necesario facilitar una comunicación efectiva que permita una comprensión clara de los requisitos extraídos y de los resultados esperados para ambas partes [30].

III. LA CADENA DE VALOR DE DATOS MULTIMODALES (M-DVC)

Con el fin de proponer una Cadena de Valor Multimodal (M-DVC por sus siglas en inglés, *Multimodal Data Value Chain*), nos apoyamos en una cadena de valor multimodal ya existente y de amplio uso [12]. Proponemos una extensión de esta DVC para superar sus limitaciones al ser aplicada a un proceso MMLA. Más en concreto, en nuestra investigación previa [17], detectamos tres características del campo MMLA que deberían ser tenidas en cuenta:

1. MMLA trabaja con datos educativos, en vez de con datos procedentes de negocios, y cada decisión debe ser apoyada en teorías educativas [31].
2. MMLA necesita atender los requisitos de múltiples partes interesadas, que cubren un espectro amplio de la jerarquía educativa, desde legisladores a estudiantes. Además, para poder desarrollar soluciones MMLA, se necesita que estos actores procedentes de un rango tan amplio se comuniquen con los actores de perfil técnico [20].
3. MMLA implica escenarios de aprendizaje heterogéneos. Estos escenarios son diversos entre sí en muchos aspectos, como el número de participantes, tipos de actividades de aprendizaje, espacios de aprendizaje, propósitos y razones para usar MMLA y modalidades.

Como primer paso en la propuesta de la M-DVC, analizamos cuatro casos reales MMLA, que se resumen en la Tabla I. Analizamos su heterogeneidad y diversidad para entender las especificidades de MMLA. Como resultado de estos análisis, extrajimos las siguientes características que distinguen los procesos MMLA de otros procesos de Big Data que la DVC pretende apoyar.

1. Necesidad de definir metas, modos y modalidades: El procesamiento de datos en Big Data y Minería de Datos tienen sentido por las siguientes razones: 1. No hay una meta concreta para el análisis, más allá de la minería aleatoria para encontrar patrones y tendencias y 2. Los datos crudos son históricos y no fueron recogidos con el mismo propósito que el del análisis. Sin embargo, en MMLA, los actores educativos tienen una meta clara y específica para el uso de MMLA. Esta meta está basada en la teoría educativa y la pedagogía subyacente del contexto educativo. Los actores educativos definen múltiples modos en los que el proceso de aprendizaje tiene lugar. Es más, los actores educativos y técnicos (personal de apoyo TIC en la escuela o expertos en tecnología educativa) necesitan hablar acerca de la configuración del escenario de aprendizaje para relacionar los modos en modalidades que puedan ser registradas para generar evidencia multimodal del aprendizaje. Por tanto, la DVC especializada debería tener en cuenta estos aspectos.
2. Inclusión de la información de contexto: A diferencia del dominio empresarial, la educación se refiere a personas, y las prácticas educativas son prácticas cognitivas. Por tanto, hay algunos eventos que no pueden ser planeados de antemano. Por ejemplo, la actividad de aprendizaje fue planificada para 12 estudiantes, pero dos estudiantes cayeron enfermos y no pudieron unirse a la hora de la puesta en marcha de la actividad de aprendizaje. Por tanto, los datos educativos suponen la existencia de información contextual que incluye metadatos desde las fases de planificación a la puesta en marcha de las actividades de aprendizaje. El análisis directo de la evidencia multimodal del aprendizaje sin incluir esta información contextual podría llevar a las distintas partes interesadas a resultados erróneos. Por tanto, la DVC

Cuadro I
 CONTEXTO DE APRENDIZAJE Y CARACTERÍSTICAS MULTIMODALES DE LOS CUATRO ESCENARIOS USADOS PARA PARTICULARIZAR LA DVC AL
 CONTEXTO DE MMLA.

Escen.	Contexto de aprendizaje			Analítica de aprendizaje multimodal	
	Participantes	Tipo de curso	Espacios	Propósito	Fuentes de datos
1	8 Profesores 3 Investigadores 20 Estudiantes	Actividades mixtas (día puertas abiertas)	Digital + Físico	Analizar el impacto de la innovación docente en la implicación grupal	logs de Graasp Observación estructurada
2	1 Profesor 1 Investigador 150 Estudiantes	Curso en línea	Digital	Estudiar el efecto de la gamificación en la implicación del estudiante	logs del LMS logs de herramientas externas logs de la plataforma de gamificación
3	2 Profesores 20 Estudiantes	Actividad de búsqueda del tesoro	Digital + Físico	Adaptar la experiencia de aprendizaje del estudiante en tiempo real	Logs de la App Observación estructurada Sensores
4	1 Instructor 1000 Estudiantes	MOOC	Digital	Ayudar al instructor a identificar estudiantes en apuros en un curso	Posts en el foro del curso Problemas reportados por los estudiantes

especializada debería tener en cuenta esta especificidad del MMLA.

3. Rendir cuentas a los actores educativos en la fase de toma de decisiones: Los actores educativos deben tomar decisiones pedagógicas basadas en los resultados analíticos en la actividad de toma de decisiones de la DVC. Esas decisiones pedagógicas deben repercutir en las prácticas educativas para poner en marcha las acciones necesarias. Esas acciones pueden lanzar cambios para la mejora de dichas prácticas. Esto podría llevar a los actores educativos a definir nuevas metas.
4. La secuencia de actividades de procesamiento de datos no es lineal: A diferencia de la DVC genérica existente, las actividades de procesamiento de datos multimodales a menudo no son lineales y unidireccionales. La principal razón detrás de esta diferencia es que en la mayoría de los casos MMLA los actores exploran los datos (debido a su escasa experiencia y la necesidad de hacer tratamiento intensivo de los datos de MMLA). Durante dicha exploración, los actores saltan a menudo de una actividad de procesamiento de datos a otra anterior.

Incorporamos estas cuatro lecciones aprendidas en la DVC existente con el fin de proponer una DVC especializada para MMLA (ver Figura 1), llamada DVC multimodal (M-DVC). En primer lugar, incorporamos la primera lección ‘necesidad de definir objetivos, modos y modalidades’, añadiendo tres nuevos pasos, 1, 2, y 3 en la M-DVC (cajas azules en la Figura 1). La segunda lección fue incorporada en la M-DVC añadiendo una conexión entre las fases 1 y 5, lo que lleva la información contextual al procesamiento de la evidencia de aprendizaje multimodal. La tercera lección fue aplicada añadiendo una flecha en el paso 10, que retroalimenta la fase 1 con nuevas metas pedagógicas (ver la flecha cerrada entre la fase 10 y la 1 en la Figura 1, y otra flecha desde el paso 10 que lleva a los cambios en la práctica existente (si se necesita), basada en la evidencia del análisis. Finalmente, incorporamos la cuarta lección aprendida añadiendo flechas recursivas desde el paso 5 al 10, lo que permite saltar de una actividad de procesamiento de datos a otra.

IV. METODOLOGÍA

La presente propuesta es parte de un proyecto más amplio basado en una metodología de investigación basada en diseño (en inglés, *Design-Based Research*, o DBR) [32], con el objetivo de dar soporte al desarrollo de soluciones MMLA. Concretamente, proponemos que la cadena de valor M-DVC descrita anteriormente podría ser una herramienta conceptual útil para apoyar la comunicación entre las partes interesadas durante dichos procesos de desarrollo. Así, para investigar esta afirmación en una primera iteración, propusimos un estudio para responder a la siguiente pregunta de investigación:

¿Hasta qué punto es la M-DVC propuesta útil como herramienta conceptual para apoyar la comunicación de partes interesadas provenientes de múltiples disciplinas, durante el desarrollo de soluciones MMLA?

Esta pregunta de investigación principal ha sido iluminada desde la perspectiva de tres temas principales, relacionados con los desafíos encontrados en la literatura sobre el desarrollo de soluciones MMLA (véase Sección II). Así, nuestra evaluación aborda la utilidad de la M-DVC para dar soporte a las decisiones de diseño del sistema, su mayor sistematización, y como apoyo para la comunicación entre las diferentes partes interesadas (provenientes, típicamente, de campos diversos). Estos tres temas se descomponen a su vez en subtemas (véase Figura 2).

En esta iteración, evaluamos nuestra propuesta siguiendo una aproximación de métodos mixtos [33] en tres casos reales de proyectos de desarrollo de soluciones MMLA, en la Universidad de Tallín (Estonia). La Tabla II resume estos tres casos. Involucramos a dos personas de cada uno de estos proyectos: una con un perfil más técnico (e.g., desarrollador o analista *software*) y otra con un perfil más educativo (e.g., docente o investigador educativo). En adelante, denotaremos estas parejas de partes interesadas simplemente con el término ‘parejas’ (o sea, seis participantes, formando tres parejas, una por cada caso/proyecto). La idea básica de la evaluación consistió en pedir a cada pareja que discutiera su proyecto MMLA para especificar los requisitos de la solución MMLA concreta en cada caso (como primer paso del proceso de

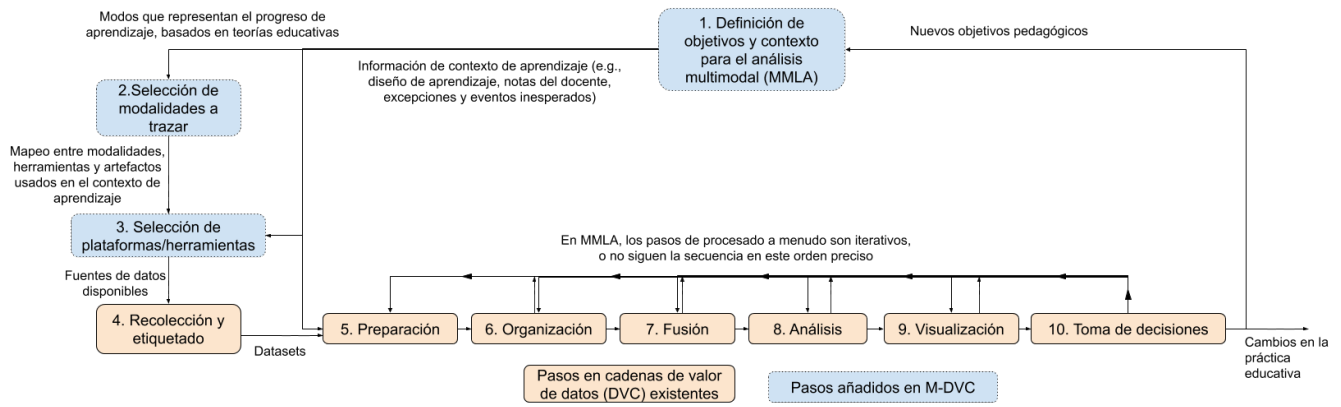


Figura 1. Cadena de valor de datos multimodales (Multimodal Data Value Chain, M-DVC) propuesta.

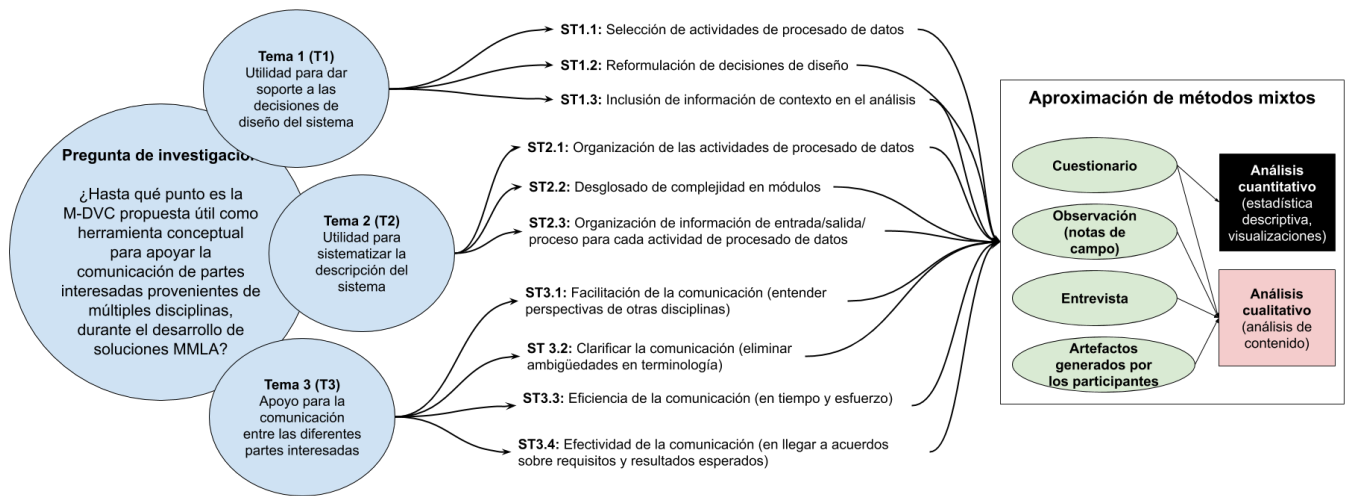


Figura 2. Pregunta de investigación, temas, subtemas investigados, y aproximación de métodos mixtos usada durante la evaluación

desarrollo), usando la M-DVC como herramienta conceptual para guiar la conversación.

Para fomentar que los participantes utilicen la M-DVC cuando se comunican, desarrollamos un instrumento (una guía de reflexión y conversación, implementada utilizando el servicio de hojas de cálculo Google Spreadsheet), que facilita el uso de la herramienta conceptual y sus conceptos relacionados (i.e., detallando las actividades de procesamiento de datos involucradas y concretando los pasos involucrados en cada actividad). Además, dos de los autores de este artículo llevaron a cabo un ejercicio similar (i.e., una conversación y relleno del formulario) sobre un escenario/sistema de MMLA preciso, de manera que se ofrecía un ejemplo de uso de dicho instrumento⁴ para que los participantes lo usasen como referencia, en caso de ser necesario. Para cada caso estudiado, se llevó a cabo una entrevista presencial semi-estructurada con la pareja de participantes, en la que el entrevistador (uno de los autores) presentó en cinco minutos la M-DVC, una copia del instrumento sin relleno, y el ejemplo. Una vez que los

participantes proporcionaron oralmente su consentimiento para participar en la investigación y para ser grabados, la pareja de participantes discutió sobre cómo debería funcionar la propuesta de MMLA, siguiendo y relleno de las distintas fases del instrumento de M-DVC (este ejercicio se prolongó durante 90, 80 y 150 minutos en los casos 1, 2 y 3 respectivamente). Durante la discusión, el entrevistador actuó como observador, tomando distintas notas y resolviendo cualquier duda que los participantes tuvieran sobre el instrumento.

Tras este ejercicio de discusión, se presentó un cuestionario a cada participante⁵ que incluía 10 preguntas para responder mediante una escala de Likert (las respuestas variaban entre 1–Totalmente en desacuerdo, a 5–Totalmente de acuerdo) relacionadas con los 10 subtemas (ST) de nuestra evaluación (véase la Figura 2). El cuestionario también incluía tres preguntas abiertas que consultaban de manera más genérica sobre las razones por las que los participantes habían puntuado cada uno de los temas (T). Una vez que estos cuestionarios se relleno, el entrevistador pidió aclaraciones o hizo preguntas que consultaban sobre las respuestas introducidas en el cuestionario y

⁴El ejemplo de uso se encuentra en la siguiente dirección: <https://tinyurl.com/yylf5wud>.

⁵El cuestionario está disponible en: <https://tinyurl.com/y6y2vya2>.

Cuadro II
DESCRIPCIÓN DE LOS PROYECTOS MMLA USADOS EN LA EVALUACIÓN

Caso	Rol del participante 1	Rol del participante 2	Objetivo(s) de investigación del caso	Contexto educativo	Objetivo del análisis multimodal	Estado actual del proyecto (a nivel educativo)	Estado actual del proyecto (a nivel técnico)
Caso 1	Investigador (educativo)	Desarrollador (técnico)	¿Es posible detectar patrones significativos de colaboración y dominación utilizando audio y trazas de software (logs)?	Trabajo en grupo dentro de clase en educación secundaria	Visualizar las contribuciones individuales y el dominio de sus integrantes al grupo	Varios experimentos piloto llevados a cabo con un prototipo	Prototipo usable disponible
Caso 2	Investigador (educativo)	Analista de software (técnico)	Ayudar al profesor comprender el desarrollo de las actividades en clase en entornos donde se utiliza tecnología, incluyendo espacios físicos y digitales	Profesores de primaria hacen preguntas sobre sus práctica docente utilizando recursos de aprendizaje digitales	Visualizar datos multimodales en clase con una herramienta de monitorización	Recursos de aprendizaje en uso	Especificación
Caso 3	Investigador (educativo)	Desarrollador (técnico)	Cómo un profesor e investigador puede comprender/comparar estrategias docentes en términos del aprendizaje (ganancia, adecuación de las tareas, tiempo empleado) y retroalimentación de los alumnos?	Desarrollo profesional del profesorado en colegios de primaria/secundaria (sobre el impacto de la práctica docente en el éxito de los alumnos)	Visualizar/Comparar datos multimodales de distintas lecciones en una herramienta de monitorización que soporta la auto-reflexión	Varios experimentos llevados a cabo para recoger datos	Especificación

las razones de las mismas, así como las impresiones generales del uso de la herramienta conceptual.

Se analizaron los datos cuantitativos de los cuestionarios utilizando estadísticas descriptivas y visualizaciones simples, con el propósito de encontrar tendencias en las opiniones de los participantes (ya que, dado el pequeño número de la muestra, inferencias estadísticas no sería apropiado). Mas la parte central del análisis tenía una naturaleza cualitativa, consistiendo en el análisis de contenido de las respuestas dadas a los cuestionarios de preguntas abiertas, los aspectos relevantes de las entrevistas y las notas tomadas por el entrevistador, codificados según los temas y subtemas de la evaluación. Los artefactos generados por los participantes (i.e., los artefactos de M-DVC que rellenaron durante las entrevistas) se utilizaron para interpretar las declaraciones en los audios grabados y en los cuestionarios.

V. EVALUACIÓN

Las Tablas III, IV y V presentan nuestras principales conclusiones y las evidencias que las soportan en relación a cada tópico y sub-tópico de la evaluación. Las evidencias cualitativas de nuestras conclusiones se etiquetan utilizando el siguiente esquema: '[Cn.I/D/AS]'. En este esquema, 'C' quiere decir 'Caso'; 'n', 'número del caso de estudio'; 'I' 'Investigador'; 'D' 'Desarrollador'; y 'AS', 'Analista de Software' (para más detalles sobre los casos de estudio, véase la Tabla II). Por ejemplo, si hay una evidencia cuyo código sea [C2.D] significa que la evidencia la proporcionó el **Desarrollador del Caso 2**. Utilizamos histogramas para representar las evidencias cuantitativas en cada uno de los subtemas, para dar una idea de la distribución de las respuestas de los participantes. Más abajo resumimos las conclusiones para cada tópico y sub-tópico de la evaluación.

Utilidad de la M-DVC como una herramienta de soporte a las decisiones (T1): Al preguntarles sobre cómo la M-DVC daba soporte a la selección de actividades de procesamiento de datos, incluyendo información contextual en el análisis y

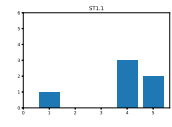
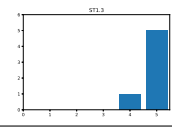
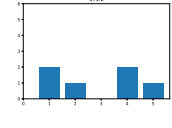
formulando decisiones de diseño, los participantes indicaron lo siguiente:

- Selección de actividades de procesamiento de datos (ST1.1) A pesar de que consideraron a la M-DVC útil para seleccionar las actividades de procesamiento de datos (5 de 6 participantes estaban de acuerdo), parece que los desarrolladores lo encuentran más útil, teniendo en cuenta que ellos son más expertos en muchas de las operaciones de pre-procesamiento de datos que ocurren antes del análisis de datos en sí [C3.I, C3.D].
- Reformular decisiones de diseño (ST1.2) La M-DVC es útil para reformular decisiones de diseño (5 de 6 participantes están totalmente de acuerdo y 1 está de acuerdo). Sin embargo, los participantes no acabaron esta tarea del ejercicio. Por ello no forjaron una opinión sobre este aspecto, ya que no discutieron mucho sobre las decisiones de diseño.
- Incluir información contextual en el análisis de MMLA (ST1.3) Los participantes ofrecieron respuestas diversas (2 estaban muy en desacuerdo, 1 en desacuerdo, 2 de acuerdo y 1 totalmente de acuerdo) sobre la utilidad de la M-DVC en incluir información contextual en MMLA. Mencionaron que la M-DVC en su versión gráfica representa la idea pero que cuando los participantes comienzan su discusión y el ejercicio de rellenar el cuestionario, olvidan esa idea [C3.D]. Hay una brecha de traducción desde la herramienta conceptual al instrumento que no permite informar explícitamente de esta idea. Además, la carga cognitiva es alta durante la discusión y ellos no se preocupan de leer las instrucciones del instrumento detenidamente, donde la idea se presenta brevemente [C3.I].

Con todo, podemos observar que a la M-DVC se la ha considerado por lo general útil como una herramienta para ayudar a la toma de decisiones, en términos de qué tipos de mecanismos de procesamiento de datos deben incluirse, y en qué orden [C2.AS]. También se reportó el aumento de la consciencia sobre las actividades de procesamiento de datos

Cuadro III

TEMA 1, SUBTEMAS, CONCLUSIONES Y EVIDENCIAS ILUSTRATIVAS DE TRES ESTUDIOS DE CASO. LAS ETIQUETAS EN CORCHETES DENOTAN EL CASO/PARTICIPANTE DE QUIEN SE EXTRAJO LA CITA

Tópico / Sub-tópico	Conclusión	Evidencia cualitativa	Evidencia cuantitativa
T1. Utilidad de la M-DVC como una herramienta de soporte a las decisiones	M-DVC es útil como una herramienta para la toma de decisiones, pero es necesario enfocarse explícitamente en el aspecto 'inclusión de la información contextual en el análisis', pues llama la atención de los participantes.	* 'El soporte de la información contextual puede ser útil.' [C1.D] * 'Esta herramienta es útil en mi camino cognitivo para tomar decisiones relacionadas con procesamiento de datos multimodales.' [C2.I] * 'DVC es muy útil en la toma de decisiones para nuestro caso y me informa sobre la secuencia de actividades de procesamiento de datos.' [C2.AS] * 'Me hizo más consciente de los aspectos relacionados con el procesamiento de datos.' [C3.I] * 'Fue MUY útil para desenterrar decisiones de diseños u operaciones de las que los participantes quizá no fuesen conscientes.' [C3.D]	(véase los histogramas de más abajo)
ST1.1 Selección de actividades de procesamiento de datos	El perfil de desarrollador tiene más conocimiento de las actividades de procesamiento de datos en general, pero los pasos requeridos en cada una de esas actividades deben ser discutidas con el perfil de investigador	* 'La DVC me hizo más consciente de las actividades de procesamiento de datos. Hasta ahora mis decisiones y mi contribución en la conversación eran muy dependientes de las preguntas de los desarrolladores.' [C3.I] * 'Yo era consciente de la DVC y del instrumento de antemano, así que era un añadido para mí. Utilicé el conocimiento que ya tenía para extraer los requisitos por parte de los investigadores, pero el soporte de la herramienta como una herramienta conceptual común durante la discusión fue útil.' [C3.D]	
ST1.2 Reformular decisiones de diseño	Los participantes parecen estar de acuerdo en que la M-DVC les ayudó a reformular decisiones de diseño		
ST1.3 Incluir información contextual en el análisis de MMLA	El traspaso de la M-DVC al instrumento tuvo algún problema y no ofreció el aspecto 'inclusión de la información contextual en el análisis' de manera clara para los participantes	* 'La información contextual no es necesaria solo en la preparación de los datos, pero también en otras actividades de procesamiento de datos. El investigador necesita discutir sobre la información contextual en casos en los que no hay ningún diseño educativo' [C1.D] * 'DVC y el instrumento han intentado incluir la información contextual en el análisis pero no han conseguido impulsar la idea - para incluir la información contextual en el análisis.' [C3.D]	

por parte de los participantes educativos, quienes tenían una cultura de datos limitada en la mayor parte de los casos [C3.I]. Además, como herramienta conceptual, M-DVC ayuda en el camino cognitivo de los participantes para reflexionar sobre sus decisiones [C2.R]. Sin embargo, M-DVC no satisface los requerimientos de incluir la información contextual en el análisis en su actual forma en la que ni la herramienta ni el instrumento alentaron estos aspectos explícitamente durante las discusiones [C3.D].

Utilidad de la M-DVC como una herramienta de sistematización (T2): Los datos recogidos sobre este tópico arrojan los siguientes resultados:

- Organizar las actividades de procesamiento de datos (ST2.1) Se consideró la M-DVC como una herramienta conceptual útil para estructurar y organizar actividades de procesamiento de datos en el proceso de desarrollar soluciones MMLA (3 de 6 participantes están de acuerdo con esta afirmación, mientras que los otros 3 están totalmente de acuerdo) [C1.D].
- Descompone en partes la complejidad que supone el desarrollo de una solución MMLA (ST2.2) Sobre este punto, los seis participantes están de acuerdo con que M-DVC es una herramienta conceptual útil para descomponer en partes la complejidad que supone el desarrollo de una solución MMLA [C1.D, C2.I].
- Necesidad de incluir información sobre organizar información, procesar y salida de datos para cada actividad de procesamiento de datos (ST2.3) Tanto a la M-DVC como al instrumento se les consideraron útiles para detectar la información relevante sobre las actividades de procesamiento de datos en un proceso de MMLA y discutir sobre ello. Solo un participante no estuvo de acuerdo ni en desacuerdo con este punto. Además, los participantes mencionaron la necesidad de clasificar los requisitos de cada actividad de procesamiento de datos

bajo las etiquetas de Entrada/Procesamiento/Salida de información [C1.D].

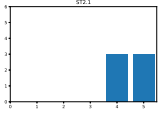
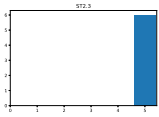
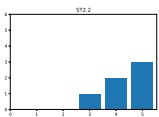
En suma, se ha considerado que M-DVC es útil como una herramienta de sistematización en el proceso de desarrollar soluciones MMLA. Sin embargo, como MMLA no es una tecnología muy diseminada todavía, algunos participantes sin experiencia previa en MMLA consideraron que utilizar M-DVC conlleva mucho tiempo [C2.I]. Además, dos desarrolladores subrayaron que ellos hubieran seguido una secuencia diferente para procesar evidencias multimodales de aprendizaje en escenarios reales sin no hubieran sido conscientes de esta herramienta conceptual [C1.D, C3.I].

Utilidad de la M-DVC como herramienta de comunicación (T3): Los resultados relacionados con este tópico son los siguientes:

- Comunicación clara (ST3.1). Se consideró que la M-DVC facilita la comunicación (4 de 6 participantes están totalmente de acuerdo mientras que los otros 2 no están ni de acuerdo ni en desacuerdo) cuando ambos participantes utilizaron la misma herramienta conceptual durante la discusión para especificar los requisitos. Los participantes dijeron que la herramienta les ayudo incluso en el caso en el que el científico educativo no tenía experiencia en expresar los requisitos de software a los desarrolladores [C2.R].
- Comunicación efectiva (ST3.2). Sobre este aspecto la M-DVC no alcanzó las expectativas de ser útil como una herramienta conceptual para una comunicación clara (3 de 6 participantes no están ni de acuerdo ni en desacuerdo, 2 están de acuerdo y otros 2 están totalmente de acuerdo). Los desarrolladores son quienes dudan especialmente sobre el propio término de 'comunicación clara' en el contexto de especificaciones de requisitos [C3.D].
- Comunicación eficiente (ST3.3). Concluimos que durante la discusión la carga cognitiva era considerada alta para

Cuadro IV

TEMA 2, SUBTEMAS, CONCLUSIONES Y EVIDENCIAS ILUSTRATIVAS DE TRES ESTUDIOS DE CASO. LAS ETIQUETAS EN CORCHETES DENOTAN EL CASO/PARTICIPANTE DE QUIEN SE EXTRAJO LA CITA

Tópico / Sub-tópico	Conclusión	Evidencia cualitativa	Evidencia cuantitativa
T2. Utilidad de la M-DVC como una herramienta de sistematización	La M-DVC es útil para el propósito de la sistematización, pero requiere tiempo para los participantes inexpertos y es necesario que se enfoque más en la información relacionada con E/S/P (Entrada/Salida/Procesamiento) para los aspectos de cada una de las actividades de procesamiento de datos	* 'Por parte de la DVC, la visualización está después de la fusión pero los casos en los que los investigadores exploran los datos recogidos la secuencia puede ser diferente, como tener una visualización preliminar antes de la fusión.' [C1.D] * 'Esta DVC es útil para nuestro caso para organizar y disponer las actividades de procesamiento de datos. Pero este ejercicio requiere mucho tiempo y es complejo de responder, especialmente en nuestro caso porque no tenemos experiencia previa.' [C2.I] * 'Puede que tenga una secuencia de actividades de procesamiento de datos diferentes para abordar mis necesidades, pero seguir este tipo de herramienta conceptual estandarizada puede ser una buena práctica para el propósito de la sistematización.' [C3.I] * 'La estructura de E/S/P no es muy explícita.' [C3.D]	(véase los histogramas de más abajo)
ST2.1 Organizar las actividades de procesamiento de datos	La M-DVC como una herramienta conceptual permite la organización práctica de las actividades de procesamiento de datos, así como el pensamiento cognitivo de los participantes	* 'A pesar de que seguimos diferentes actividades de procesamiento de datos para abordar nuestros requisitos, esta herramienta nos ayuda a estructurar y organizar las actividades de procesamiento de datos y, además, alimenta nuestra repetición cognitiva.' [C1.D]	
ST2.2 Descompone en partes la complejidad que supone el desarrollo de una solución MMLA	El M-DVC maneja este aspecto correctamente con su estructura modular para presentar todos los conceptos relacionados en el procesamiento de evidencias de aprendizaje multimodales	* 'A pesar de que nuestra aproximación no era modular en el pasado, esta herramienta me da una mejor visión para trabajar de manera modular en el desarrollo de soluciones MMLA.' [C1.D] * 'Ayuda a descomponer el problema grande en pequeños problemas para ayudar a nuestro pensamiento conceptual.' [C2.I]	
ST2.3 Es necesario incluir información sobre organizar información, procesar y salida de datos para cada actividad de procesamiento de datos	Las fases de M-DVC ayudan a los participantes a extraer los requisitos para cada una de las actividades de procesamiento de datos, pero es necesario clasificar los requisitos en las categorías de información relacionada con E/S/P para cada actividad.	* 'Como herramienta, la DVC es útil para mí para extraer los requisitos del investigador para cada una de las actividades de procesamiento de datos.' [C1.D]	

ambos participantes. Así, 4 de cada 6 participantes no está de acuerdo ni en desacuerdo cuando se le pregunta si la M-DVC es útil como una herramienta conceptual para una comunicación eficiente en el proceso de desarrollar una solución MMLA.

- Facilita la comunicación (ST3.4). Se consideró que la M-DVC es útil como una herramienta conceptual para la comunicación efectiva [C2.AD], donde 4 de los 6 participantes están totalmente de acuerdo con ello mientras que 1 está de acuerdo y 1 ni está de acuerdo ni en desacuerdo.

Globalmente, los participantes encuentran el modelo M-DVC útil para facilitar la comunicación sencilla y eficiente durante la especificación de requisitos para el desarrollo de una solución MMLA. Esto es cierto para todos los participantes, especialmente los de perfil educativo que tenían una limitada cultura de datos. Sin embargo, no es tan efectivo como se esperaba en la claridad y la eficiencia de la comunicación debido a la versión actual de la M-DVC y porque el instrumento actual utiliza términos técnicos que no son sencillos de interiorizar para la mayoría de los participantes de perfil educativo.

Es interesante resaltar aquí que todos los participantes cuyo rol era de investigador en los tres casos estaban de acuerdo o totalmente de acuerdo con T3 excepto en el aspecto de la eficiencia del caso 3. Los desarrolladores consideraron que la herramienta todavía no es útil como una herramienta de comunicación, especialmente con los aspectos relacionados con la comunicación *clara y eficiente*. El aspecto mejor valorado en T3 es 'facilita la comunicación', en el que cuatro de seis participantes estaban totalmente de acuerdo con la afirmación. En este sentido, los investigadores del caso 2 mencionaron que 'es muy sencilla mi comunicación con el analista de sistemas. Pero todavía no estoy familiarizado con los términos y la

terminología utilizada en la herramienta. Es un vocabulario totalmente nuevo para mí'.

VI. DISCUSIÓN

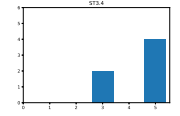
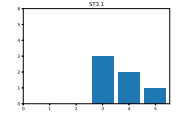
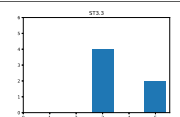
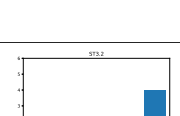
La sección anterior muestra los resultados de los tres temas definidos en la sección IV. En esta sección presentamos los principales resultados y sus implicaciones para el campo del MMLA.

El análisis de nuestros escenarios muestra la necesidad de una atención explícita en la inclusión de información contextual para evitar resultados analíticos confusos [T1]. La idea de incluir información contextual en el análisis es buscada por muchos investigadores [19], [18]. En la mayoría de los escenarios de aprendizaje, la información contextual, o bien no se guarda o bien se guarda en documentos relacionados con la orquestación del aprendizaje, como el diseño de aprendizaje o las notas de los profesores. Además, en aquellos casos en los que hay una recogida explícita de estos documentos, éstos consisten habitualmente en texto sin estructura, muy difícil de procesar por un sistema técnico. Esto impide la inclusión de la información contextual en el análisis de aprendizaje, e incrementa la complejidad de tenerlo en cuenta para las distintas partes interesadas. Por tanto, necesitan un impulso externo en su práctica habitual a través de una herramienta conceptual, como M-DVC, que puede llamar la atención de forma explícita sobre este aspecto.

Hemos visto también que no hay una secuencia lineal clara de las actividades en un M-DVC [T1]. En la mayoría de los proyectos MMLA, las distintas partes interesadas exploran las bases de datos heterogéneas para entender la evidencia multimodal sobre el aprendizaje [16]. En esta exploración, los actores pueden realizar actividades de procesamiento de

Cuadro V

TEMA 3, SUBTEMAS, CONCLUSIONES Y EVIDENCIAS ILUSTRATIVAS DE TRES ESTUDIOS DE CASO. LAS ETIQUETAS EN CORCHETES DENOTAN EL CASO/PARTICIPANTE DE QUIEN SE EXTRAJO LA CITA

Tópico / Sub-tópico	Conclusión	Evidencia cualitativa	Evidencia cuantitativa
T3. Utilidad de la M-DVC como herramienta de comunicación	M-DVC es útil para la comunicación sencilla y efectiva incluso si los participantes tienen una cultura de datos limitada, pero hay una necesidad de simplificar los términos técnicos de manera que los participantes no técnicos puedan formar parte en el desarrollo de las soluciones MMLA utilizando M-DVC como herramienta conceptual.	* 'Verdaderamente facilita mi comunicación con el analista de sistemas. Pero todavía no estoy familiarizado con los términos y las terminologías utilizados en esta herramienta. Todo este vocabulario es nuevo para mí.' [C2.I] * 'Cada vez que se involucra a humanos es difícil minimizar todos estos problemas relacionados con la comunicación. La idea es buena pero no estoy seguro cómo funcionaría en situaciones reales (entre investigadores y desarrolladores).' [C2.AS] * 'Nos ayudó a negociar el significado entre nosotros, incluso teniendo una cultura de datos limitada.' [C3.I] * 'Creo que la terminología todavía está demasiado orientada hacia aspectos técnicos. Hay ambigüedades (e.g., operaciones que pueden ser clasificadas como organización y como preparación). Creo que es una herramienta eficiente, no por el resultado final sino por las discusiones que ayudó a llevar a cabo.' [C3.D]	(véase los histogramas de más abajo)
ST3.1 Facilita la comunicación al seguir una sola herramienta conceptual	La M-DVC facilita la comunicación porque ambos participantes utilizan la misma herramienta conceptual, i.e., la M-DVC.	'Nunca había pensado en una herramienta conceptual para el procesamiento de datos. Tengo miedo de hablar con personal técnico cuando discuto de requisitos para las herramientas que solicito. Utilizando esta DVC, me siento como un científico educativo que puede hablar con el personal técnico.' [C2.I]	
ST3.2 Aclara la comunicación sin ambigüedades en la comprensión de términos y terminologías debido a haber involucrado distintos perfiles	Los términos utilizados en M-DVC están más inclinados hacia los aspectos técnicos que a los aspectos educativos de MMLA	* 'No estoy seguro del aspecto relacionado con la comunicación clara porque hay muchas definiciones y parámetros involucrados en una comunicación clara.' [C3.D]	
ST3.3 Comunicación eficiente sin malgastar tiempo y coste para alcanzar los resultados esperados	Los investigadores educativos no pueden juzgar los aspectos relacionados con la eficiencia de la comunicación proporcionados por la M-DVC ya que requiere tiempo y la carga cognitiva es alta durante la discusión.	'El ejercicio que incluye la discusión y rellenar el instrumento requiere mucho tiempo, especialmente en nuestro caso, en el que no tenemos ninguna experiencia previa con MMLA.' [C2.I]	
ST3.4 La comunicación efectiva acordando los resultados y los requisitos esperados	La M-DVC es útil para la comunicación efectiva en general, incluso en los casos en los que los participantes no tienen experiencia técnica o educativa relacionada con MMLA	'Utilizando esta herramienta podemos negociar las expectativas y los requisitos. Antes, era casi imposible. Pero todavía soy escéptico sobre si esto funcionaría en cualquier situación.' [C2.AS]	

datos en secuencias diferentes a la presentada en el M-DVC. Por ejemplo, podrían necesitar visualizar cada una de las bases de datos antes de la actividad de fusión de datos. Sin embargo, una vez la exploración está terminada, los actores podrían seguir una secuencia estándar de actividades de procesamiento de datos, como el M-DVC.

Es necesario también mejorar la estructuración y organización de información relacionada con la especificación de entradas/procesado/salidas (E/P/S) para cada actividad de procesamiento de datos [T2]. En este sentido, el M-DVC es una herramienta útil para la estructuración y extracción de requisitos de una solución MMLA. Este es el caso especialmente cuando varias partes interesadas están involucradas, porque puede servir para guiar sus deliberaciones. El tiempo que se necesita para rellenar el instrumento, y la cantidad de datos de las copias rellenas revelan que cada una de las actividades de procesamiento de datos proporciona una cierta cantidad de información relevante, que es extraída durante la comunicación con las diferentes partes interesadas para la especificación de requisitos. La versión actual del M-DVC clasifica esta información relevante en información relacionada con E/P/S para cada una de las actividades de procesamiento. Sin embargo, se necesita prestar más atención para ver qué información relevante puede ser clasificada bajo las etiquetas E/P/S.

Finalmente, hemos encontrado la necesidad de facilitar los términos técnicos incluidos en el M-DVC para los actores educativos [T3]. En la mayor parte de los casos, los actores

MMLA no tienen experiencia previa con soluciones MMLA y muchos actores del campo de la educación se enfrentan a problemas relacionados con conocimientos en tratamiento de datos, lo que a veces obstaculiza su comprensión del M-DVC. Los participantes - especialmente los investigadores procedentes de Ciencias de la Educación- resaltan que los términos del M-DVC propuesta son demasiado técnicos. Para algunos de ellos, el M-DVC presenta un vocabulario completamente nuevo que requiere un esfuerzo significativo para su comprensión.

Debemos de ser conscientes de las limitaciones de nuestro estudio: en primer lugar, la evidencia ha sido recogida solamente en tres casos, lo que no da suficiente evidencia para todos los temas y subtemas. Segundo los tres casos MMLA analizados son de la Universidad de Tallín, lo que significa que comparten el aspecto socio-cultural. Tercero, planeamos llevar a cabo el estudio en una hora en cada uno de los casos, pero en ninguno de ellos pudimos acabar en esa hora prevista. El ejercicio en el que los participantes tenían que comunicar su caso MMLA y rellenar la información en el instrumento requiere más tiempo, sobre todo en los casos en los que los participantes no tienen experiencia previa en procesamiento de datos multimodal. En cuarto lugar, no llevamos a cabo otros estudios similares que nos podrían facilitar la realización de un estudio comparativo. Finalmente, se pidió a los dos participantes de un caso que rellenaran el instrumento con la información requerida en diferentes fases de la propuesta M-DVC. Adicionalmente, el instrumento no establecía quién de los dos participantes de cada

fase debía rellenar el instrumento. Esta falta de guía llevó a cierta confusión entre los participantes, que se reflejó en algunas respuestas negativas en los cuestionarios. Se ha visto la necesidad de un mapeado entre los participantes (basado en un rol a elegir) y la actividad de procesamiento de datos necesaria para ese rol, donde el instrumento declare claramente quién es responsable de rellenar la información en cada parte.

VII. CONCLUSIONES Y TRABAJO FUTURO

Este artículo se basa en trabajos anteriores en los que se ha aplicado una cadena de valor de datos masivos (DVC) [12] a cuatro escenarios de análisis multimodal de aprendizaje (MMLA) cuyo objetivo es el procesamiento de evidencias de aprendizaje multimodales [22]. A partir de ese estudio preliminar, hemos extraído requisitos con el propósito de especializar dicha DVC para el campo de MMLA. En este artículo, hemos presentado una cadena de valor de datos multimodales (*Multimodal Data Value Chain*, o M-DVC) que incorpora dichas lecciones aprendidas, para su uso como herramienta en el soporte a las decisiones de diseño, sistematización y comunicación durante el desarrollo de soluciones MMLA.

Con el objeto de evaluar la utilidad de la M-DVC para dicho propósito, hemos formulado una pregunta de investigación (*¿Hasta qué punto es la M-DVC propuesta útil como herramienta conceptual para apoyar la comunicación de partes interesadas provenientes de múltiples disciplinas, durante el desarrollo de soluciones MMLA?*), que hemos examinado en tres casos reales de MMLA llevados a cabo en la Universidad de Tallín (Estonia). Nuestra evaluación usando métodos mixtos (ver Sección IV para más detalles) muestra que la M-DVC propuesta es útil como herramienta de soporte a decisiones, sistematización y comunicación entre partes interesadas durante el desarrollo de soluciones MMLA. Nuestros resultados subrayan su utilidad como herramienta conceptual, incluso para participantes no acostumbrados al uso de datos complejos. La M-DVC potencia la capacidad de investigadores educativos para comunicarse con desarrolladores y otras partes interesadas del área técnica durante el desarrollo de *software* MMLA. Finalmente, los resultados también indican que esta cadena de valor de datos especializada es útil para participantes que no tienen experiencia anterior con MMLA.

Nuestro trabajo futuro se centra principalmente en dos áreas. La M-DVC (y el instrumento de comunicación basado en ella) debe mejorarse para dar respuesta a algunas de las limitaciones observadas durante la evaluación: 1. abordar explícitamente la idea de “incorporar información contextual”; 2. asignar más claramente las distintas actividades de proceso de datos (y pasos subyacentes a las mismas) a los distintos perfiles de los participantes, para evitar confusión y esfuerzos innecesarios al rellenar el instrumento; y 3. Clarificar y adaptar la terminología técnica utilizada en el instrumento para beneficio de las partes interesadas del lado más educativo.

El otro área de trabajo futuro que estamos abordando se centra en dar soporte adicional a investigadores, educadores y plantilla técnica para desarrollar soluciones MMLA de una manera más eficiente. Planeamos construir una infraestructura MMLA que utilice la M-DVC como marco conceptual, para

modelar las actividades de procesamiento de datos (y los pasos subyacentes detectados). Esta infraestructura permitirá una mayor reutilización durante el desarrollo de soluciones MMLA para distintos escenarios de aprendizaje en la práctica educativa real. Una vez disponible, emplearemos dicha infraestructura, junto con la M-DVC, para dar soporte a escenarios reales de MMLA en diferentes instituciones educativas. De esta manera, recogeremos evidencia adicional de la utilidad de (y posibles mejoras en) la cadena de valor de datos multimodales.

AGRADECIMIENTOS

Esta investigación ha sido parcialmente financiada por: la Unión Europea a través del Fondo Europeo de Desarrollo Regional en el marco del proyecto CEITER (no. 669074), el Fondo Europeo de Desarrollo Regional y el Consejo Regional de Educación de Castilla y León (proyecto no. VA257P18), así como por el FEDER y la Agencia Nacional de Investigación del Ministerio de Ciencia, Innovación y Universidades de España (proyecto TIN2017-85179-C3-2-R).

REFERENCIAS

- [1] R. Ferguson, “Learning analytics: drivers, developments and challenges,” *International Journal of Technology Enhanced Learning*, vol. 4, no. 5/6, pp. 304–317, 2012.
- [2] A. Merceron, P. Blikstein, and G. Siemens, “Learning analytics: from big data to meaningful data,” *Journal of Learning Analytics*, vol. 2, no. 3, pp. 4–8, 2015.
- [3] X. Ochoa, A. C. Lang, and G. Siemens, “Multimodal learning analytics,” *The Handbook of Learning Analytics*, vol. 1, pp. 129–141, 2017.
- [4] P. Blikstein, “Multimodal learning analytics,” in *Proceedings of the third international conference on learning analytics and knowledge*. ACM, 2013, pp. 102–106.
- [5] F. H. Weinberg, “An experimental investigation of the interaction between sensory modality preference and mode of presentation in the instruction of arithmetic concepts to third grade underachievers,” Ph.D. dissertation, St. John’s University (New York), 1984.
- [6] W. QinHong, “On mode, medium, and modality in multimodal discourse [j],” *Foreign Language Education*, vol. 4, 2009.
- [7] C.-C. Chang, H. Lei, and J.-S. Tseng, “Media presentation mode, english listening comprehension and cognitive load in ubiquitous learning environments: Modality effect or redundancy effect?” *Australasian Journal of Educational Technology*, vol. 27, no. 4, 2011.
- [8] X. Ochoa and M. Worsley, “Augmenting learning analytics with multimodal sensory data,” *Journal of Learning Analytics*, vol. 3, no. 2, pp. 213–219, 2016.
- [9] L.-P. Morency, S. Oviatt, S. Scherer, N. Weibel, and M. Worsley, “Icmi 2013 grand challenge workshop on multimodal learning analytics,” in *Proceedings of the 15th ACM on International conference on multimodal interaction*. ACM, 2013, pp. 373–378.
- [10] D. Di Mitri, M. Scheffel, H. Drachler, D. Börner, S. Ternier, and M. Specht, “Learning pulse: a machine learning approach for predicting performance in self-regulated learning using multimodal data,” in *Proceedings of the seventh international learning analytics & knowledge conference*. ACM, 2017, pp. 188–197.
- [11] M. E. Porter and C. Advantage, “Creating and sustaining superior performance,” *Competitive advantage*, vol. 167, 1985.
- [12] H. G. Miller and P. Mork, “From data to decisions: a value chain for big data,” *It Professional*, no. 1, pp. 57–59, 2013.
- [13] A. Ruiz-Calleja, M. J. Rodríguez Triana, L. P. Prieto, K. Poom-Valickis, and T. Ley, “Towards a living lab to support evidence-based educational research and innovation,” Tallinn University, Tech. Rep., 2017.
- [14] W. N. A. Bork, “Multimodal teaching analytics: Spanish language influence on machine learning models for teaching analytics tools,” *williambork.com*, 2018.
- [15] D. Di Mitri, J. Schneider, R. Klemke, M. Specht, and H. Drachler, “Read between the lines: An annotation tool for multimodal data for learning,” in *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. ACM, 2019, pp. 51–60.

- [16] M. Worsley and R. Martínez-Maldonado, "Multimodal learning analytics' past, present, and potential futures." in *CrossMMLA@ LAK*, 2018.
- [17] S. K. Shankar, L. P. Prieto, M. J. Rodríguez-Triana, and A. Ruiz-Calleja, "A review of multimodal learning analytics architectures," in *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 2018, pp. 212–214.
- [18] B. Schneider and P. Blikstein, "Unraveling students' interaction around a tangible interface using multimodal learning analytics," *Journal of Educational Data Mining*, vol. 7, no. 3, pp. 89–116, 2015.
- [19] S. Oviatt, "Ten opportunities and challenges for advancing student-centered multimodal learning analytics," in *Proceedings of the 2018 on International Conference on Multimodal Interaction*. ACM, 2018, pp. 87–94.
- [20] L. P. Prieto, M. J. Rodríguez-Triana, R. Martínez-Maldonado, Y. Dimitriadis, and D. Gašević, "Orchestrating learning analytics (orla): Supporting inter-stakeholder communication about adoption of learning analytics at the classroom level," *Australasian Journal of Educational Technology*, 2019.
- [21] S. Asghar and M. Umar, "Requirement engineering challenges in development of software applications and selection of customer-off-the-shelf (cots) components," *International Journal of Software Engineering*, vol. 1, no. 1, pp. 32–50, 2010.
- [22] S. Kant Shankar, A. Ruiz Calleja, S. Serrano Iglesias, A. Ortega Arranz, P. Topali, A. Martínez Monés *et al.*, "A data value chain to model the processing of multimodal evidence in authentic learning scenarios," in *LASI Spain*. CEUR Workshop Proceedings, 2019.
- [23] W. Greller and H. Drachler, "Translating learning into numbers: A generic framework for learning analytics," *JSTOR*, 2012.
- [24] M. A. Chatti, A. L. Dyckhoff, U. Schroeder, and H. Thüs, "A reference model for learning analytics," *International Journal of Technology Enhanced Learning*, vol. 4, no. 5-6, pp. 318–331, 2013.
- [25] H. Drachler and W. Greller, "Privacy and analytics: it's a delicate issue a checklist for trusted learning analytics," in *Proceedings of the sixth international conference on learning analytics & knowledge*. ACM, 2016, pp. 89–98.
- [26] L. Ali, M. Asadi, D. Gašević, J. Jovanović, and M. Hatala, "Factors influencing beliefs for adoption of a learning analytics tool: An empirical study," *Computers & Education*, vol. 62, pp. 130–148, 2013.
- [27] Y.-S. Tsai and D. Gasevic, "Learning analytics in higher education—challenges and policies: a review of eight learning analytics policies," in *Proceedings of the seventh international learning analytics & knowledge conference*. ACM, 2017, pp. 233–242.
- [28] Y.-S. Tsai, P. M. Moreno-Marcos, K. Tammets, K. Kollom, and D. Gašević, "Sheila policy framework: informing institutional strategies and policy processes of learning analytics," in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. ACM, 2018, pp. 320–329.
- [29] A. Ruiz-Calleja, S. Garcia, K. Tammets, C. Aguerrebere, and T. Ley, "Scaling learning analytics up to the national level: the experience from estonia and uruguay," *CEUR Proceedings*, 2019.
- [30] S. K. Shankar, A. Ruiz-Calleja, L. P. Prieto, M. J. Rodríguez-Triana, and P. Chejara, "An architecture and data model to process multimodal evidence of learning," in *International Conference on Web-Based Learning*. Springer, 2019, pp. 72–83.
- [31] D. M. Kaufman, "Applying educational theory in practice," *Bmj*, vol. 326, no. 7382, pp. 213–216, 2003.
- [32] F. Wang and M. J. Hannafin, "Design-based research and technology-enhanced learning environments," *Educational technology research and development*, vol. 53, no. 4, pp. 5–23, 2005.
- [33] J. W. Creswell, "Mixed-method research: Introduction and application," in *Handbook of educational policy*. Elsevier, 1999, pp. 455–472.

Shashi Kant Shankar obtuvo su Máster en Tecnología (Ciencias de la Computación e Ingeniería) en la Lovely Professional University (India) en 2016 y su Máster en Computación aplicada por la Sikkim Manipal University (India) en 2014. Es estudiante de doctorado de tercer año en la Escuela de Tecnologías Digitales, Universidad de Tallin (Estonia). Sus intereses de investigación se centran en el análisis de aprendizaje multimodal como herramienta de apoyo a la toma de decisiones basada en evidencia en las prácticas de enseñanza-aprendizaje.

María Jesús Rodríguez-Triana se doctoró en Tecnologías de la Información y la Comunicación por la Universidad de Valladolid (España) en 2014, realizando sus estudios postdoctorales en la École Polytechnique Fédérale de Lausanne (Suiza). Actualmente, es investigadora sénior en la Escuela de Tecnologías Digitales de la Universidad de Tallin (Estonia). Sus intereses de investigación se centran en el diseño y el análisis (multimodal) del aprendizaje para apoyar la orquestación en el aula, la investigación del profesorado y la toma de decisiones institucionales.

Adolfo Ruiz Calleja recibió sus títulos de maestría y doctorado en Ingeniería de las Telecomunicaciones por la Universidad de Valladolid (España) en 2007 y 2013, respectivamente. Tras 5 años como investigador postdoctoral e investigador senior en la Universidad de Tallin (Estonia), actualmente trabaja como investigador en el grupo de investigación GSIC-EMIC de la Universidad de Valladolid. Sus intereses de investigación incluyen análisis de aprendizaje, sistemas distribuidos, datos abiertos y su aplicación en el aprendizaje formal e informal.

Luis P. Prieto es investigador principal de la Facultad de Ciencias de la Educación de la Universidad de Tallin (Estonia). Recibió su Ph.D. de la Universidad de Valladolid (España) en 2012 y fue becario Marie Curie en la École Polytechnique Fédérale de Lausanne (Suiza). Sus intereses de investigación incluyen análisis de aprendizaje, especialmente análisis multimodal de los procesos de enseñanza-aprendizaje, el estudio de la orquestación por parte de los docentes y su aplicación para el desarrollo profesional docente.

Pankaj Chejara es estudiante de doctorado en la Escuela de Tecnologías Digitales de la Universidad de Tallin. Recibió su Máster en Tecnología (Ingeniería Informática) del Malviya National Institute of Technology (India) en 2012. Su investigación se centra en el análisis multimodal del aprendizaje colaborativo.

Alejandra Martínez-Monés es profesora de Interacción Persona-Ordenador en la Escuela de Ingeniería Informática de la Universidad de Valladolid, España. Es doctora por esa misma Universidad de Valladolid en 2003. Su trabajo de investigación se centra en el diseño y uso de tecnologías interactivas para apoyar la innovación pedagógica del profesorado, y en el uso de analíticas de aprendizaje para favorecer la orquestación de situaciones de aprendizaje apoyadas por tecnología.