

# Descubrimiento de modelos de comportamiento de perfiles de jugadores en juegos serios

Juan Antonio Caballero-Hernández<sup>1</sup>[0000-0001-9012-8961], Manuel Palomo-Duarte<sup>2</sup>[0000-0001-6002-0319], Juan Manuel Dodero<sup>2</sup>[0000-0002-4105-5679], and Antonio Balderas<sup>2</sup>[0000-0003-0026-7410]

<sup>1</sup> Grupo EVALfor, Universidad de Cádiz, Puerto Real, España  
juanantonio.caballero@uca.es

<sup>2</sup> Dep. de Ingeniería Informática, Universidad de Cádiz, Puerto Real, España  
{manuel.palomo, juanma.dodero, antonio.balderas}@uca.es

**Resumen** Los juegos serios facilitan el aprendizaje de capacidades en entornos educativos por su carácter lúdico. El desarrollo de juegos personalizados facilita alinear la experiencia lúdica con los intereses de aprendizaje concretos, a la vez que permite el registro de la actividad de los jugadores. Sin embargo, la información que se genera puede ser demasiado voluminosa para analizarse con técnicas manuales. En este artículo proponemos aplicar minería de procesos a los datos resultantes de un juego serio. En concreto, se aplica descubrimiento de modelos a un juego serio desarrollado al efecto para aprender modelado conceptual de datos en un sistema informático. En el juego se registra información detallada de cada partida, incluyendo no solo el resultado final sino también cómo el jugador ha llegado a él. Los resultados preliminares en el contexto real de una asignatura del Grado en Ingeniería Informática permiten analizar de manera escalable la información producida por los estudiantes mediante los perfiles que se generan. De este modo, se obtienen evidencias del comportamiento de perfiles específicos de estudiantes para proporcionar retroalimentación o incluso evaluar a los alumnos.

**Palabras claves:** Juegos serios · Aprendizaje basado en juegos · Perfiles de usuario · Minería de procesos · Descubrimiento de modelos.

**Abstract.** Serious games facilitate the learning of skills in educational environments, due to their entertaining nature. Ad-hoc developed serious games ease aligning the ludic experience with specific learning goals, allowing the recording of player activity. However, the data generated can be too large to be analysed through manual techniques. In this paper we propose applying process mining techniques to analyse the data resulting from a serious game experience. Specifically, model discovery is applied to a serious game specifically developed to work skills on conceptual data modelling in a computer system. Detailed data about each game experience is recorded, including not only the final outcome but also how the player reached it. Preliminary results in an actual case study

conducted in a course of a Degree in Computer Science and Engineering, proved to conduct a scalable analysis of the data produced by the students through profile generation. In this way, evidence of the behaviour for specific student profiles is obtained to provide feedback or even assess students.

## 1. Introducción

Los juegos serios se han visto beneficiados de la alta presencia de videojuegos en múltiples entornos, demostrando que pueden ser algo más que simples herramientas de entretenimiento. En general, los juegos serios son juegos y simulaciones que no presentan únicamente propósitos lúdicos, sino también propósitos de aprendizaje [3]. Según indica el informe sobre juegos serios elaborado por la Conferencia de Directores y Decanos de Ingeniería Informática (CODII) de España: “el análisis de los principios de diseño de los videojuegos los identifica como un medio idóneo para la educación: un videojuego es una experiencia de aprendizaje diseñada para mantener el nivel de desafío teniendo al jugador y, además, mantiene la atención del jugador” [7].

Dentro de un juego pueden ocurrir una alta cantidad de interacciones según las características del juego: seleccionar opciones, interactuar con un personaje, usar objetos, etc. Estas interacciones pueden almacenarse como registros secuenciales de eventos formando grandes conjuntos de datos. Sin embargo, el análisis de estos conjuntos de datos puede acarrear problemas de escalabilidad, de forma que llegue a no ser abarcable mediante un análisis manual [11].

La naturaleza secuencial de los juegos serios hace que sean una fuente de datos apropiada para las técnicas de análisis de secuencias. Dentro de estas técnicas, la minería de procesos (*process mining*) proporciona técnicas de análisis de eventos para descubrir, monitorizar y mejorar procesos reales a través de la extracción de conocimiento de registros de eventos producidos por dichos procesos [12]. En este estudio proponemos la aplicación de técnicas de process mining para analizar de forma escalable el comportamiento de perfiles específicos de estudiantes en una experiencia de aprendizaje basada en juegos serios. Para ello, proponemos un método para descubrir modelos de comportamiento que posteriormente es implementado y validado en un caso de estudio para responder a la siguiente pregunta de investigación: **¿Pueden las técnicas de process mining proporcionar evidencias del comportamiento de perfiles específicos de estudiantes en experiencias de aprendizaje basadas en juegos serios?**.

El resto del trabajo está estructurado de la siguiente manera: la sección 2 presenta un contexto general sobre process mining. En la sección 3 describimos el método propuesto. La sección 4 describe el caso de estudio llevado a cabo. La sección 5 proporciona un resumen del juego serio empleado. En la sección 6 realizamos un análisis del modelo de comportamiento obtenido. Finalmente, la sección 7 presenta las conclusiones del estudio.

## 2. Process mining

Por un lado, process mining es una disciplina situada entre la inteligencia computacional y la minería de datos. Por otro lado, está situada entre el modelado y el análisis de procesos [12]. El punto de inicio para aplicar técnicas de process mining es un registro secuencial de eventos de una actividad concreta que pertenezcan a un caso particular. En general, un registro de eventos puede considerarse como una colección de casos, mientras que un caso hace referencia a una traza de eventos. Existen tres tipos de técnicas de minería de procesos: descubrimiento (*discovery*), conformidad (*conformance*) y mejora (*enhancement*) [1].

En primer lugar, las técnicas de descubrimiento toman como entrada un registro de eventos y producen modelos sin usar ninguna otra información adicional. Los modelos pueden ser proporcionados en notaciones diferentes: Redes de Petri, grafos de dependencia, BPMN (del inglés, Business Process Model and Notation), etc. Destacar que BPMN se ha convertido en un estándar para el modelado de procesos en la industria <sup>3</sup>. En segundo lugar, las técnicas de conformidad comparan un modelo de procesos ya existente con un registro de eventos del mismo proceso para validar si los eventos registrados encajan con el comportamiento asociado al modelo y viceversa. Finalmente, las técnicas de mejora se centran en extender o enriquecer modelos de proceso aplicando nueva información incluida en un registro de eventos adicional.

Un algoritmo de descubrimiento de proceso se define como “una función que mapea un registro de eventos como un modelo de proceso, el cual es representativo del comportamiento observado en el registro de eventos” [2]. Existen cuatro criterios que marcan la calidad de un modelo descubierto: idoneidad (de ahora en adelante, *fitness*), precisión, generalización y simplicidad.

*Fitness* mide la capacidad de un modelo para reproducir las trazas de eventos almacenadas en un registro. La precisión mide la capacidad que un modelo posee para aceptar comportamiento no relacionado, por lo que un modelo con una baja precisión es un modelo demasiado genérico. La generalización mide si el modelo puede generalizar el comportamiento analizado. Como contrapartida de la precisión, un modelo con una baja generalización será demasiado específico y adaptado al comportamiento del registro de eventos. Finalmente, la simplicidad indica que el modelo descubierto debería ser el modelo más simple que represente el comportamiento de un registro de eventos.

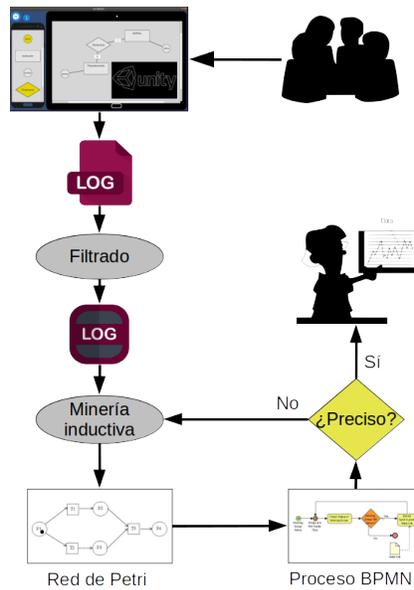
Cabe destacar que process mining ha sido aplicado de forma activa en entornos educativos [8], siendo conocido como “Educational Process Mining” (EPM) [5]. EPM se centra en el uso de los registros de eventos obtenidos en entornos educativos para descubrir, analizar y detectar comportamientos comunes llevados a cabo durante el proceso de aprendizaje. Aunque existen estudios donde se aplican técnicas de analíticas de juego para analizar el comportamiento [4], el uso de EPM en experiencias basadas en juegos no ha sido tratado en profundidad.

---

<sup>3</sup> <https://www.omg.org/spec/BPMN/2.0/>

### 3. Método de descubrimiento de modelos

En este trabajo proponemos el uso de un método basado en técnicas de process mining para descubrir modelos que faciliten el análisis de comportamiento llevado a cabo por perfiles específicos de estudiantes durante una experiencia de aprendizaje basada en juegos serios. La arquitectura del método propuesto se muestra en la Fig. 1.



**Figura 1.** Método para descubrir modelos de comportamiento de perfiles de estudiante en experiencias de aprendizaje basadas en juegos serios.

Nuestro punto de entrada es un registro de eventos con todas las interacciones realizadas por cada jugador durante la experiencia de juego. Durante el juego, las interacciones realizadas por los estudiantes son almacenadas en un registro donde cada instancia corresponde a una traza de eventos. Estos eventos deben incluir algunos campos obligatorios: el CaseID correspondiente (un identificador único por cada instancia de proceso), la actividad (interacción llevada a cabo) y una marca de tiempo. Según el contexto, pueden incluirse otros campos externos: calificaciones, errores cometidos, etc.

El registro de eventos es filtrado para obtener un subconjunto según el perfil de estudiante que queramos analizar, por lo que la herramienta de process mining a utilizar para implementar el método debe proporcionar dicha funcionalidad. Estos perfiles de estudiante pueden basarse en cualquier campo que esté incluido en el registro: edad, género, calificación, duración de la partida, etc.

Una vez obtenido el sub-registro correspondiente a un perfil de estudiante, aplicamos técnicas de descubrimiento de modelos de proceso para identificar patrones de comportamiento. En nuestro método hemos usado el algoritmo B', cuya implementación es conocida como *minería inductiva* (IM, del inglés, *Inductive Miner*). IM descubre un conjunto de modelos de proceso que encajen en el comportamiento observado en un orden de tiempo polinómico al tamaño de los registros, medido como el número de actividades que contiene el registro.

Sin embargo, en caso de que el sub-registro empleado presente una variabilidad de comportamiento demasiado amplia, los modelos obtenidos con IM podrían ser demasiados genéricos. Existe una variante del algoritmo IM llamada *IM - infrequent* (IMi), la cual permite filtrar comportamientos infrecuentes de todos los pasos de IM [9]. A diferencia de otros enfoques, IMi puede descubrir los llamados *modelos 80 %* usando el principio de Pareto, mientras garantiza la obtención de un modelo de proceso sólido en un corto espacio de tiempo.

El proceso proporciona como salida un modelo en notación de red de Petri, la cual puede transformarse a un modelo BPMN. En caso de que el modelo obtenido no sea lo suficientemente preciso, se pueden llevar a cabo nuevas iteraciones aplicando IMi con diferentes parámetros, filtrando caminos infrecuentes.

#### 4. Caso de estudio

En esta sección, implementamos y validamos nuestro método a través de un caso de estudio. Las técnicas de process mining del método fueron implementadas a través del framework open source ProM [13]. ProM permite descubrir modelos que reflejan la dinámica incluida en un registro de eventos. El perfil de estudiante analizado en este caso de estudio corresponde a aquellos estudiantes que obtuvieron en la calificación final de la asignatura un sobresaliente (9 o más sobre 10).

El caso de estudio fue llevado a cabo en la asignatura “Bases de Datos”, obligatoria para los estudiantes del Grado en Ingeniería Informática de la Universidad de Cádiz durante el segundo semestre del segundo curso. 110 estudiantes matriculados en la asignatura “Bases de Datos” participaron en la experiencia de aprendizaje basada en juegos serios. En el experimento utilizamos un videojuego que propone el diseño de una especificación conceptual de datos a través de un diagrama de Entidad Relación (E/R). El juego fue desarrollado utilizando el motor Unity, es multiplataforma (Windows, GNU/Linux y MacOs) y posee enlace de descarga permanente [6].

El jugador realiza un diagrama E/R en base a una serie de requisitos textuales sobre una universidad: profesores, alumnos, asignaturas, etc. El problema planteado dentro del juego está basado en el ejemplo práctico presentado en el anexo de [10], una referencia ampliamente utilizada para la enseñanza de fundamentos de bases de datos. La solución propuesta contiene 6 entidades, 19 atributos, 7 relaciones y 14 cajas donde indicar cardinalidades.

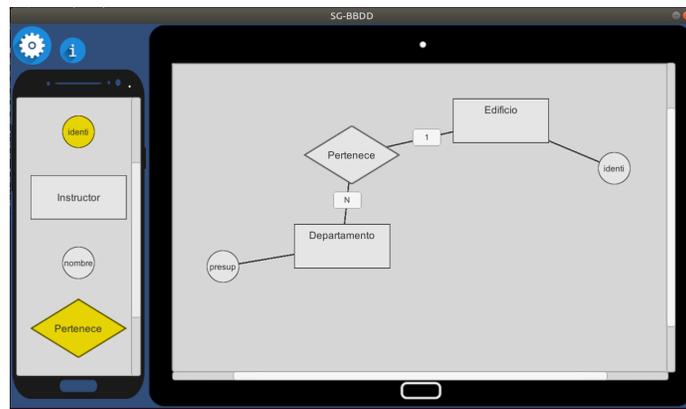
## 5. Resumen del juego serio

Inicialmente, se le solicita al jugador un identificador único, el cual se incluye en el registro de eventos. Tras introducirlo, el jugador puede usar un menú para navegar por las diferentes pantallas del juego: el mapa, el bloc de notas y el editor de diagrama E/R. Además, el menú proporciona una opción para confirmar el diagrama E/R y salir del juego.

En la pantalla del mapa el jugador debe obtener los requisitos textuales para el ejercicio propuesto visitando diferentes edificios. Al seleccionar un edificio, un personaje (un delegado de clase, un profesor o una directora de departamento) proporcionará nuevos requisitos textuales.

La pantalla del bloc de notas permite consultar todos los requisitos recolectados por el jugador. Cada requisito está formado por una o dos frases donde la información más relevante para el diseño del E/R (palabras clave) están resaltadas en un color diferente.

Por último, el jugador debe diseñar el diagrama E/R usando las herramientas proporcionadas en la pantalla del editor, el cual consta de una barra de inventario y una zona de trabajo. La Fig. 2 muestra un ejemplo de diseño de un E/R. En primer lugar, el inventario contiene las palabras clave de los requisitos textuales obtenidos, de forma que el jugador debe elegir el tipo de elemento al que corresponde una palabra (entidad, atributo o relación) antes de incorporarlo al modelo. En segundo lugar, la zona de trabajo permite al usuario organizar los elementos del modelo, relacionarlos entre sí, seleccionar sus cardinalidades (0, 1 ó N) o eliminarlos. Cuando un jugador confirma su diseño del diagrama E/R, se crea un registro de eventos con todas las interacciones realizadas durante el juego.



**Figura 2.** Ejemplo de un diagrama E/R siendo diseñado en la pantalla del editor. La barra de inventario es mostrada a la izquierda. La zona de trabajo corresponde al resto de la ventana.

Una vez recolectados los registros de eventos de todos los estudiantes, estos fueron procesados para unificarlos de forma automática en un único conjunto de datos. En este procesado se añadieron las calificaciones finales que los estudiantes obtuvieron en la asignatura, lo que permitió analizar el comportamiento de perfiles específicos de estudiantes en base a sus resultados.

El registro de eventos procesado contuvo 9.402 eventos, donde cada evento corresponde a una interacción, y 110 trazas, donde cada traza corresponde a la experiencia de juego de un estudiante. En total existen 18 tipos de interacciones relacionadas con la construcción del diagrama E/R. Estas interacciones abarcan las diferentes operaciones que el jugador puede aplicar sobre los diferentes elementos del diagrama E/R (entidad, atributo y relación): añadir, borrar, unir y separar. Además, existe una interacción por cada tipo de cardinalidad asignada a la unión de una entidad con una relación (0,1,N).

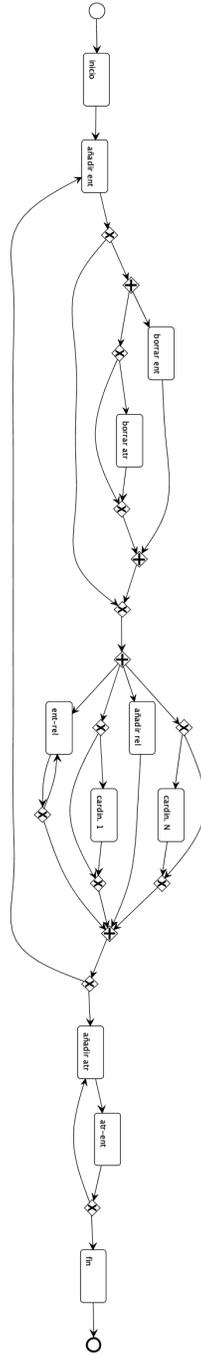
## 6. Análisis del modelo de comportamiento

El registro de eventos obtenido en la experiencia de aprendizaje fue importado en ProM. Tras filtrar el registro según las calificaciones de los estudiantes, obtuvimos un sub-registro con los eventos de aquellos estudiantes con una calificación de sobresaliente (10 estudiantes). El sub-registro fue utilizado como entrada para aplicar técnicas de descubrimiento de modelos basadas en el algoritmo IMi. Buscando un balance entre precisión y generalización, filtramos el 20 % de los caminos menos frecuentes. El algoritmo generó el modelo de proceso mostrado en la Fig. 3.

El modelo es presentado siguiendo una notación BPMN. En esta sintaxis, existen dos círculos vacíos que representan el inicio y el final del modelo, respectivamente. Todos los nodos se muestran como cajas etiquetadas con una acción de evento, lo que en nuestro contexto corresponde con un tipo de interacción durante la experiencia de juego. Los nodos están enlazados por flechas direccionadas que representan la secuencialidad entre los eventos. Además, existen dos tipos de cajas en forma de rombo para representar el control de flujo. Por un lado, los rombos con un símbolo “+” corresponden a uniones de los caminos de entrada y a una división en paralelo de los caminos de salida. Por otro lado, los rombos con un símbolo “X” corresponden a puntos de decisión donde el camino se bifurca en caso de tener múltiples salidas. Finalmente, los nombres de los eventos fueron acortados para mejorar la visualización del modelo: entidad (ent), atributo (atr) y relación (rel).

En general, el modelo presenta una estructura lineal pero incluyendo ciertas iteraciones y bifurcaciones, alejándose de estructuras demasiado adaptadas al registro (modelos lineales con escasa variedad secuencial) y de estructuras sobregeneralizadas (modelos “en flor” donde cualquier secuencia es posible). Por tanto, consideramos que es un modelo representativo y apropiadamente balanceado entre precisión y generalización.

El modelo contiene cuatro etapas diferenciadas y una iteración principal que va desde el final de la tercera etapa hasta el principio del modelo. La iteración re-



**Figura 3.** Modelo de comportamiento de los estudiantes con calificación de sobresaliente.

presenta un proceso repetitivo donde los estudiantes recorrieron secuencialmente las tres primeras etapas antes de pasar a la cuarta y finalizar el diseño.

En primer lugar, los estudiantes comenzaron con la interacción de añadir entidades, lo cual tiene sentido debido a que las entidades pueden considerarse como el punto de inicio del diseño de un diagrama E/R.

La segunda etapa incluye dos interacciones de borrado de elementos en el diagrama: entidades y atributos. Debido a las bifurcaciones incluidas, solo parte de los alumnos eliminaron entidades y/o atributos, por lo que algunos diseñaron el diagrama E/R a la perfección sin tener que realizar correcciones. Además, evidencia que solo se eliminaron elementos en ciertas iteraciones del proceso.

La tercera etapa incluye cuatro tipos de interacciones: añadir relaciones, unir las con las entidades e indicar cardinalidades con valor 1 y/o N. Esta etapa representa la creación de relaciones entre entidades. Las interacciones de añadir relaciones y de unir entidades con relaciones son obligatorias cada vez que un estudiante pase por esta etapa, mientras que las cardinalidades pueden ser evitadas debido a las bifurcaciones incluidas. Esto evidencia que los estudiantes dejaron pendientes ciertas cardinalidades para ser indicadas posteriormente, probablemente una vez el diagrama E/R estuviera más completo. Además, el bucle unido a la interacción de unir entidades y relaciones indica que al igual que con las cardinalidades, los estudiantes pudieron retrasar la inclusión de relaciones entre entidades para iteraciones posteriores.

Tras esta tercera etapa, el modelo presenta una bifurcación que permite volver al inicio del modelo o seguir avanzando. Finalmente, la cuarta etapa incluye dos interacciones: añadir atributos y unirlos con las entidades. Esta etapa incluye un pequeño bucle para repetir este proceso tantas veces como sea necesario.

El modelo analizado evidencia ciertos patrones de comportamiento propios de este grupo de estudiantes. Realizaron un proceso estructurado e iterativo, el cual se basó en añadir entidades, relacionar aquellas que el estudiante considerara oportuno, y volver a añadir entidades para repetir el proceso. Una vez indicadas todas las relaciones entre entidades, se añadieron y relacionaron todos los atributos con las entidades correspondientes.

Cabe destacar que el modelo carece de ciertas interacciones incluidas en el registro, por lo que los estudiantes no borraron relaciones ni separaron elementos una vez unidos. Esto evidencia una clara y adecuada interpretación de los requisitos textuales. Además, solo se indicaron cardinalidades 1 y N, evitando cardinalidades con valor 0 y que no existen en la solución proporcionada.

## 7. Conclusiones

En este trabajo se propone un método para analizar de manera escalable el comportamiento de perfiles específicos de estudiantes a través de la aplicación de técnicas de process mining. Para implementar y validar nuestro método, se llevó a cabo un caso de estudio en una experiencia de aprendizaje basada en juegos serios en estudios universitarios de grado.

Nuestro método produjo un modelo de comportamiento para aquellos estudiantes que obtuvieron una calificación final de sobresaliente en la asignatura. El modelo proporcionó varias evidencias de los patrones de comportamiento seguidos por este perfil de estudiantes durante la experiencia de juego que permite dar feedback y descubrir errores comunes.

Los resultados permiten afirmar que la propuesta facilitó el análisis de comportamiento de los estudiantes en experiencias de aprendizaje basadas en juegos serios. El método propuesto proporcionó un análisis escalable y detallado, proporcionando evidencias del comportamiento desempeñado. Por tanto responde afirmativamente a la pregunta de investigación: **¿Pueden las técnicas de process mining proporcionar evidencias del comportamiento de perfiles específicos de estudiantes en experiencias de aprendizaje basadas en juegos serios?**. Como trabajo futuro proponemos el uso de estas evidencias mediante la inclusión del descubrimiento de modelos como parte de un método más amplio para evaluar competencias educativas. Además, proponemos aplicar el descubrimiento de modelos para comparar dinámicas de comportamiento de diferentes grupos de estudiantes.

## Agradecimientos

Este trabajo ha sido financiado parcialmente por el proyecto VISAIGLE (código TIN2017-85797-R).

## Referencias

1. van der Aalst, W.M.P.: Process Mining Discovery, Conformance and Enhancement of Business Processes. Berlin Heidelberg: Springer (2011). <https://doi.org/10.1007/978-3-642-19345-3>
2. van der Aalst, W.M.P.: Process Mining Data Science in Action. Berlin Heidelberg: Springer, 2nd ed. edn. (2016). [https://doi.org/10.1007/978-3-662-49851-4\\_1](https://doi.org/10.1007/978-3-662-49851-4_1)
3. Abt, C.: Serious games: The art and science of games that simulate life. New York: Viking Press (1970)
4. Bicalho, L.F., Baffa, A., Feijó, B.: A game analytics model to identify player profiles in singleplayer games. In: 2019 18th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames). pp. 11–20 (2019)
5. Bogarín, A., Cerezo, R., Romero, C.: A survey on educational process mining. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery **8**(1), e1230 (jan 2018). <https://doi.org/10.1002/widm.1230>, <http://doi.wiley.com/10.1002/widm.1230>
6. Caballero-Hernández, J.A.: Supporting skill assessment in learning experiences based on serious games through process mining techniques (2020). <https://doi.org/10.6084/m9.figshare.c.4916412.v1>, <https://doi.org/10.6084/m9.figshare.8984090.v1>
7. Fernández-Manjón, B., Moreno-Ger, P., Freire, M., Martínez-Ortiz, I.: Juegos serios. Tech. rep., Conferencia de Directores y Decanos de Ingeniería Informática (CODII) - Universidad Complutense de Madrid (2016), <https://goo.gl/1k1EPn>

8. Garcia, C.d.S., Meinheim, A., Faria Junior, E.R., Dallagassa, M.R., Sato, D.M.V., Carvalho, D.R., Santos, E.A.P., Scalabrin, E.E.: Process mining techniques and applications – A systematic mapping study. *Expert Systems with Applications* **133**, 260–295 (nov 2019). <https://doi.org/10.1016/j.eswa.2019.05.003>
9. Leemans, S.J.J., Fahland, D., van der Aalst, W.M.P.: Discovering Block-Structured Process Models from Event Logs Containing Infrequent Behaviour. In: Lohmann, N., Song, M., Wohed, P. (eds.) *Business Process Management Workshops. BPM 2013. Lecture Notes in Business Information Processing*, vol 171, pp. 66–78. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-06257-0\\_6](https://doi.org/10.1007/978-3-319-06257-0_6)
10. Silberschatz, A., Korth, H.F., Sudarshan, S.: *Database system concepts*. New York: McGraw-Hill, 6th ed. edn. (2011)
11. Slater, S., Joksimović, S., Kovanovic, V., Baker, R.S., Gasevic, D.: Tools for educational data mining: A review. *Journal of Educational and Behavioral Statistics* **42**(1), 85–106 (2017). <https://doi.org/10.3102/1076998616666808>
12. Van Der Aalst, W., Adriansyah, A., De Medeiros, A.K.A., Arcieri, F., Baier, T., Blickle, T., Bose, J.C., Van Den Brand, P., Brandtjen, R., Buijs, J., et al.: Process mining manifesto. In: *International Conference on Business Process Management*. pp. 169–194. Springer (2012)
13. Verbeek, H.M.W., Buijs, J.C.A.M., Van Dongen, B.F., van der Aalst, W.M.: ProM: The Process Mining Toolkit. In: *International Conference on Business Process Management Demonstration Track*. pp. 34–39. Hoboken, New Jersey (2010)