

EDUCACIÓN, CREATIVIDAD E INTELIGENCIA ARTIFICIAL: NUEVOS HORIZONTES PARA EL APRENDIZAJE. ACTAS DEL VIII CONGRESO INTERNACIONAL SOBRE APRENDIZAJE, INNOVACIÓN Y COOPERACIÓN, CINAIC 2025

María Luisa Sein-Echaluce Lacleta, Ángel Fidalgo Blanco y Francisco José García Peñalvo (coords.)

1ª Edición. Zaragoza, 2025

Edita: Servicio de Publicaciones. Universidad de Zaragoza.



Servicio de
Publicaciones
Universidad Zaragoza

EBOOK ISBN 978-84-10169-60-9

DOI 10.26754/uz.978-84-10169-60-9



Esta obra se encuentra bajo una licencia Creative Commons Reconocimiento – NoComercial (ccBY-NC). Ver descripción de esta licencia en <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Referencia a esta obra:

Sein-Echaluce Lacleta, M.L., Fidalgo Blanco, A. & García-Peñalvo, F.J. (coords.) (2025). *Educación, Creatividad e Inteligencia Artificial: nuevos horizontes para el Aprendizaje. Actas del VIII Congreso Internacional sobre Aprendizaje, Innovación y Cooperación. CINAIC 2025 (11-13 de Junio de 2025, Madrid, España)*. Zaragoza. Servicio de Publicaciones Universidad de Zaragoza. DOI 10.26754/uz.978-84-10169-60-9

Una Herramienta para Evaluación de Equipos de Trabajo en un marco de Metodologías Ágiles

A Tool for Evaluating Work Teams in an Agile Methodology Framework

Patricio Orlando Letelier Torres, Antonio Molina Marco, José Fabián Reyes Román
letelier@dsic.upv.es, amolina@upv.es, jreyes@pros.upv.es

Valencian Research Institute for Artificial Intelligence (VRAIN)
Universitat Politècnica de València
Valencia, España

Resumen- El Aprendizaje Basado en Proyectos (ABP) es un método activo que ha ganado creciente atención. No obstante, la evaluación justa y formativa en contextos de trabajo en equipo sigue siendo un desafío al introducir ABP. En este trabajo se presenta el módulo de evaluación de Worki, desarrollado para apoyar la gestión ágil de proyectos y validado en asignaturas de desarrollo de software donde los estudiantes trabajan en equipo. Worki combina la evaluación desde diferentes perspectivas: autoevaluación, coevaluación entre integrantes y evaluación externa por parte del profesorado. Su módulo de evaluación facilita tanto la introducción de evaluaciones como el cálculo de las notas individuales. Gracias a este enfoque, se ha logrado implementar una evaluación continua, formativa y eficaz, valorando tanto el rendimiento del equipo como el de cada miembro. Los resultados obtenidos muestran un impacto positivo y con una alta aceptación de parte de los estudiantes.

Palabras clave: Proyecto de Desarrollo de Software, Metodologías Ágiles, Método de Evaluación, Worki.

Abstract- Project-based Learning (PBL) is an active learning method that has been gaining increasing attention. However, implementing fair and formative assessment in team-based contexts remains a challenge. This paper presents the evaluation module of Worki, a tool developed to support agile project management and validated in software development courses where students work in teams. Worki integrates the evaluation from multiple perspectives: i) self-evaluation, ii) peer evaluation among team members, and iii) external evaluation by the teacher. Its evaluation module streamlines the input of assessments and the calculation of individual grades. This approach has enabled the effective implementation of continuous and formative evaluation, addressing both team performance and individual contributions. The results demonstrate a positive impact and high levels of student acceptance.

Keywords: Software Development Project, Agile Methodologies, Evaluation Method.

1. INTRODUCCIÓN

La habilidad para trabajar eficazmente en equipo es actualmente una de las más solicitadas por las empresas en sus puestos de trabajo. Esta necesidad se ha trasladado al ámbito académico como un resultado de aprendizaje clave para los estudiantes. Sin embargo, la evaluación del desempeño individual es uno de los desafíos cuando el trabajo del estudiante se desarrolla trabajando en equipo. Suele ser difícil

desde un punto de vista externo al equipo de trabajo diferenciar el desempeño individual de cada integrante. El problema es mayor cuando se trabaja con muchos estudiantes agrupados en un número importante de equipos de trabajo.

Nuestro enfoque lo hemos desarrollado en el marco de proyectos de desarrollo de software y aplicando metodologías ágiles (Letelier Torres & Cornide Reyes, 2020). Sin embargo, este enfoque es fácilmente extrapolable a otros contextos de asignaturas en los cuales se trabaje en equipo. El objetivo de este trabajo es presentar la herramienta de gestión de evaluaciones, además de ofrecer información asociada a la validación de nuestro enfoque con la aplicación en los dos últimos cursos, en cada uno de ellos involucrando a más de 100 estudiantes y alrededor de 20 equipos de trabajo.

La Figura 1 muestra cómo diferentes perspectivas de evaluación de se pueden adoptar en un marco metodológico ágil, particularmente en uno que utiliza las definiciones de responsabilidades establecidas por el método Scrum (Schwaber & Sutherland, 2020). El Scrum Team está formado por el Product Owner (PO), el Scrum Master (SM) y los Desarrolladores. El PO realiza una evaluación desde una perspectiva de satisfacción con el resultado conseguido. El SM se centra en una evaluación desde un punto de vista metodológico y de mejora del desempeño de los Desarrolladores. Además, cada desarrollador realiza una autoevaluación de su desempeño y una coevaluación del desempeño de sus compañeros de equipo.

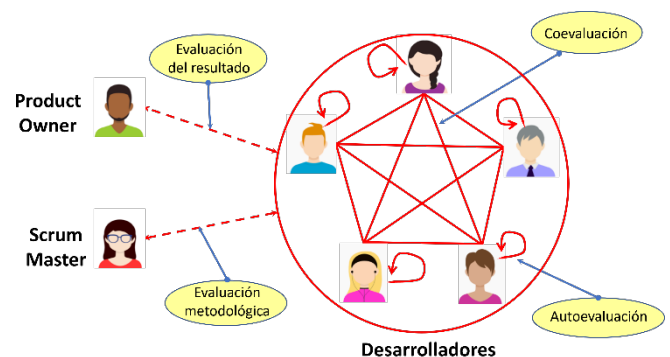


Fig. 1. Tipos de evaluación en un contexto Scrum.

En este trabajo, además de mejorar la precisión y equidad en la evaluación del desempeño individual dentro de los equipos de desarrollo software, también se fomenta y potencia una retroalimentación formativa, promoviendo una cultura de mejora continua alineada con los principios del enfoque ágil. La incorporación de una herramienta de soporte facilita la gestión de la evaluación por parte del profesorado, reduciendo la gestión y permitiendo un seguimiento más detallado del progreso de los estudiantes. En este sentido, nuestro enfoque contribuye a una evaluación más integral y transparente, adaptable a distintos entornos académicos y profesionales donde el trabajo en equipo es un pilar fundamental.

A. Trabajos Relacionados

Hoy en día, en el entorno académico e industrial se han adoptado de forma significativa la aplicación de metodologías ágiles para el desarrollo de productos software. Las metodologías ágiles (Álvarez & Andrés, 2021) han demostrado ser eficaces para mejorar la productividad de los equipos de desarrollo y la calidad del software entregado (Akhtar et al., 2022). El desafío actual radica en cómo llevar a cabo la evaluación del desempeño de los equipos de trabajo, ya que en muchas ocasiones los parámetros para evaluar no reflejan en su totalidad el flujo de trabajo colaborativo y la autogestión propia en los equipos ágiles.

En esta línea de trabajo se han realizado diversos estudios y/o trabajos de investigación centrados en abordar la evaluación de equipos de trabajo en un contexto de metodologías ágiles. En (López Trujillo et al., 2018) se centran en la identificación de roles y a través del uso de escalas de valoración se busca medir el grado de cumplimiento de estos roles, para ello aportan una herramienta de medición centrada en comportamientos/roles, más que en resultados técnicos. El trabajo de (Hamer et al., 2021) se centra en el uso de métricas de Git (commits, merges, líneas de código) para medir la contribución individual y grupal de estudiantes en proyectos académicos de desarrollo, su principal desventaja es que no considera aspectos cualitativos como la calidad del código o la colaboración entre equipos. El trabajo reportado en (Matthies et al., 2016) presenta una propuesta interesante para contextos académicos integrando métricas ágiles (velocidad del equipo, tareas finalizadas y participación en reuniones), las cuales están alineadas con Scrum y facilitan una evaluación continua del desempeño del equipo. Finalmente, en la propuesta de (Fernandes et al., 2011) el principal aporte se basa en mostrar la aplicabilidad práctica y los desafíos reales en contextos productivos.

Este artículo se organiza de la siguiente forma: la **Sección 2** describe el método propuesto y la herramienta de evaluación; la **Sección 3** presenta su validación; y la **Sección 4** expone conclusiones y futuras líneas de mejora para optimizar la evaluación del desempeño en equipos ágiles.

2. DESCRIPCIÓN DEL MÉTODO Y APOYO CON LA HERRAMIENTA DE EVALUACIÓN

El proyecto se divide en Sprints, o ciclos de trabajo cortos, al final de los cuáles se realiza una evaluación desde cada una de las perspectivas mencionadas anteriormente. Utilizamos la herramienta Worki para gestionar los proyectos. Worki ofrece diversas funcionalidades para que cada equipo gestione su trabajo, entre ellas: tableros *kanban* para visualizar el estado del trabajo, registro de estimaciones y tiempos, gestión de pruebas de aceptación, un *dashboard* con métricas y gráficas para el

seguimiento de los Sprints y del Proyecto, etc. Adicionalmente, en Worki hemos ido desarrollando un módulo integrado para la evaluación del desempeño del equipo y de sus integrantes.

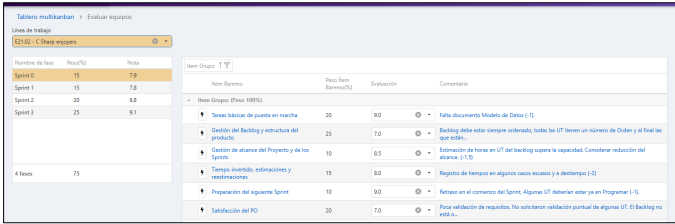


Fig. 2. Evaluación del equipo.

En la Figura 2 se muestra la evaluación de un Sprint realizada por el PO y el SM. En este caso los ítems del baremo y sus pesos también son configurables. La idea es definir algunos ítems que son evaluados por el SM y otros por el PO. La escala de evaluación también es configurable, en este caso los valores van de 0 a 10 cada 0.5 puntos. El PO o SM pueden introducir un comentario respecto de la evaluación de cada ítem del baremo, lo cual orienta al equipo respecto de lo que está fallando. El promedio ponderado de los ítems del baremo da como resultado la evaluación de un Sprint para el equipo.

En la Figura 3 se muestra la interfaz que tiene cada estudiante para autoevaluarse y coevaluar a sus compañeros. El baremo de autoevaluación y coevaluación puede ser configurado. Como se observa, este ejemplo se ha configurado con criterios de evaluación relativos a competencias transversales, la mayor parte asociadas a trabajo en equipo. Estos criterios son una adaptación de los criterios utilizados por el Instituto de Ciencias de la Educación de la UPV en el diseño de las rúbricas de evaluación de las competencias transversales. En las columnas aparecen los integrantes del equipo, incluyendo una columna para que el propio estudiante realice su autoevaluación. Para cada integrante y cada dimensión del baremo de evaluación se debe introducir un valor de desempeño. La escala de valores de desempeño también es configurable, en este ejemplo se han utilizado los valores: 2-Nunca, 5-Casi Nunca, 7.5-A veces, 9-Casi siempre y 10-Siempre. La autoevaluación de momento solo la utilizamos para que el estudiante la realice y tome conciencia de cómo evaluará a sus compañeros. Los pesos de cada una de las dimensiones que se evalúan también son configurables.

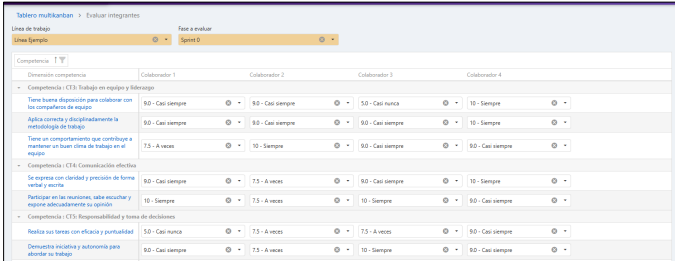


Fig. 3. Autoevaluación y coevaluación de integrantes de equipo.

La autoevaluación y coevaluación de cada Sprint está abierta durante toda la duración del Sprint pudiéndose modificar. Una vez terminado un Sprint se cierra su evaluación y se abre la del Sprint siguiente. Los estudiantes pueden acceder a una interfaz similar a la de la Figura 3 para ver la evaluación del equipo, obviamente sin poder modificar la evaluación establecida.

Sistema de Evaluación - Evaluaciones Individuales							
Unidad de Trabajo		Colaborador					
11102 - Orden Tema							
Nombre de la fila	Peso(%)	Autoevaluación	Coevaluación promedio	Coevaluación promedio resto equipo	Diferencia	Nota Equipo	Nota ajustada
Sprint 0	15	9.1	8.4	9.2	-0.2	8.5	8.6
Sprint 1	15	8.2	8.8	9.4	-0.6	8.7	8.1
Sprint 2	20	9.2	8.2	9.4	-1.2	8.6	7.4
Sprint 3	20	8.8	7.3	9.2	-1.9	9.1	7.2
4 Sprints	75						
Competencias							
Competencia: CTS. Trabajo en equipo y liderazgo (Peso 40%, Promedio 7.5)		Peso Dimensión (%)		Evaluación promedio			
Aplica correctamente la metodología de trabajo		30		5.7			
Tiene buena disposición para colaborar con los compañeros de equipo		40		6.3			
Tiene un comportamiento que contribuye a mantener un buen clima de trabajo en el equipo		30		6.3			
Competencia: CTS. Comunicación efectiva (Peso 30%, Promedio 7.8)							
Participa en las reuniones, sabe escuchar y expone adecuadamente su opinión		50		6.0			
Se expresa con claridad y precisión de forma verbal y escrita		50		7.5			
Competencia: CTS. Responsabilidad y toma de decisiones (Peso 30%, Promedio 6.6)							
Demuestra iniciativa y autonomía para decidir su trabajo		50		7.2			
Realiza sus tareas con eficacia y puntualidad		50		5.7			

Fig. 4. Evaluaciones de un integrante.

Para determinar la evaluación de cada integrante en cada Sprint se aplica un ajuste de la nota del equipo y basado en la coevaluación recibida por sus compañeros/as. En cada Sprint se calcula la coevaluación promedio otorgada a un integrante (sin contar con su autoevaluación) y se calcula la coevaluación promedio del resto de integrantes del equipo. La diferencia entre la coevaluación del integrante respecto de la media de coevaluaciones del resto del equipo se aplica sumando o restando a la nota del equipo para establecer así la nota individual ajustada del integrante en el Sprint. Si para sus compañeros el integrante ha tenido un desempeño destacado dicha diferencia será positiva y se le otorgará una recompensa sobre la nota puesta al equipo. En el caso opuesto si su desempeño ha sido menor respecto del promedio de coevaluaciones del resto de sus compañeros se verá penalizado pues se le restará dicha diferencia.

En la Figura 4 se muestran las evaluaciones obtenidas por un integrante en cada Sprint (en este caso está seleccionado el Sprint 0). En la tabla superior de dicha figura se muestra en cada fila un Sprint con los valores de autoevaluación, coevaluación promedio, coevaluación del resto de integrantes, diferencia entre coevaluaciones, nota del equipo (otorgada por SM y PO) y finalmente la nota ajustada que se le otorga al integrante. Vemos en este ejemplo que el estudiante no ha tenido un buen desempeño respecto de su equipo y ha tenido una importante penalización. Por ejemplo, en el Sprint 3 la nota del equipo ha sido 9.1 pero dado que la diferencia entre coevaluaciones personal y la del resto de su equipo es -1.9, su nota ajustada para ese Sprint es 7.2.

Finalmente, el profesorado cuenta con una interfaz para poder consultar todas las evaluaciones de los estudiantes al terminar cada Sprint. Esta información ordenada por las diferencias de coevaluación permite detectar casos de desempeño deficiente, o pistas de algún conflicto en el equipo.

En el marco de la Reunión de Retrospectiva cada equipo debe reflexionar respecto de su rendimiento teniendo en cuenta los comentarios que el PO y SM han puesto en su evaluación del equipo. En dicha reunión el profesor actuando como SM comenta con el equipo la evaluación de desempeño del equipo y de cada uno de sus integrantes. En esta reunión se pone especial atención a las diferencias de coevaluación negativas, aquellas más significativas, confirmando si están bien valoradas las coevaluaciones asociadas, y de ser así, se intenta reconducir la situación para conseguir mejorar el desempeño del integrante en el siguiente Sprint. Así la Reunión de Retrospectiva es el mecanismo que garantiza una evaluación formativa, pues los estudiantes al finalizar cada Sprint tienen una evaluación de su desempeño como equipo e

individualmente, teniendo la posibilidad de mejorar en los siguientes Sprints. Cabe comentar que con estos mecanismos de evaluación a veces se detectan interesantes situaciones de conflicto. La más usual es la “todos contra uno”, es decir, un integrante del equipo es coevaluado muy mal por el resto de los integrantes. Esto a veces tiene como respuesta de dicho integrante también una coevaluación mala para el resto de sus compañeros, lo cual puede reducir algo las diferencias de todos. Otra situación anómala que se comenta con el equipo es que a veces las coevaluaciones son muy buenas pero el desempeño del equipo evaluado por el profesorado no es tan bueno. Sin embargo, si todos los integrantes se coevalúan de forma similar las diferencias son cercanas a cero, con lo cual la nota individual es prácticamente la misma nota que ha otorgado el docente. Finalmente, en pocas ocasiones hemos detectado un caso bastante interesante, el conflicto puntual entre dos estudiantes. En esta situación los dos estudiantes en conflicto se coevalúan significativamente mal, pero sus compañeros no les otorgan una mala coevaluación. Así, solo los que están en conflicto ven disminuida su nota, y el resto se ve beneficiado de dicho conflicto. El profesorado en la Reunión de Retrospectiva puede confirmar estas situaciones anómalas e intentar establecer acuerdos de mejora con los involucrados.

3. VALIDACIÓN Y RESULTADOS

El enfoque propuesto se ha validado con estudiantes de varias asignaturas de grado. Se utilizó un cuestionario de satisfacción que incluía cinco cuestiones relacionadas directamente con el sistema de evaluación, en una escala Likert de cinco opciones: “Totalmente en desacuerdo” (TD), “Más bien en desacuerdo” (D), “Indiferente” (I), “Más bien de acuerdo” (A) y “Totalmente de acuerdo” (TA). A continuación, presentamos los resultados obtenidos en la asignatura Proyecto de Ingeniería del Software (PIN) que se cursa en cuarto curso del Grado en Ingeniería Informática de la Escola Tècnica Superior d'Enginyeria Informàtica de la Universitat Politècnica de València. En el curso 24-25 la asignatura tuvo 102 estudiantes matriculados, de los que 95 respondieron el cuestionario (93%).

Las preguntas que incluía el cuestionario fueron las siguientes:

- **Pregunta 1. Evaluación Justa:** el método de evaluación me parece justo (evaluación a nivel de equipo realizada por el Scrum Master, incluyendo un ajuste asociado a la evaluación de competencias que me hacen mis compañeros/as de equipo).
- **Pregunta 2. Reflexión Individual:** a lo largo del proyecto he revisado las evaluaciones de competencias que me han hecho mis compañeros/as de equipo y me ha servido para mejorar y/o reforzar mi desempeño en el proyecto.
- **Pregunta 3. Reflexión Grupal:** en mi equipo hemos comentado las evaluaciones de competencias que hemos obtenido.
- **Pregunta 4. NO Conflictos:** en mi equipo la coevaluación de competencias NO creo conflictos o mal clima en el equipo.
- **Pregunta 5. Desempeño continuo:** el hecho de que la coevaluación de competencias por Sprint pudiera afectar a mis notas durante todo el proyecto contribuyó a que no me relajara o desconectara del proyecto.

El resumen de las respuestas se muestra en la Figura 5.

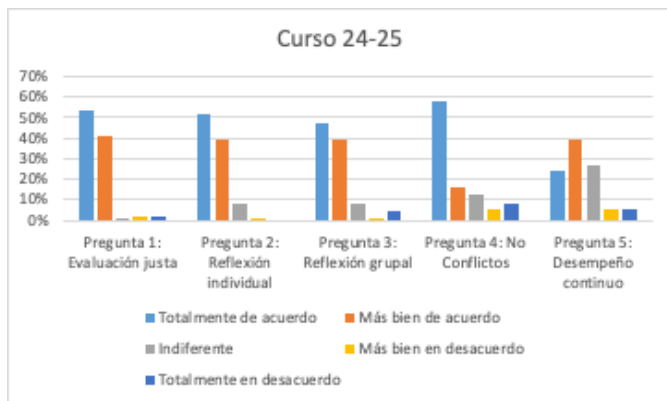


Fig. 5. Opinión de los estudiantes del curso 2024-25

Como se puede observar, la percepción general de los estudiantes sobre el sistema de evaluación fue muy positiva y mayoritariamente aprovecharon la información que les proporcionaba para mejorar su desempeño individual y como equipo. Destacamos el hecho que una gran mayoría de los estudiantes consideró que el método de evaluación diseñado es justo (Pregunta 1, TA: 53,7% y A: 41,1%). Es decir, les parece justo que el resultado de la coevaluación realizada por sus compañeros repercuta en su calificación individual. Los resultados muestran que la evaluación ha fomentado entre los estudiantes la reflexión sobre su desempeño tanto de manera individual (Pregunta 2, TA: 51,6% y A: 38,9%) como grupal (Pregunta 3, TA: 47,4% y A: 38,9%). Esta reflexión es un punto de partida para la mejora del desempeño, tanto individual como del equipo. Por otra parte, no parece que el hecho de coevaluarse provoque situaciones de conflicto, aunque se observa un porcentaje de estudiantes (Pregunta 4, TA: 8,4% y A: 5,3%) que sí que señalan lo contrario. Esto indica que el sistema diseñado puede facilitar la detección temprana de situaciones conflictivas en los equipos. Finalmente, el hecho de aplicar la evaluación periódicamente durante el proyecto ha ayudado a que una gran parte de los estudiantes no se desconectase, lo que pensamos que de alguna manera ha favorecido que no se den muchos conflictos y que se mejore el desempeño general (Pregunta 5, TA: 24,2% y A: 38,9%). Este cuestionario lo hemos aplicado en varios cursos y asignaturas y los resultados son bastante similares al expuesto como ejemplo.

4. CONCLUSIONES

En este trabajo hemos presentado un método para la evaluación de desempeño de integrantes de un equipo de trabajo, en un contexto de enseñanza basada en ABP. Este método de evaluación integra varias perspectivas de evaluación: la externa (desde la visión del PO y el SM) y la interna (desde la visión del integrante y de sus compañeros de equipo). Gracias al proceso iterativo, al terminar cada interacción (Sprint) se tiene la oportunidad para aplicar la evaluación y reflexionar al respecto. De esta forma se consigue una evaluación continua y formativa pues el estudiantado recibe la valoración de sus profesores (PO y SM) y cada integrante

además recibe la valoración de su desempeño por parte de sus compañeros de equipo. Estas valoraciones permiten mejorar el desempeño durante el desarrollo de la asignatura. El apoyo brindado por la herramienta Worki permite que la introducción y visualización de la información sea muy sencilla, tanto para el estudiantado como para el profesorado. Además, existe total transparencia y precisión respecto de los criterios de evaluación pues están accesibles en Worki desde el inicio de la asignatura. Por las encuestas que hemos realizado al finalizar cada asignatura podemos afirmar que ha tenido una muy buena acogida por los estudiantes. Actualmente estamos trabajando en incorporar mecanismos para detectar anomalías en la coevaluación, entre ellas: conflictos entre estudiantes que se coevalúan muy negativamente siendo que sus compañeros no les otorgan mala coevaluación, o cuando las coevaluaciones no son coherentes respecto de la nota del equipo y/o de las autoevaluaciones de los integrantes.

La presentación de este trabajo ha sido financiada por la Escola Tècnica Superior d'Enginyeria Informàtica de la Universitat Politècnica de València y la Generalitat Valenciana a través del proyecto CoMoDID (CIPROM/2021/023).

REFERENCIAS

- Akhtar, A., Bakhtawar, B., & Akhtar, S. (2022). EXTREME PROGRAMMING VS SCRUM: A COMPARISON OF AGILE MODELS. *International Journal of Technology Innovation and Management (IJTIM)*, 2(2), 80-96. <https://doi.org/10.54489/ijtim.v2i2.77>
- Alvarez, S., & Andrés, H. (2021). Metodologías Ágiles-IS248-202101. <http://hdl.handle.net/10757/664702>
- Fernandes, P., Sales, A., Santos, A. R., & Webber, T. (2011). Performance Evaluation of Software Development Teams: A Practical Case Study. *Electronic Notes in Theoretical Computer Science*, 275, 73-92. <https://doi.org/10.1016/j.entcs.2011.09.006>
- Hamer, S., Quesada-López, C., Martínez, A., & Jenkins, M. (2021). Using git metrics to measure students' and teams' code contributions in software development projects. *CLEI electronic journal*, 24(2). <https://doi.org/10.19153/cleiej.24.2.8>
- Letelier Torres, P., & Cornide Reyes, H. (2020). Evaluación de equipos de trabajo en el marco de Aprendizaje Basado en Proyectos y Métodos Ágiles. *Actas de las Jenui*, 5, 165-172. https://aenui.org/actas/pdf/JENUI_2020_023.pdf
- López Trujillo, Y., André Ampuero, M., Infante Abreu, A. L., Escalera Fariñas, K., & Verona Marcos, S. (2018). Evaluación del desempeño de roles en equipos de desarrollo de software. Utilización de escalas de valoración. *Ingeniare. Revista chilena de ingeniería*, 26(3), 486-498. <http://dx.doi.org/10.4067/S0718-33052018000300486>
- Matthies, C., Kowark, T., Uflacker, M., & Plattner, H. (2016). Agile metrics for a university software engineering course. *2016 IEEE Frontiers in Education Conference (FIE)*, 1, 1-5. <https://doi.org/10.1109/fie.2016.7757684>
- Schwaber, K., & Sutherland, J. (2020). The Scrum Guide. Recuperado de: <https://scrumguides.org>.