

Jornada de Innovación: Big Data en Educación

Mejorando el proceso del aprendizaje con técnicas de minería de datos

Pedro J. Muñoz-Merino

Email: pedmume@it.uc3m.es

Twitter: [@pedmume](https://twitter.com/pedmume)

- *"Quite simply, knowledge is the engine of our economy. And data is its fuel."*



Neelie Kroes

Former Vicepresident
European Commission,
Responsible for Digital Agenda

- *"Data is a precious thing and will last longer than the systems themselves."*



Tim Berners-Lee

Director W3C

Inventor of the WWW

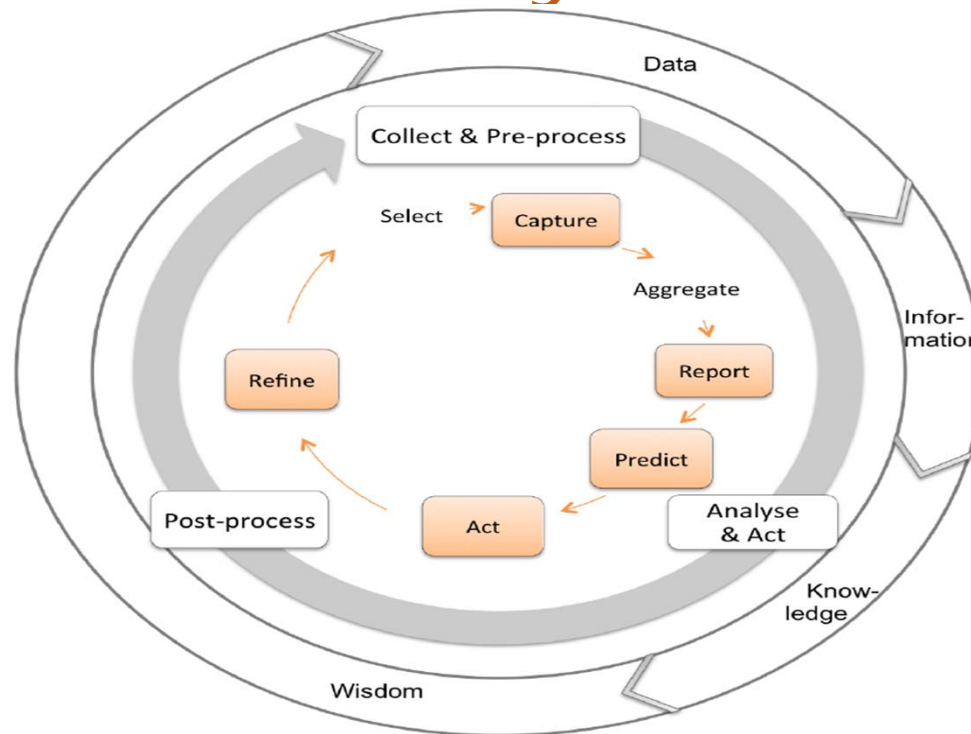
- Sector educativo

- Mejorar como aprenden los alumnos
- Mejorar interés de los alumnos
- Mejorar habilidades de los alumnos
- Mejorar a los profesores
- Toma de decisiones en la institución
- Mejorar los contenidos y las actividades
- Mejorar las metodologías y pedagogías
- Poder abarcar a un mayor número de personas

- Analítica del aprendizaje
- Minería de datos educativa
- Análisis de datos educativos
- Analítica de negocio
- Descubrimiento de conocimiento en los datos

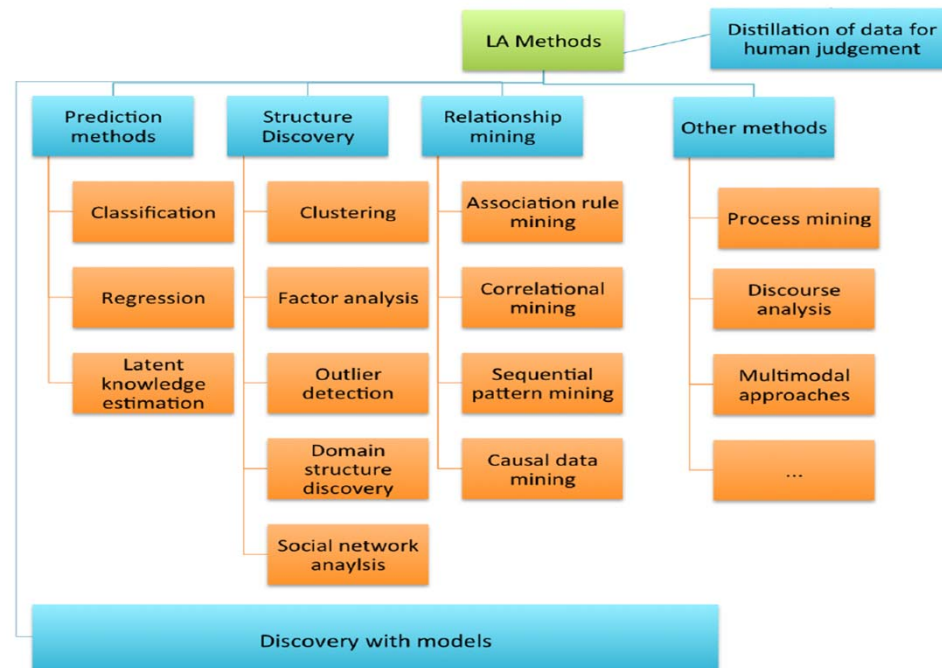
- LAK: International Conference on Learning Analytics and Knowledge
- EDM: International Conference on Educational Data Mining
- L@S: Learning at Scale
- ITS: Intelligent Tutoring Systems
- AIED: Artificial Intelligence in Education
- EC-TEL: European Conference on Technology Enhanced Learning
- eMOOCs

Fases de learning analytics



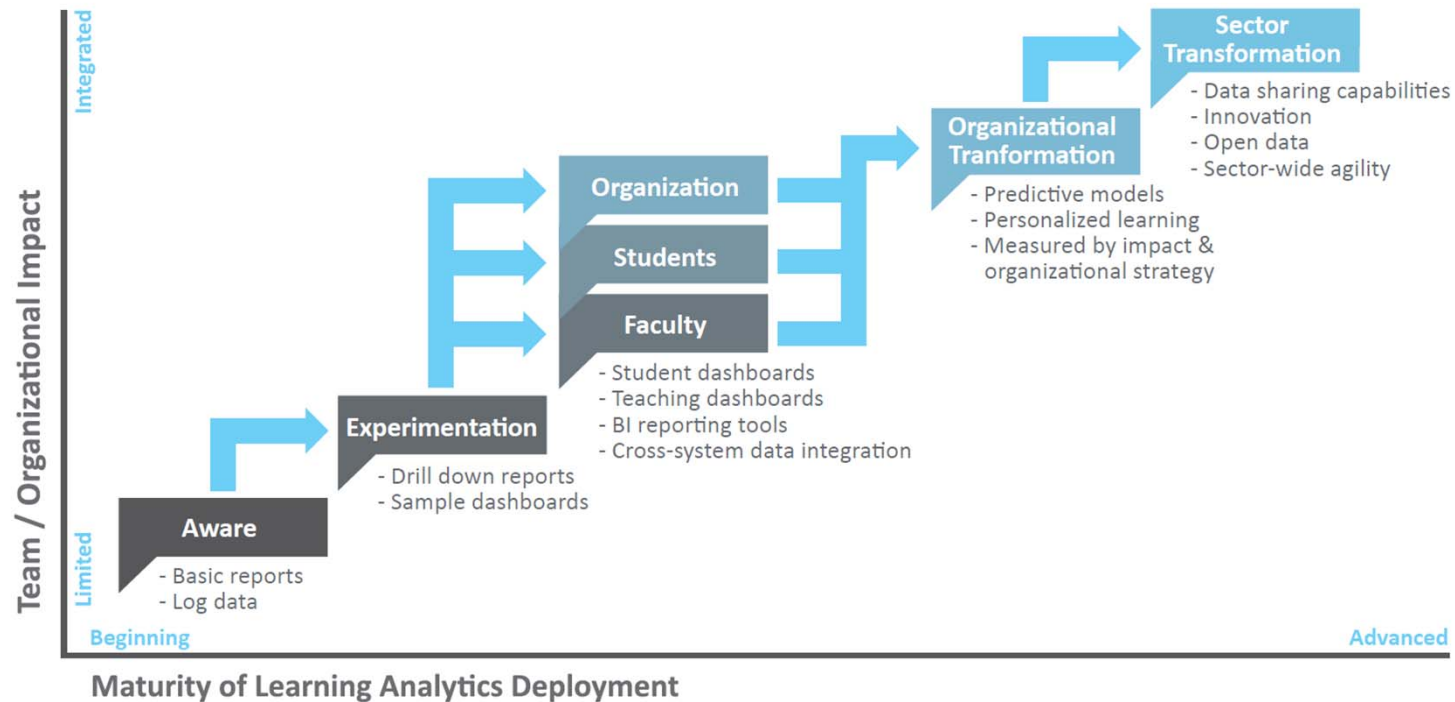
Steiner, C., Kickmeier-Rust, M., Türker, M. A. (2014). Review Article about LA and EDM Approaches, Deliverable D3.1 – LEA’s Box Project, <http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D3.1.pdf>

Métodos en learning analytics



Steiner, C., Kickmeier-Rust, M., Türker, M. A. (2014). Review Article about LA and EDM Approaches, Deliverable D3.1 – LEA’s Box Project, <http://css-kmi.tugraz.at/mkrwww/leas-box/downloads/D3.1.pdf>

Adopción de learning analytics



Siemens, G., Dawson, S., & Lynch, G. (2014). Improving the Quality and Productivity of the Higher Education Sector - Policy and Strategy for Systems-Level Deployment of Learning Analytics. Canberra, Australia: Office of Learning and Teaching, Australian Government. Retrieved from http://solaresearch.org/Policy_Strategy_Analytics.pdf

Política de learning analytics

- Aspectos a considerar
 - Privacidad
 - Seguridad
 - Objetivos de la organización
 - Transmisión de la importancia
 - Aplicaciones y herramientas
 - Metodologías
 - Entrenamiento a profesores y alumnos
 - ...

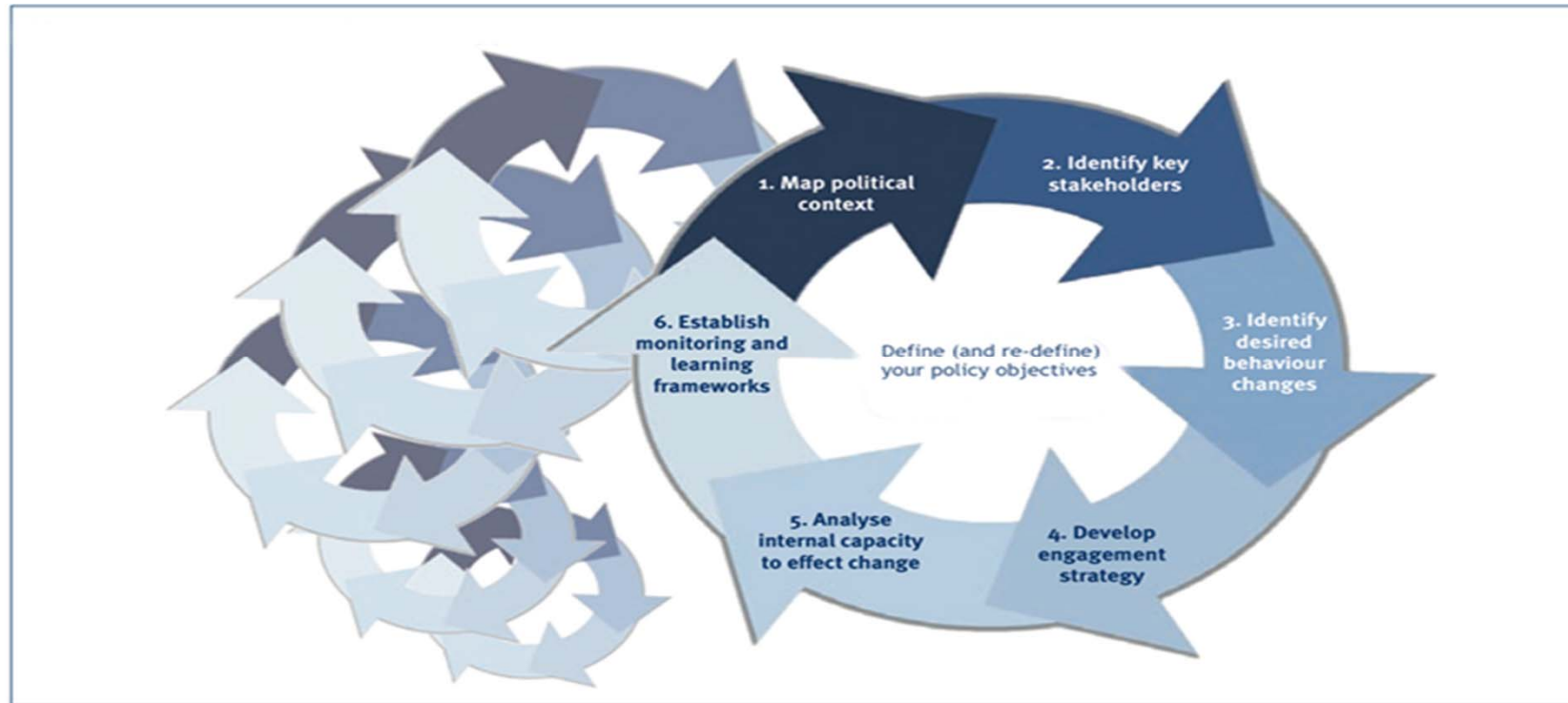
- Constitución española (1978)
 - Artículo 18:
Derecho a la privacidad personal
- AEPD: Agencia Española de Protección de Datos
 - Desde 1993
- Directiva de protección de datos 95/46/EC
- LOPD: Ley Orgánica de Protección de Datos 15/1999
 - Derecho acceso, modificación, cancelación
- Agencia de Protección de Datos de la Comunidad de Madrid
 - 2001-2012
- Reglamento de Medidas de Seguridad (RD 1720/2007)
 - Medida de seguridad de los datos
- Universidades: Comisión LOPD
 - Da recomendaciones y aplica ley en universidades

- SHEILA: Supporting Higher Education to Integrate Learning Analytics
- Partners
 - Brussels Education Services (BES)
 - University of Edinburgh (UE)
 - Open University Netherlands (OUNL)
 - Tallinn University (TLU)
 - Universidad Carlos III Madrid (UC3M)
 - European Association for Quality Assurance in Higher Education (ENQA)
 - Erasmus Student Network (ESN)

SHEILA: Objetivos

- Conocer las perspectivas de los principales actores en Europa
- Conocer las evidencias de aproximaciones de learning analytics efectivas en instituciones
- Proponer un framework para el desarrollo de políticas de learning analytics
- Llevar a cabo casos de estudio que sigan ese framework
- Documentar y diseminar la experiencia

SHEILA: Metodología



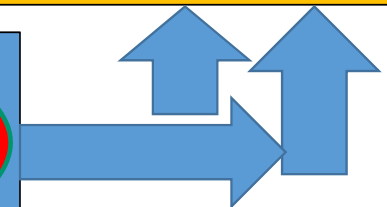
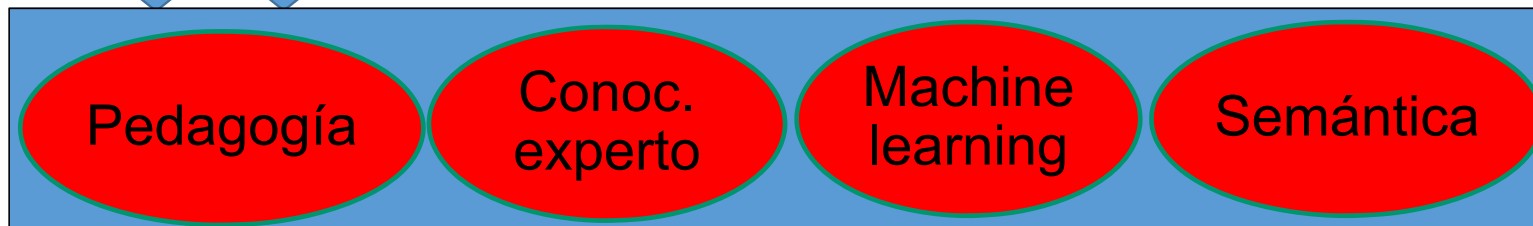
Macfadyen, L., Dawson, S., Pardo, A., Gašević, D., (2014). The learning analytics imperative and the sociotechnical challenge: Policy for complex systems. *Research & Practice in Assessment*, 9(Winter 2014), 17-28.

Ejemplos de Aplicaciones

- 1) Inferencia de información inteligente
- 2) Sistemas de visualización en plataformas e-learning
- 3) Evaluación de experiencias del aprendizaje
- 4) Adaptación del aprendizaje. Aprendizaje personalizado
- 5) Sistemas de predicción y recomendación

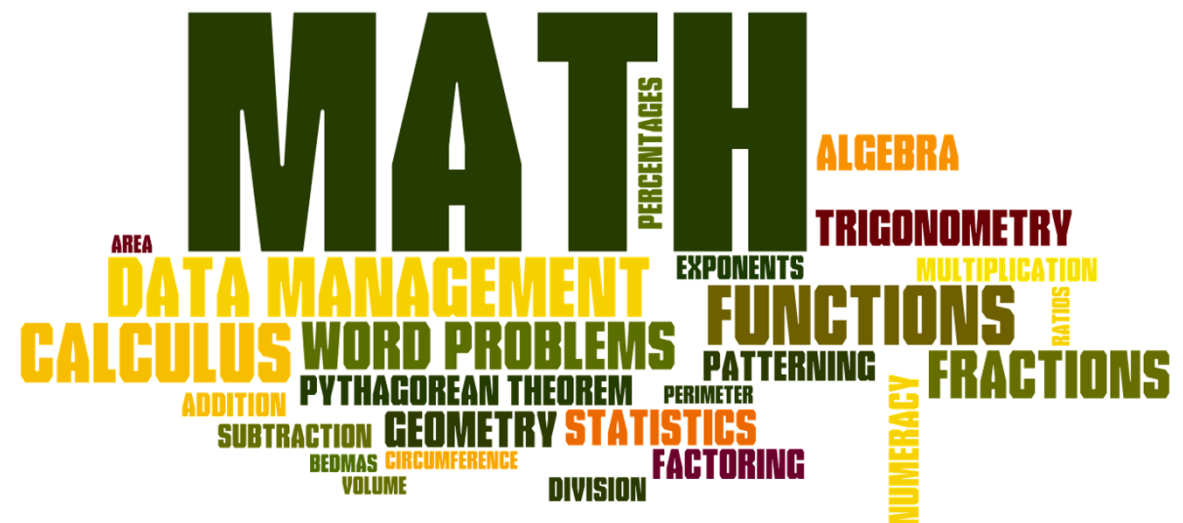
1) Inferencia de información inteligente

Imagen tomada de <http://www.datadial.net/blog/index.php/2011/08/24/why-your-website-isnt-as-fast-as-it-should-be/>



- Inferencia de habilidades de un alumno:

1. *Knowledge Spaces*
2. *Item Response Theory*
3. *Bayesian networks*
4. *Semantic based*

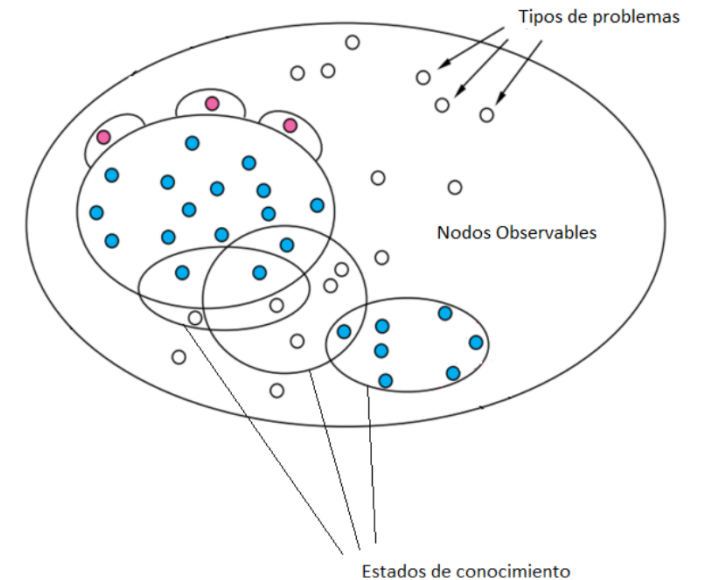


Knowledge spaces

¿Qué sabe un alumno?

¿Qué está preparado un alumno para saber?

- Conocimiento estructurado
- Un número limitado de estados de conocimiento
- Diferentes caminos de aprendizaje



Ejemplo: Modelado de habilidades

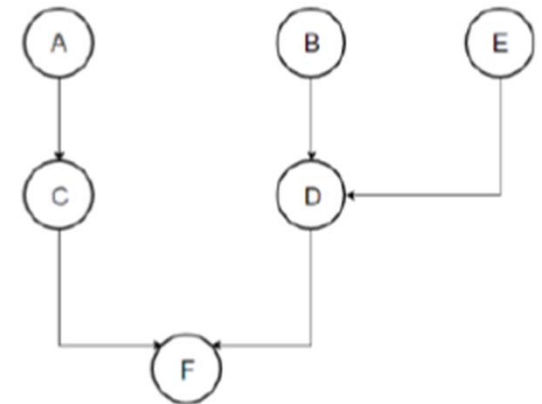
$$K = \left\{ \emptyset, (A), (B), (E), (A, B), (A, E), (B, E), (A, B, E), (A, C), (A, C, B), \right. \\ \left. (A, C, E), (A, C, B, E), (B, E, D), (B, E, D, A), (B, E, D, A, C), Q \right\}$$

Diferentes caminos de aprendizaje \longrightarrow Personalización

$$\text{Path 1} = \{ \emptyset \rightarrow (A) \rightarrow (A, C) \rightarrow (A, C, B) \rightarrow (A, C, B, E) \\ \rightarrow (A, C, B, E, D) \rightarrow Q \}$$

$$\text{Path 2} = \{ \emptyset \rightarrow (B) \rightarrow (B, E) \rightarrow (B, E, D) \rightarrow (B, E, D, A) \\ \rightarrow (A, C, B, E, D) \rightarrow Q \}$$

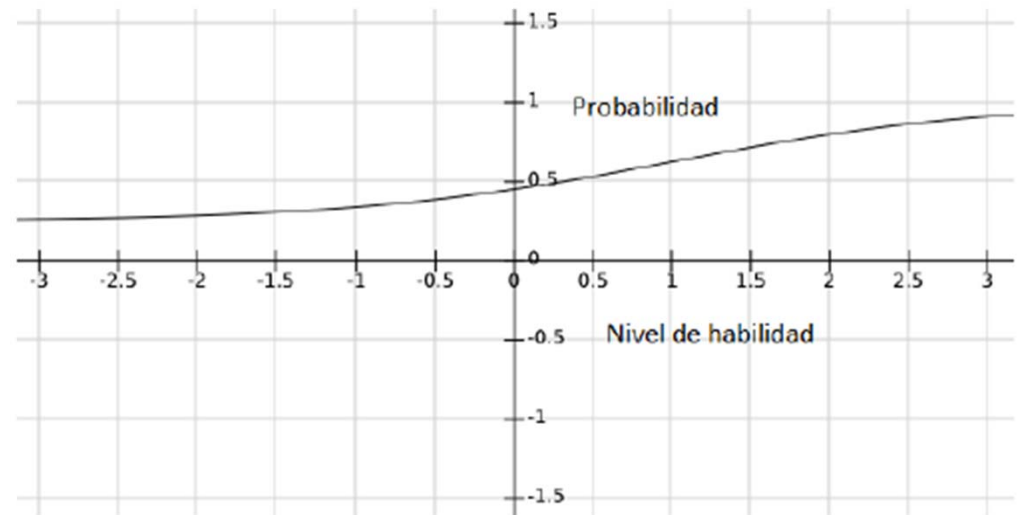
$$\text{Path 3} = \{ \emptyset \rightarrow (A) \rightarrow (A, B) \rightarrow (A, C, B) \rightarrow (A, C, B, E) \\ \rightarrow (A, C, B, E, D) \rightarrow Q \}$$



Item Response Theory

- Solo una habilidad
- Uno, dos o tres parámetros:
 1. Dificultad
 2. Pendiente Necesidad de calibración
 3. Probabilidad de adivinar

ITEM CHARACTERISTIC CURVE (ICC)



Ejemplo: Modelado de habilidades

ITEM RESPONSE FUNCTION (IRF)

$$P(x_i|\theta) = c + \frac{1 - c}{1 + e^{-a_i(\theta - \beta_i)}}$$

- Probabilidad de responder correctamente un ítem dependiendo de la habilidad
- Independencia local de los ítems
- Estimación de la habilidad del alumno

Ejemplo: Modelado de habilidades

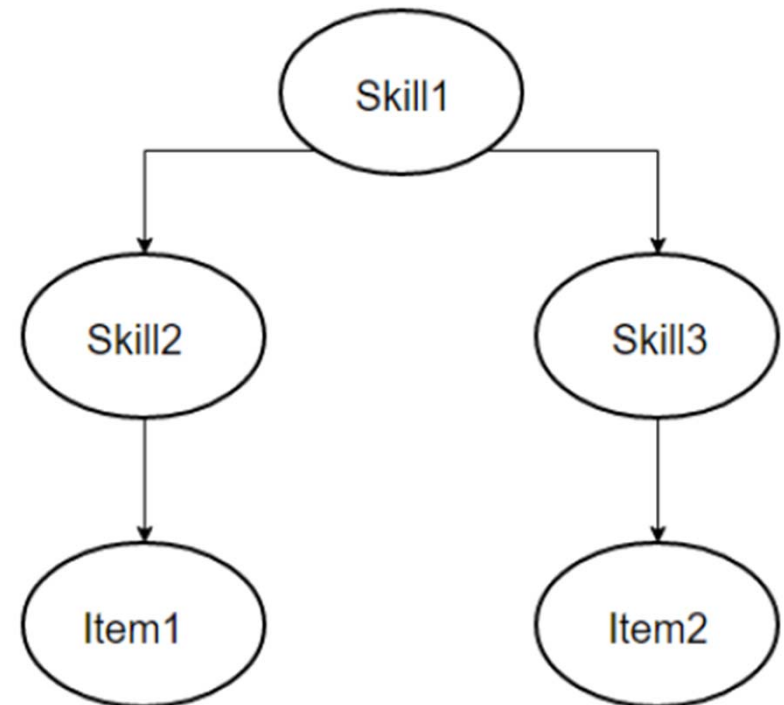
Redes bayesianas

- Los nodos son habitualmente ejercicios y habilidades
- Cada nodo tiene una tabla condicional de probabilidad que denota la probabilidad de realizar correctamente un ejercicio dependiendo de los resultados de los nodos padre
- Teorema de Bayes para realizar la inferencia
- Independencia condicional



Redes bayesianas

- Un evento actualiza la red
- Es posible calcular las diferentes probabilidades



BAYESIAN NETWORKS

- + Más información semántica
- + Varias habilidades
- + Más complejidad
- + Necesidad de realizar etiquetados, realización de la estructura y asignación de probabilidad

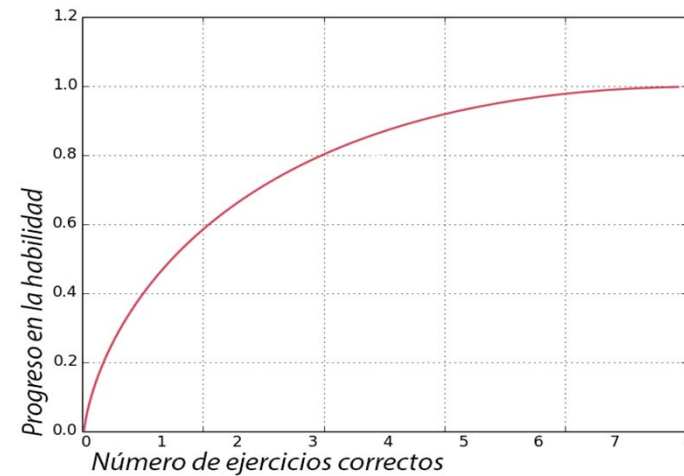
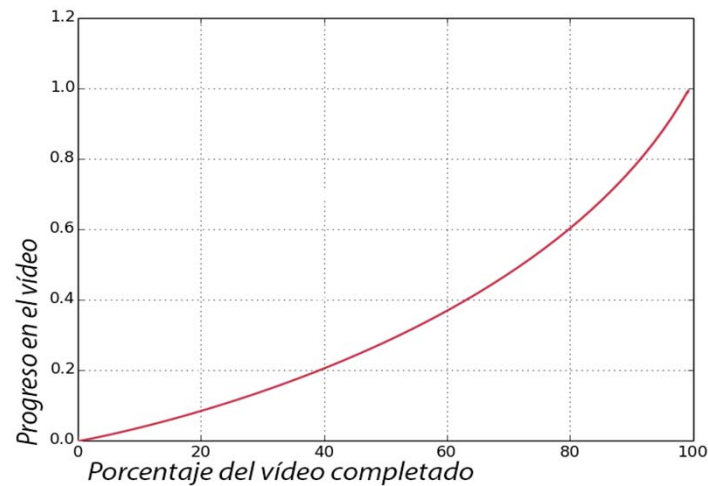
ITEM RESPONSE THEORY

- +No hay información semántica entre contenidos
- +No hay necesidad de construir una red
- +Una sola habilidad
- +Calibración de parámetros

KNOWLEDGE SPACES

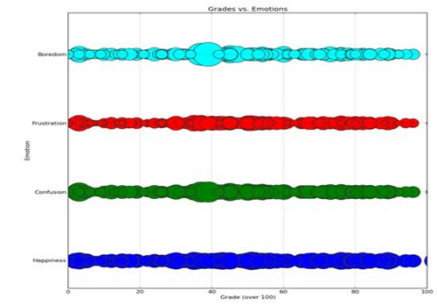
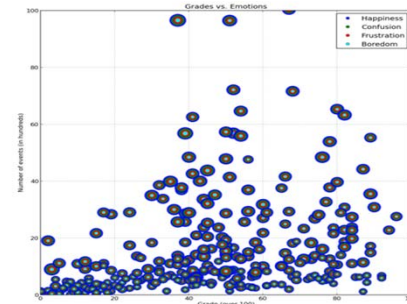
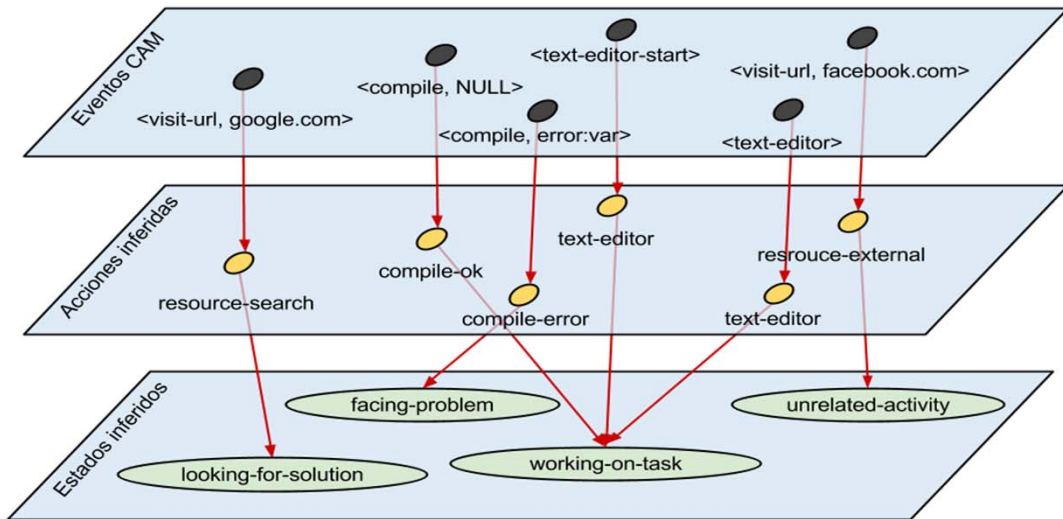
- +No hay necesidad de calibración
- +Predefinida estructura de conocimiento
- +Sin nodos escondidos

Ejemplo: Efectividad



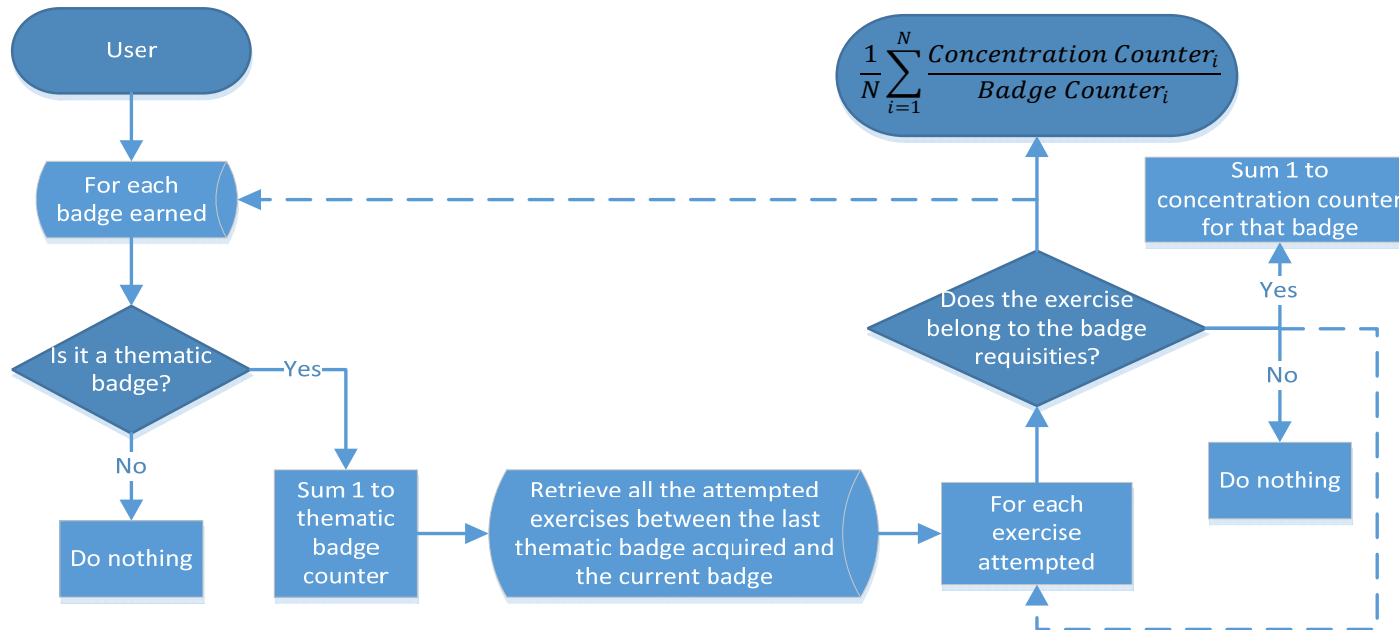
PJ Muñoz-Merino, JA Ruipérez-Valiente, C Alario-Hoyos, M Pérez-Sanagustín, C Delgado Kloos: *"Precise effectiveness strategy for analyzing the effectiveness of students with educational resources and activities in MOOCs"*, Computers in Human Behavior, vol. 47, pp. 108–118 (2015)

Ejemplo: Emociones



D Leony, PJ Muñoz-Merino, A Pardo, C Delgado Kloos: *“Provision of awareness of learners’ emotions through visualizations in a computer interaction-based environment”*, Expert Systems With Applications, vol. 40, no. 3 (2013), pp. 5093-5100

Ejemplo: Gamificación



D Leony, PJ Muñoz-Merino, A Pardo, C Delgado Kloos: *“Provision of awareness of learners’ emotions through visualizations in a computer interaction-based environment”*, Expert Systems with Applications, 40:3 (2013), 5093-5100

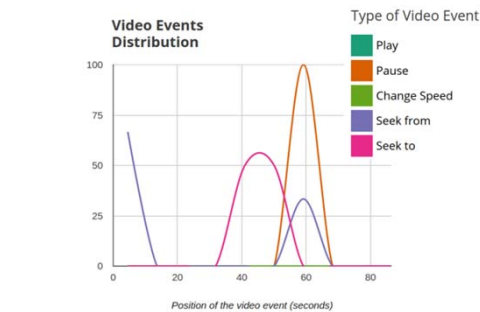
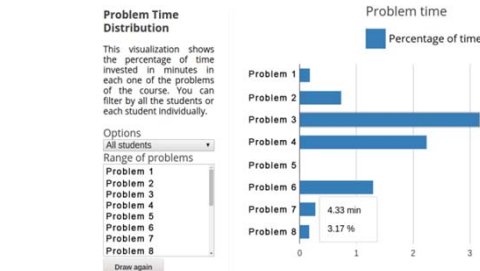
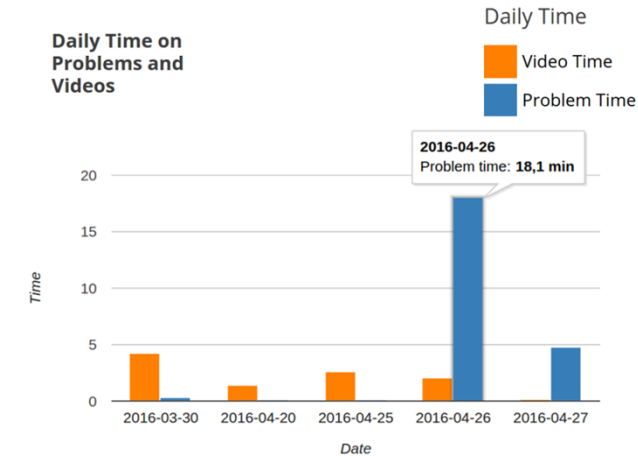
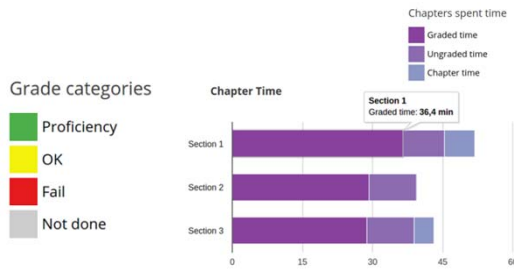
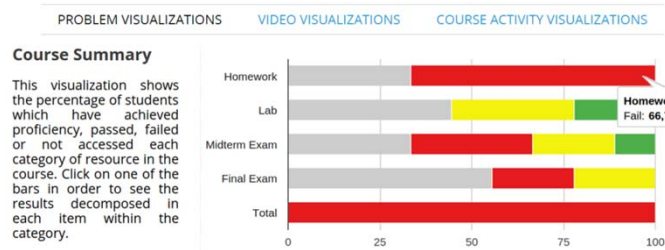
2) Sistemas de Visualización



- El grupo ha realizado módulos para las plataformas
 - Open edX -> ANALYSE
 - Khan Academy -> ALAS-KA
 - Entorno programación
 - Google Course Builder
- Evaluación de usabilidad, efectividad y utilidad
- Para gestores, profesores y alumnos

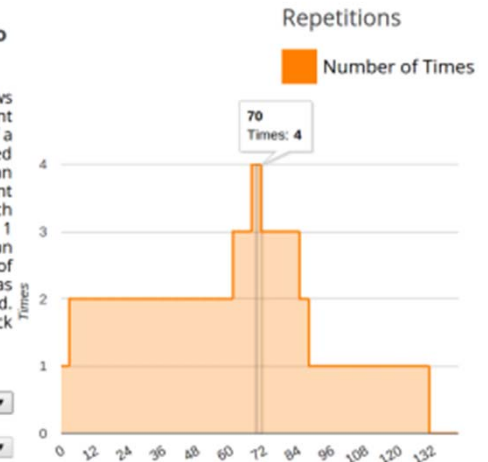
2) Sistemas de Visualización

DASHBOARD FOR INSTRUCTORS



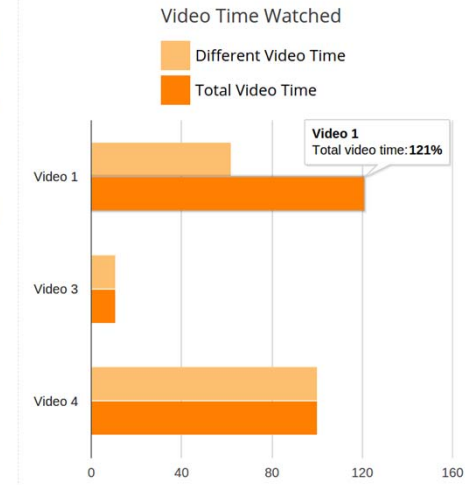
Repetitions of Video Intervals

This visualization shows the number of different times that each second of a video have been watched by all the class. You can select two different options; either each student can count only 1 time, or each student can count up to the number of times that he/she has watched that video second. You can also pick students individually.

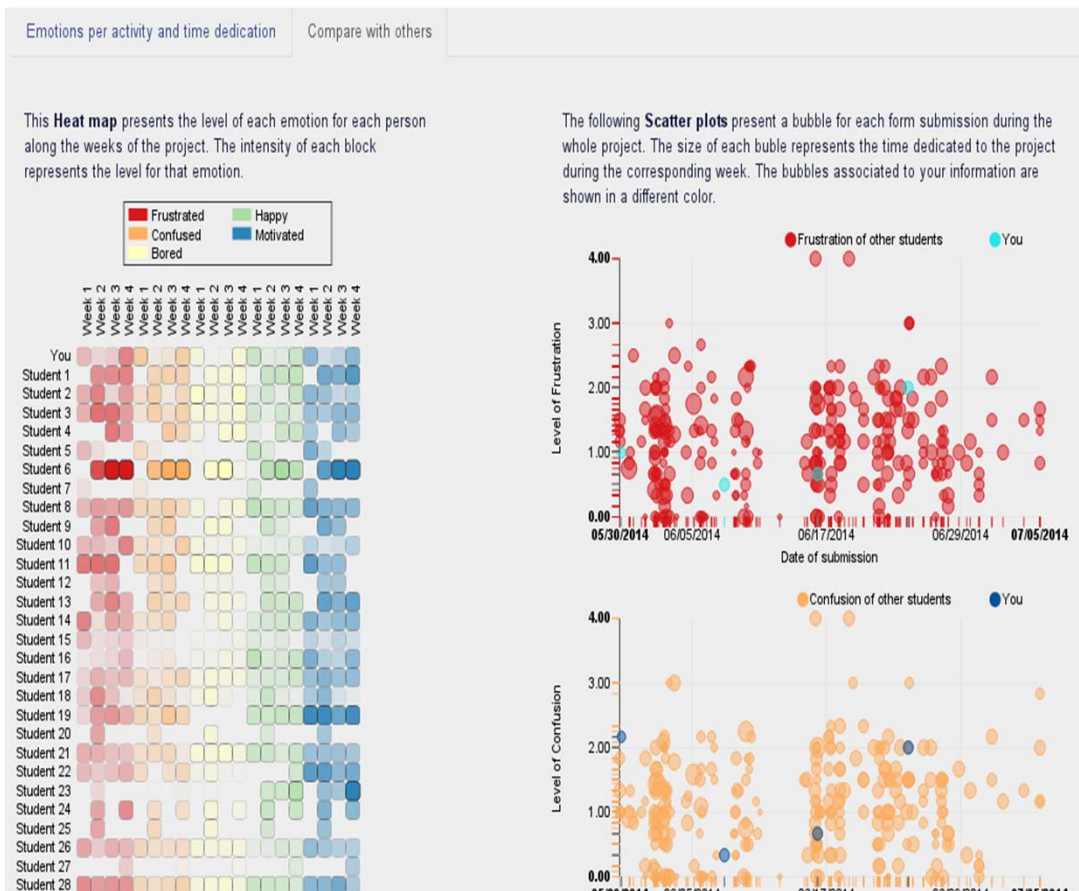


Video Time Watched

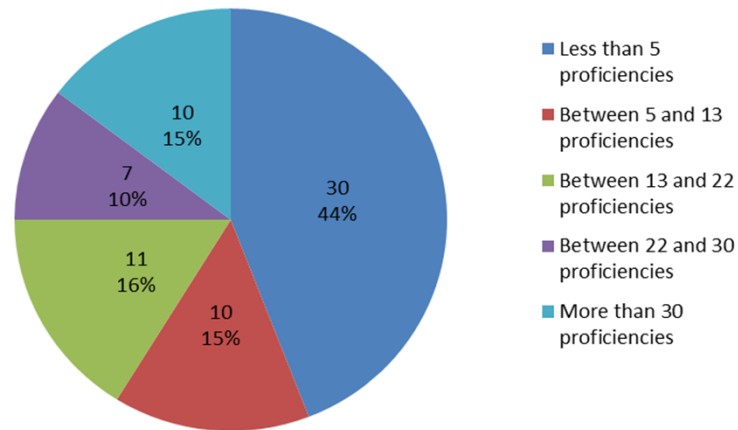
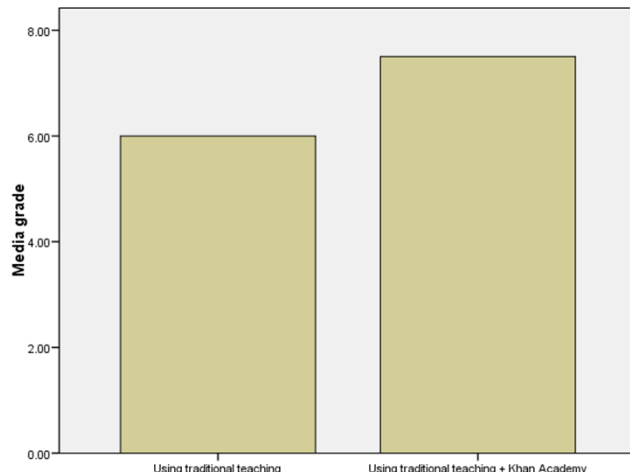
This visualization shows in light orange the percentage of different video watched (it does not count if the same parts are watched several times, the max. is the 100% of the video) and in dark orange the percentage of total video watched (total amount of time spent compared to the length of the video in percentage). You can filter by all the students or each student individually.



2) Sistemas de Visualización



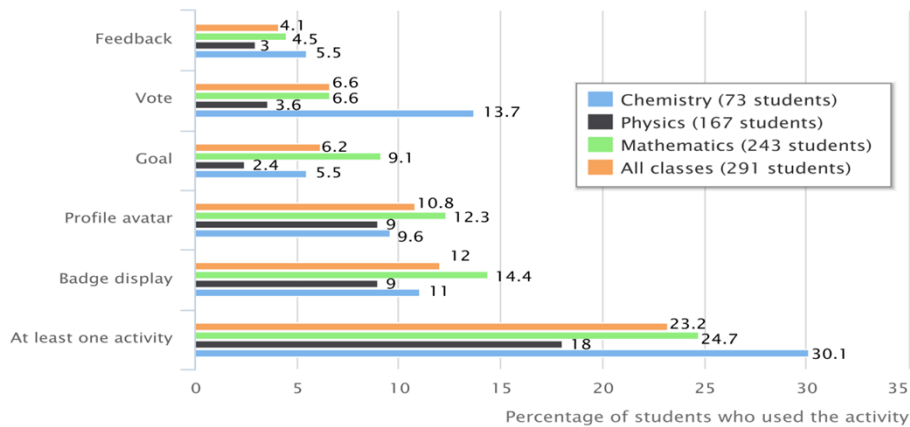
3) Evaluación del proceso de aprendizaje



- El grupo ha participado en la evaluación de multitud de experiencias educativas: universidad, colegios, empresas
- Para la evaluación utilizamos: indicadores de sistemas, tests, encuestas, entrevistas, etc.
- Se evalúan materiales, alumnos, herramientas y el proceso del aprendizaje

3) Ejemplo: Evaluación del aprendizaje

Use of Optional Activities by Class

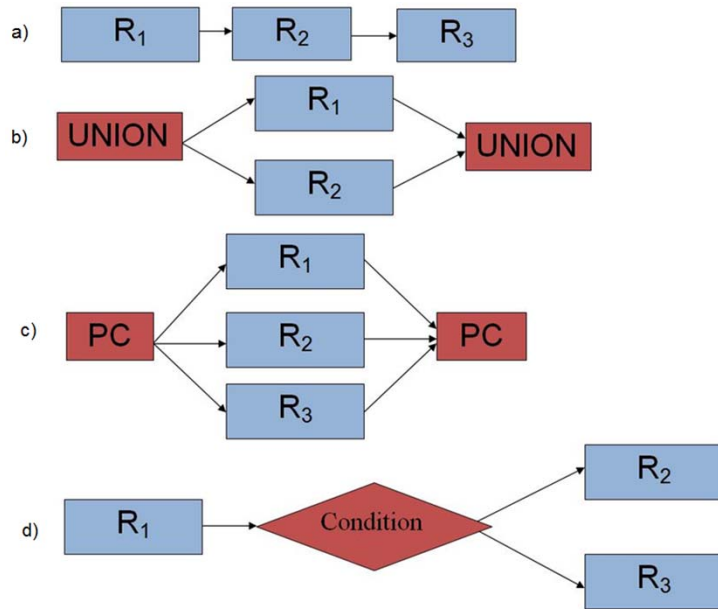


- Los alumnos completaron los objetivos que se pusieron en más del 50% de las ocasiones
- Los alumnos votaron los comentarios de sus compañeros de una manera positiva
- El género y el tipo de curso tuvieron influencia en qué actividades opcionales fueron utilizadas

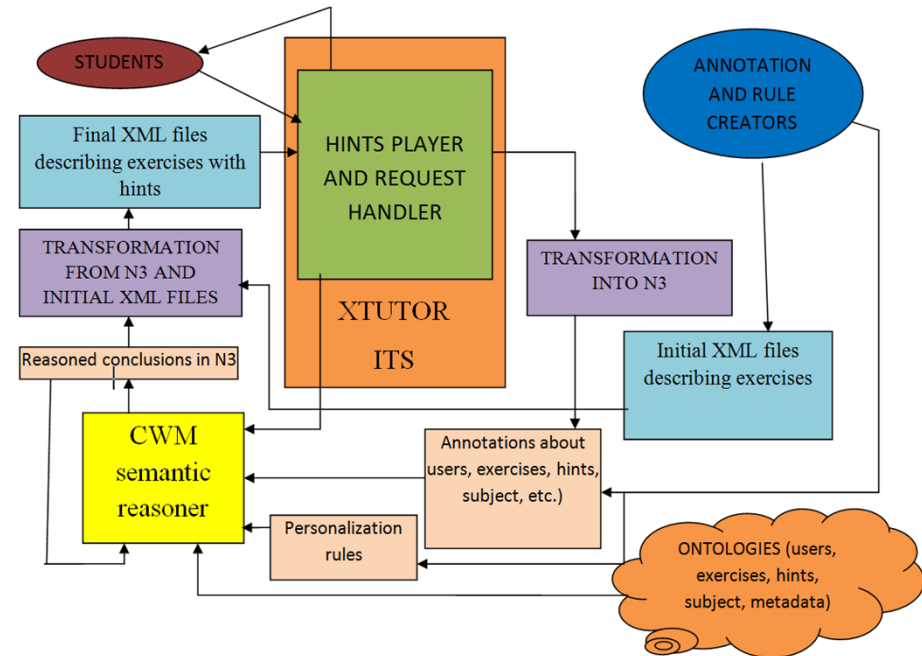
Pearson Correlation. Learning gain sig. (2-tailed) N = 69	Optional activities: 0.293** (p=0.015)	Goal: 0.102 (p=0.406)	Feedback: 0.219 (p=0.071)
	Vote: 0.333* (p=0.005)	Avatar: 0.221 (p=0.068)	Display badges: 0.296** (p=0.013)
***Partial Correlation. Learning gain sig. (2-tailed) N = 69	Optional activities: 0.142 (p=0.260)	Goal: - 0.070 (p=0.581)	Feedback: 0.124 (p=0.323)
	Vote: 0.214 (p=0.087)	Avatar: 0.170 (p=0.176)	Display badges: 0.261 (p=0.036)

$$LG = \begin{cases} \frac{post_{test} - pre_{test}}{max_{score} - pre_{test}}, & pre_{test} \leq post_{test} \\ \frac{post_{test} - pre_{test}}{pre_{test}}, & pre_{test} > post_{test} \end{cases}$$

4) Adaptación del aprendizaje. Aprendizaje personalizado



- Arquitecturas para realización de reglas adaptativas



- Adaptación de contenidos, usuarios, etc. en base a diferentes modelos de usuario

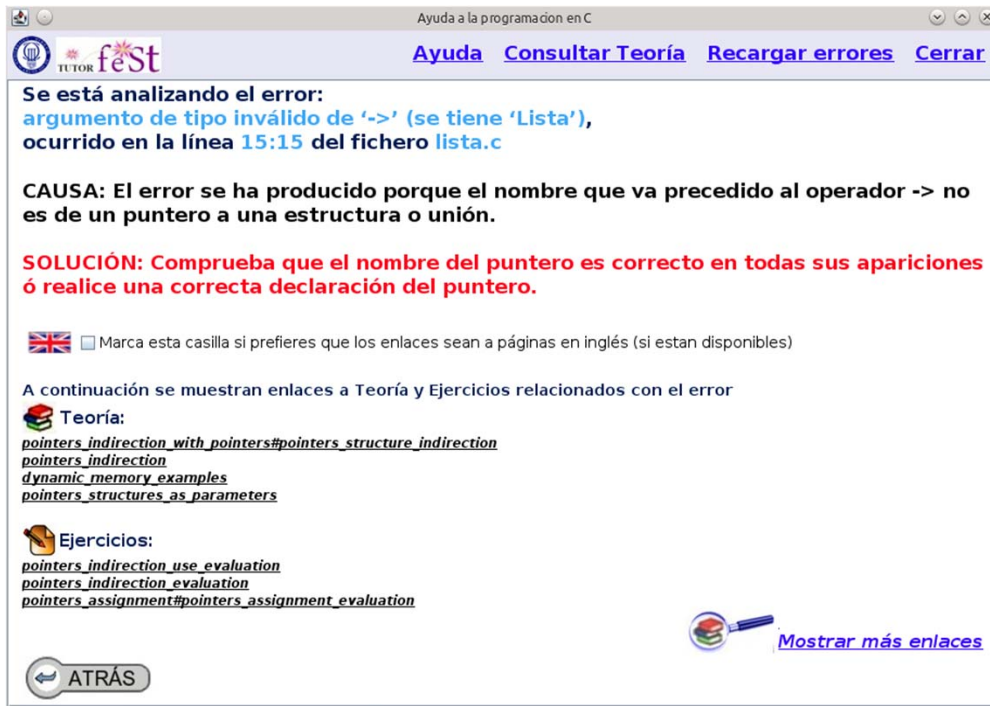
4) Adaptación del aprendizaje. Aprendizaje personalizado

Problem data: <ul style="list-style-type: none">- Title: Lenguaje Java- Problem max. score: 1.0- Type: choiceMultiple	Time left: 8 min, 3 seg
Problem: <p>Select the correct answers about Java:</p> <ul style="list-style-type: none"><input type="checkbox"/> It is an object oriented programming language<input type="checkbox"/> It is a portable programming language<input type="checkbox"/> Generate code that depends on the architecture<input type="checkbox"/> It is a script language, this is the origin of JavaScript<input type="checkbox"/> None of the previous answers	My information: <ul style="list-style-type: none">- NIA: 800000084- Name: 800000084- Average score: 0.0%- Tournament Points: 0.0 / 2.0- Round points: 0.0 / 7.0- Resolving: 1st problem Opponent information: <ul style="list-style-type: none">- NIA: 800000085- Name: 800000085- Average score: 0.0%- Tournament Points: 0.0 / 2.0- Round points: 1.0 / 7.0- Resolving: 4th problem

Submit answer

Next problem

5) Sistemas de predicción y recomendación




Ayuda a la programación en C

TUTOR **feSt** [Ayuda](#) [Consultar Teoría](#) [Recargar errores](#) [Cerrar](#)


Se está analizando el error:
argumento de tipo inválido de '->' (se tiene 'Lista'),
ocurrido en la línea **15:15** del fichero **lista.c**


CAUSA: El error se ha producido porque el nombre que va precedido al operador -> no es de un puntero a una estructura o unión.


SOLUCIÓN: Comprueba que el nombre del puntero es correcto en todas sus apariciones ó realice una correcta declaración del puntero.


 Marca esta casilla si prefieres que los enlaces sean a páginas en inglés (si están disponibles)

A continuación se muestran enlaces a Teoría y Ejercicios relacionados con el error

 Teoría:
[pointers_indirection_with_pointers#pointers_structure_indirection](#)
[pointers_indirection](#)
[dynamic_memory_examples](#)
[pointers_structures_as_parameters](#)

 Ejercicios:
[pointers_indirection_use_evaluation](#)
[pointers_indirection_evaluation](#)
[pointers_assignment#pointers_assignment_evaluation](#)

 [Mostrar más enlaces](#)

 ATRÁS

- Predicción sobre:
 - Ganancias de aprendizaje
 - Obtención de certificados
 - Actividad social
 - Realización de trampas en el sistema
 - Comportamientos en el sistema
- Recomendación sobre:
 - Recursos de diferente tipo
 - Otros participantes

Diseño e Implementación de aplicaciones telemáticas

- Aplicaciones de e-learning
- Aspectos eficiencia de los procesados
- Interoperabilidad de los datos
- Tecnologías
 - Web: J2EE, PHP, Django, ...
 - Base de datos: mysql, MongoDB, Oracle, ...
 - Cliente: HTML, JavaScript, CSS, ...

Proyectos de investigación activos



Proyectos de investigación activos relacionados con analítica del aprendizaje



Universidad
Carlos III de Madrid

- “SHEILA: Supporting Higher Education to Incorporate Learning Analytics”, Proyecto Europeo. 2016-2018
- MOOC-Maker, Proyecto europeo, 2015-2018
- SNOLA: Spanish Network of Learning Analytics. Proyecto RETOS-Redes
- “Aplicaciones de analítica y sistemas de recomendación”. Proyecto artículo 83. 2016-2017
- “Análisis, evolución, propuestas de mejora y desarrollo del sistema de aprendizaje adaptativo y analítica del aprendizaje de la plataforma smartick”. Proyecto artículo 83. 2015-2019
- “mapaTIC”, Proyecto Nacional. 2015-2016
- “eMadrid: Investigación y Desarrollo de tecnologías educativas en la Comunidad de Madrid,” 2014-2018
- “RESET”, Proyecto RETOS-Investigación. 2015-2017
- “Simlap”, Proyecto RETOS-Colaboración, 2014-2016

Jornada de Innovación: Big Data en Educación

Mejorando el proceso del aprendizaje con técnicas de minería de datos

Pedro J. Muñoz-Merino

Email: pedmume@it.uc3m.es

Twitter: [@pedmume](https://twitter.com/pedmume)