استخدام أدوات ذكاء الأعمال لتحليل البيانات الناتجة عن الأنظمة التعليمية الافتراضية

تقديم

آلاء الحموي محمد باسل الشمالي

محمد غانم هانية المالكي

# تمهيد:

كان تأثير التطور التكنولوجي الذي اجتاح معظم مناحي الحياة على المجال التعليمي كبير، فبدأت المؤسسات التعليمية والأكاديمية باقتحام هذه التكنولوجيا في طرق التعلم، ومع مرور الوقت ظهرت العديد من الجامعات التي تعتمد بشكل كامل على الحاسوب والإنترنت، مما سهل عملية التعلم وجعلها متاحة لأي فرد أينما كان. إلا أن التحديات التي واجهت هذا النوع من التعلم كثيرة لانعدام التواصل المباشر بين المعلمين والطلاب وأعداد المتعلمين الهائلة والتي أدت لصعوبة ضبط المتعلمين ومساعدتهم. ونتيجة لامتلاك الأنظمة التعليمية التي تعتمد بطبيعتها على الحاسوب القدرة على تسجيل كامل تفاصيل تفاعل الطالب مع النظام، ولما كانت تأثير هذه البيانات على دعم القرار المتخذ في المؤسسات التعليمية والأكاديمية كبير، كانت الحاجة للاستفادة من هذه البيانات الخام المجنية من هذه الأنظمة وتحليلها بهدف تخطي التحديات التي تقف فيه وجه التعلم الالكتروني والوصول للمعرفة المخبئة في هذه البيانات وتحسين وضع الطلاب.

يمكن تحقيق بذلك باتباع بعض الطرق والاستراتيجيات *مثل* التنقيب عن المعرفة في المجال التعليميand Learning Analytics Educational Data Mining وهو مجال بحثي جديد يُعرّف على أنه استخدام تقنيات التنقيب في البيانات الخام المستخرجة من الأنظمة والعمليات التعليمية بهدف مساعدة مصممي ومستخدمي هذه الأنظمة لتحسين أدائها واستخلاص معلومات مفيدة عن سلوك الطلاب في عملية التعلّم. كما أن استخدام تقنيات التنقيب في أنظمة التعلم الافتراضية يمكن أن يلعب دوراً هاماً وفعالاً في حل مشاكل التعلم وتقديم المعلومات التي تساعد أصحاب القرار من طلاب ومعلمين وواضعي الخطة الدراسية باتخاذ قرارات مدعومة بالمعلومات عن كيفية التفاعل مع المصادر التعليمية.

وبذلك تكون عملية التنقيب السابقة تطبيق للأساليب الآلية بغية الكشف عن الأنماط في مجموعات كبيرة من البيانات التعليمية. قد يكون من الصعب أو المستحيل تحقيق ذلك نتيجة حجم البيانات الهائلة المستخلصة من الأنظمة التعليمية، إلا أن الجهود قد تكثفت في الآونة الأخيرة لإيجاد أدوات تسهل وتمهد الطريق أمام التنقيب بكمية هائلة من المعطيات وتعمل على تجهيز هذه المعطيات لمرحلة التنقيب قبل البدء بها.

الفهرس:

[تمهيد: 1](#_Toc479594238)

[الفصل الأول: مقدمة عامة 4](#_Toc479594239)

[مقدمة المشروع: 5](#_Toc479594240)

[توصيف المشكلة: 6](#_Toc479594241)

[هدف المشروع العام: 6](#_Toc479594242)

[دوافع المشروع: 7](#_Toc479594243)

[المتطلبات البرمجية والعتادية للنظام: 8](#_Toc479594244)

[حدود المشروع: 8](#_Toc479594245)

[الحدود الوظيفية: 8](#_Toc479594246)

[الحدود الزمانية: 8](#_Toc479594247)

[الخرج المتوقع للنظام: 9](#_Toc479594248)

[القيود المفروضة على المشروع: 9](#_Toc479594249)

[خاتمة: 10](#_Toc479594250)

[الفصل الثاني: الدراسة النظرية. 11](#_Toc479594251)

[مقدمة: 12](#_Toc479594252)

[التنقيب في المعطيات: 12](#_Toc479594253)

[هدف عمليات التنقيب: 12](#_Toc479594254)

[لمحة تاريخية: 13](#_Toc479594255)

[مجالات التنقيب في البيانات: 14](#_Toc479594256)

[أهداف التنقيب عن المعطيات في الأنظمة التعليمية: 18](#_Toc479594257)

[الطرق المستخدمة في تنقيب البيانات التعليمية: 20](#_Toc479594258)

[تصنيف البيانات المحللة: 22](#_Toc479594259)

[بعض الأدوات المستخدمة في عملية التحليل: 23](#_Toc479594260)

[تحليل التعلّم Learning Analytics: 25](#_Toc479594261)

[الفرق بين التنقيب في المعطيات الناتجة عن الأنظمة التعليمية وتحليل التعلم: 25](#_Toc479594262)

[النموذج المرجعي في تحليل التعلم: 26](#_Toc479594263)

[بعض المعايير المستخدمة في الأنظمة التعليمية: 38](#_Toc479594264)

[Xapi: 38](#_Toc479594265)

[Caliper: 44](#_Toc479594266)

[خاتمة: 45](#_Toc479594267)

[الفصل الثالث: الدراسة المرجعية. 46](#_Toc479594268)

[مقدمة: 47](#_Toc479594269)

[بعض الأنظمة التي تُعنى بمعالجة المعطيات التعليمية: 48](#_Toc479594270)

[Completing the Loop: 48](#_Toc479594271)

[Open University Analyse: 51](#_Toc479594272)

[Student Insight: 55](#_Toc479594273)

[PANDA: 57](#_Toc479594274)

[مقارنة بين الأنظمة المشابهة: 59](#_Toc479594275)

[نقاط الضعف في الأنظمة السابقة: 66](#_Toc479594276)

[النظام المقترح: 67](#_Toc479594277)

[خاتمة: 68](#_Toc479594278)

[الفصل الرابع: الدراسة التحليلية. 69](#_Toc479594279)

[مقدمة: 70](#_Toc479594280)

[دراسة جدوى المشروع: 70](#_Toc479594281)

[الجدوى التقنية: 70](#_Toc479594282)

[الجدوى المالية: 70](#_Toc479594283)

[الجدوى التشغيلية: 71](#_Toc479594284)

[الإطار القانوني: 71](#_Toc479594285)

[الإطار التنظيمي: 71](#_Toc479594286)

[الجدوى التنظيمية: 71](#_Toc479594287)

[جدوى الموارد البشرية: 71](#_Toc479594288)

[الفاعلون في النظام: 71](#_Toc479594289)

[المتطلبات الوظيفية: 72](#_Toc479594290)

[المتطلبات غير الوظيفية: 73](#_Toc479594291)

[بنية النظام: 74](#_Toc479594292)

[الخاتمة: 76](#_Toc479594293)

[الفصل الخامس: الدراسة التصميمية. 77](#_Toc479594294)

[مقدمة: 78](#_Toc479594295)

[عملية ما قبل المعالجة: 78](#_Toc479594296)

[حالات الاستخدام: 79](#_Toc479594297)

# الفصل الأول: مقدمة عامة

## **مقدمة المشروع:**

مر التّعلم عن بُعد بأطوار عديدة حتى وصل إلى شكله الحالي، فقبل ثورة الانترنت كان يتم الاعتماد على تسجيل المحاضرات على أشرطة الفيديو والصوت والتلفاز، وكان التواصل بين الطلاب والمعلم يتم عن طريق التلفون أو الرسائل. إلا أن تأثير ظهور الانترنت في مجال التعلم عن بُعد كان ملحوظ، فلم يعد التعلم محصور بمكان ولا زمان محددين وأتاح مرونة كبيرة بعملية التعلم وأمّن معلومات شاملة عن العملية التعليمية للطلاب مما دفع الباحثين لتكثيف جهودهم للاستفادة من هذه البيانات ومحاولة إيجاد الأسئلة المتعلقة بكيفية التنبؤ بنجاح الطلاب، تسلسل السيناريو الأمثل والأكثر ملائمة لكل طالب في العملية التعليمية، ما هي المؤشرات المستوحاة من سلوك الطلاب والتي تدل على إمكانية نجاح أو فشل الطالب، ما هي صفات البيئة التعليمية الأمثل والتي تساهم إلى حد كبير في نجاح الطلاب.. وما إلى ذلك من الأسئلة التي تؤثر بشكل أو بآخر على الطلاب ونجاحهم أو فشلهم. والاستفادة من أجوبة هذه الأسئلة وغيرها في تكثيف الجهود لرفع مستوى الطالب ومساعدته وتخفيض نسبة الفشل لدى هؤلاء الطلاب.

إنّ مصدر البيانات التي يتم تطبيق عمليات التنقيب والتحليل عليها لا يقتصر على البيانات الناتجة عن الحالات الفردية للطلاب (كسلوك الملاحة في النظام، التمارين التفاعلية والامتحانات)، وإنما يشمل أيضاً البيانات الناتجة من تعاون الطلاب، البيانات الإدارية، أو حتى البيانات الديموغرافية (كالعمر، الجنس..). يمكن أيضاً الاستفادة من البيانات التي تعبر عن تأثر الطلاب والحالات العاطفية لهم والتي يمكن استنتاجها من التعابير المستخدمة وغيرها.

أما عن كمية البيانات التي عادة ما يتم معالجتها فهي ضخمة جداً، فمع تزايد استخدام أجهزة المحمول والتقنيات الـ online في العمليات التعليمية، أُتيحت كمية هائلة من البيانات تِبعاً لإمكانية الحصول على نتيجة كامل العمليات وأشكال تفاعل الطالب مع المحتوى التعليمي وتسجيلها “logging” وجعلها متاحة للأغراض التحليلية. كما أنه أصبح من المتوفر الآن كميات كبيرة من البيانات للباحثين من خلال السجلات العامة مثل مركز “Pittsburgh Science of Learning Center Datashop”.

## توصيف المشكلة:

تزايد الإقبال على التعلم بشكل كبير في الآونة الأخيرة سواء التعليم الأكاديمي عبر الالتحاق بالجامعات أو التعلم الحر عبر التسجيل بالدورات على شبكة الإنترنت. ونتيجة للأعداد الهائلة للطلاب ومع ظهور الإنترنت وسيطرة المعلوماتية على معظم مجالات الحياة بما فيها السيرك التعليمي، توجهت الجامعات والمؤسسات الأكاديمية إلى محاولة أتمتة عملية التعلم قدر الإمكان، فظهرت نظم إدارة التعلم ونظم التعلم الذكي التي جعلت الحاسب جزء لا يتجزأ من العملية التعليمية. إلا أنه ومع التزايد المستمر بعدد الطلاب بات أمر ضبط الطلاب وملاحقتهم وإسدال النصائح لهم عملية شبه مستحيل نتيجة لمحدودية عدد المعلمين والمرشدين في المؤسسات الأكاديمية، وهنا ظهرت الحاجة إلى تقديم العون إلى المرشدين والمعلمين عن طريق تسخير البيانات الناتجة عن تفاعل الطلاب مع هذه الأنظمة التعليمية والعلامات وغيرها بتقديم معرفة تساعد وتسهل عملية الإرشاد قدر الإمكان، وتعود بالنفع على الطلاب والمؤسسات التعليمية وتساعد أصحاب القرار باتخاذ القرارات الصائبة بالوقت الصحيح بناء على المتغيرات التي تفرضها أوضاع الطلاب.

## هدف المشروع العام:

يعمل المشروع على الاستفادة من البيانات التعليمية التي قد يختلف مصدرها باختلاف التجربة التعليمية المتبعة من بيانات ناتجة عن تفاعل الطلاب مع بيئات التعلم الافتراضية، بيانات العلامات والمعدلات الناتجة عن الاختبارات وغيرها في خلق وسائل تعمل على:

* تحسين أوضاع الطلاب ومستواهم التعليمي.
* تسهل عملية متابعتهم ومساءلتهم.
* زيادة نسبة احتفاظهم وتقليل نسب فشلهم.
* مساعدتهم على تخطي التحديات التي يواجهونها.

كل ما سبق يؤدي إلى دعم العملية التعليمية وإثرائها وتوجيهها نحو النجاح، وذلك باستفادة من تقنيات تحليل المعطيات والتنقيب بهدف الوصول إلى معرفة من شأنها المساعدة على كشف الخلل الحاصل في الأنظمة التعليمية للعمل على إصلاحه.

## دوافع المشروع:

تعمل معظم الأنظمة التي تسعى لتحسين أوضاع الطلاب في المؤسسات التعليمية على استخدام تقنيات معقدة كالتنقيب مثلاً وعدم الاكتراث لأهمية عرض النتائج بطريقة مبسطة يفهمها المستخدمون النهائيون كالمعلمين، مصممي المناهج التعليمية، الإداريين والطلاب دون الحاجة إلى معرفة تقنية مسبقة أو تدريب خاص على استخدام هذه الأنظمة، فعادة ما تتصف هذه الأنظمة بتعقيدها وصعوبة التعامل معها. إضافة إلى تصميم هذه الأنظمة بما يتوافق مع أنظمة تعليمية محددة وصعوبة تعميم التجربة باختلاف بنية الأنظمة التعليمية، وصعوبة مكاملة البيانات عند استخدام مؤسسة تعليمية تعتمد عدة مصادر وأنظمة مختلفة البنية. من هنا كانت الحاجة للعمل على نظام سهل الاستخدام، يوفر التوافقية مع أنظمة مختلفة وقادر على مكاملة البيانات مهما تعددت مصادرها للحصول على معرفة مُثلى تغطي كامل جوانب العملية التعليمية.

## المتطلبات البرمجية والعتادية للنظام:

To do

## حدود المشروع:

الحدود الوظيفية:

يشكل النظام خدمة تسعى لتحقيق أهداف عمليات التحليل والتنقيب في المعطيات التعليمية. يتكون النظام من مكونات منفصلة تحقق بتكاملها مع بعضها تجربة تعليمية أمثلية تساعد على تحسّن أوضاع الطلاب وتحقق الأهداف المذكورة سابقاً. يتمتع النظام بمرونة عالية نتيجة إمكانية تطوير النظام لدعمه خدمات أخرى، سهولة ربطه مع الأنظمة التعليمية المستخدمة، والتعامل مع بيانات ناتجة من أنظمة تعليمية مختلفة بشرط أن تكون مهيكلة وفق معيار الهيكلة الموحد.

### الحدود الزمانية:

إن المشروع الذي يتم تطويره له بداية ونهاية ومحدود بفترة عمل معينة كأي مشروع برمجي أو غير برمجي، ولما كان تنظيم العمل من أساسيات نجاح أي مشروع، تم وضع مخطط زمني Time Line لتطوير المشروع لتمكن أعضاء الفريق من تقييم عملهم وتقدمهم في الإنجاز، كما تم وضع مخطط غانت للمشروع Gant.

## الخرج المتوقع للنظام:

يقدم النظام خدمة تحليل للبيانات الناتجة عن العمليات التعليمية الافتراضية. يُتَوَقع أن يستفيد الطلاب من هذه الخدمات بمعرفة وضعهم وتوقع نتيجتهم في الاختبارات والحصول على تغذية راجعة لحالتهم بناء على نشاطاتهم خلال الفصل. بالإضافة لإمكانية الحصول على بعض الدراسات الإحصائية المبيّنة بطريقة رسومية مرئية سهلة وواضحة تتعلق بوضعهم كطلاب مع إمكانية مقارنة الطالب مع أقرانه.

كما يسهل النظام على المدرسين الحصول على معرفة متكاملة عن أوضاع الطلاب ويقدم دراسات إحصائية عن المحتوى ومدى تفاعل الطلاب معه مما يساعد المعلمين على اتخاذ قرارات متعلقة بتغيير الأسلوب المتبع في التدريس أو حتى العمل على تغيير المحتوى والاعتماد على وسيلة تعليمية دون أخرى تبعاً لمدى فعالية الوسيلة التعليمية، ويتم تقديم معلومات عن الطلاب ومستواهم وأوضاعهم.

ويمكن لمصممي المناهج التعليمية والإداريين الاستفادة من الخدمات التي يقدمها النظام والحصول على معرفة تشمل المدرسين والطلاب على سواء.

يسهل على المطورين لنظم التعلم الافتراضي ربط الأنظمة الموجودة والتي تعمل عليها المؤسسة مع النظام المقترح بناءه نظراً لاتباع سياسة محددة ومعيار موحد للبيانات. يقوم المطور بهيكلة البيانات وفق هذا النموذج وبعد ذلك تصبح الخدمات التي يقدمها النظام متوفرة للمؤسسة التعليمية.

## القيود المفروضة على المشروع:

To do

## خاتمة:

تناول هذا الفصل لمحة عن محتوى المشروع والأهداف العامة التي يعمل على تحقيقها، وذلك بعد تحديد المشكلة التي استدعت العمل على المشروع، ووضع تصور مبدأي للخدمات والمهام التي يعمل النظام على تحقيقها.

# الفصل الثاني: الدراسة النظرية.

## مقدمة:

يتناول هذا الفصل دراسة نظرية لبعض المجالات، المصطلحات والمعايير التي قد تساهم في بناء نظام يّقدّم خدمات تحليل للبيانات الناتجة عن الأنظمة التعليمية. سيتم التطرق إلى التنقيب في البيانات وكيفية توظيفها في المجال التعليمي، مع التعرّف على الميزات الأساسية لعمية التنقيب في المعطيات التعليمية وتحليل التعلم وأهدافهما والفروق الأساسية بينهما، كما سيتم ذكر بعض المعايير المستخدمة في هيكلة البيانات التعليمية.

## التنقيب في المعطيات:

### هدف عمليات التنقيب:

تّعنى التنقيب في البيانات بتحليل هذه البيانات من أُطر ونواحي وأبعاد مختلفة بهدف الوصول للأنماط المخبأة في مجموعة كبيرة ومعقدة من المعطيات Big Data باستخدام خوارزميات رياضية وحسابية، مما جعل البعض يطلق عليها اكتشاف المعرفة Knowledge discovery. تلعب عملية التنقيب هذه دوراً هام في عمليات دعم القرار وتسمح بالقيام باستعلامات دقيقة باستمرار دون الحاجة إلى وجود هدف محدد. ولا تقوم عملية التنقيب في المعطيات على التصفح العشوائي لقواعد المعطيات وإنما تكون عملية مدروسة بدقة تامة ومصممة لاستخلاص نتائج وقرارات مفيدة غنية بالمعلومات المهمة والمطلوبة، بالرغم من ذلك تتم عملية التنقيب دون فرضيات مسبقة. تقوم عملية التنقيب هذه على بيانات ضخمة متراكمة على مدى سنين لتحقيق نتائج أدق وأفضل. وتعمل أنظمة التنقيب على الحصول على معرفة ناتجة عن التنقيب في بيانات ناتجة عن أنظمة مختلفة بعد إجراء عمليات مختلفة عليها بغية مكاملتها والتأكد من صحتها، وهنا يجب الانتباه لأهمية إجراء عمليات التنقيب على بيانات صحيحة وخالية من الأخطاء قدر المستطاع، لأن إجراء التنقيب على بيانات خاطئة يؤدي للوصول لمعرفة خاطئة وغير دقيقة.

### لمحة تاريخية:

علم التنقيب في المعطيات Data Mining ليس بجديد، وخاصة نتيجة اعتماده بشكل أساسي على علم الإحصاء الذي تعود أصوله إلى منتصف القرن الثامن عشر. فالوضع الحالي الذي آل إليه التنقيب عن المعطيات جاء نتيجة جهود تراكمت من سنين عديدة، إلا أن هذا المجال شهد تطوراً ملحوظ في الآونة الأخيرة. يبين الشكل التالي مراحل تطور هذا المجال

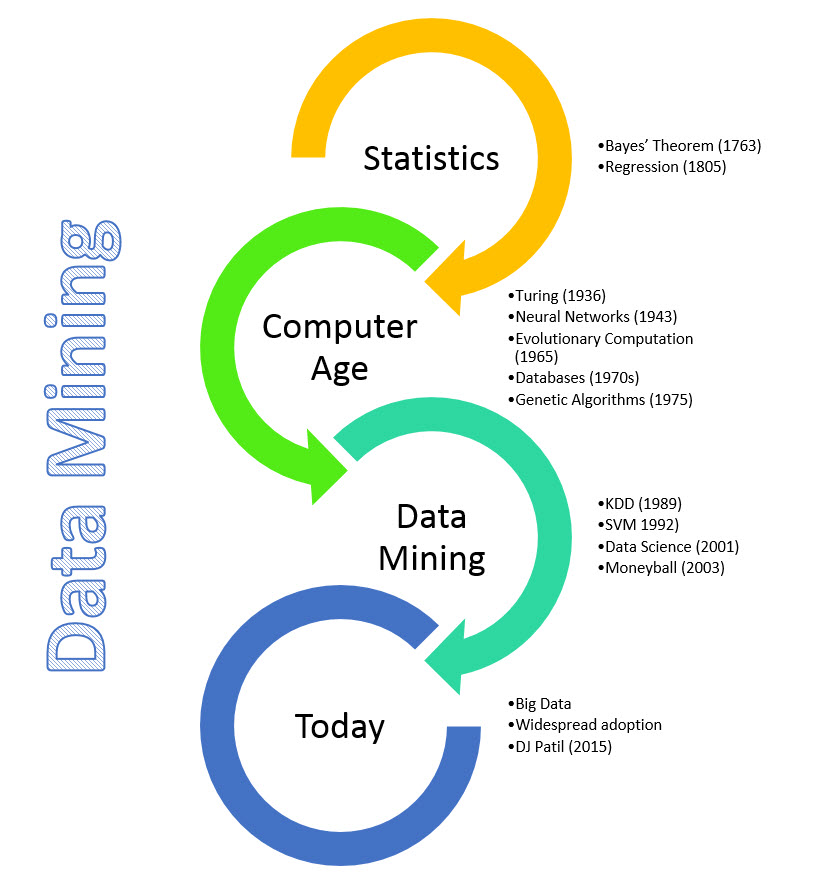


Figure : Data Mining Timeline

بالعودة إلى جذور التنقيب عن المعرفة في الأنظمة التعلمية نجد أن الأبحاث في هذا المجال جديدة نسبياً. فبالرغم من أن عمليات البحث والتحليل للبيانات الناتجة عن البرمجيات التعلمية بدأت من مدة طويلة، إلا أن التنقيب عن المعرفة في الأنظمة التعليمية Educational Data Mining (EDM) تم اعتباره كحقل مستقل بحد ذاته مؤخراً، من خلال المؤتمرات (Internal Conference on Educational Data Mining, started in 2008) والمجلات العلمية (Journal of Educational Data Mining “JEDM”, first issue published 2009). كما تم تقديمها وتعزيزها في مؤتمرات تكنولوجيا التعليم مثل:

* the International Conference on Educational Data Mining
* the Conference on Learning Analytics and Knowledge
* the International Conference on Artificial Intelligence in Education (AIED)
* the International Conference on Intelligent Tutoring Systems (ITS)
* the International Conference on User Modeling, Adaptation, and Personalization (UMAP)

### مجالات التنقيب في البيانات:

#### التنقيب في البيانات في المجال الطبي:

تحوي قواعد معطيات المشافي والمراكز الصحية كمية كبيرة من البيانات حول المرضى ووضعهم الصحي، وبدراسة العلاقات والأنماط بين هذه البيانات يمكن الوصول إلى معرفة طبية جديدة تفيد في اكتشاف مبكر للأمراض، والمساعدة بتشخيص الأمراض. كما أن عملية التنقيب هذه تعمل على ضمان تلقي المريض العلاج المناسب في الزمان والمكان المناسبين. من أحد أهم تطبيقات التنقيب في المجال الطبي الضمان الصحي واكتشاف حالات الاحتيال. يعتقد الباحثون أن التنقيب في المعطيات الطبية يمكن أن يحسّن المستوى الصحي العام بنسبة تصل إلى 30%.

#### التنقيب في البيانات في مجال القطاع المصرفي والمالي:

تستفيد المؤسسات المالية الضخمة كالبنوك من عمليات التنقيب في اكتشاف الربح والخسارة، واكتشاف الاحتيال، ودراسة الحملات التسويقية لاتخاذ القرار الأمثل اتجاه زبائنها للمحافظة عليهم وزيادة عددهم تبعاً لنوع الخدمات المصرفية المتنوعة التي تقدمها من قروض واستثمار وغير ذلك. تبرز أهمية التنقيب في المعطيات المالية نظراً لكمية البيانات الهائلة التي تُولّد بشكل سريع جداً مما يمنع الخبراء من المتابعة الدائمة والدقيقة لهذ المعطيات. أما عن مجال المبيعات فيعتمد التنقيب فيه على ربط شراء منتج من المنتجات بشراء منتج آخر. تسمح هذه العملية للتجار بفهم سلوك الزبائن ومعرفة احتياجاتهم للعمل على تلبيتها وزيادة الأرباح. كما تساعدهم في دراسة المنافسين لوضع خطط المبيعات الأمثل للتغلب على المنافسين.

#### التنقيب في البيانات في مجالات العلوم والهندسة:

تستخدم التنقيب في البيانات لتحليل الصور الفلكية والهندسة الكيميائية والمجال الحيوي. كما يمكن الاستفادة من التنقيب في مجال الأبحاث. إذ تقدم عمليات التنقيب للباحثين القدرة على إيجاد بيانات مماثلة والتي قد تسهل أو تغيير مجرى البحث.

#### التنقيب في البيانات في المؤسسات التعليمية:

يوجد عدة تعريفات للتنقيب في البيانات التعليمية من بينها تعريف ذُكر في مجلةEducational Data Mining و Baker الذي يعد من أوائل الباحثين في التنقيب بالمجال التعليمي بأنها تطوير الأساليب لاستكشاف الأنواع الفريدة من البيانات التي تأتي من البيئات التعليمية، واستخدام تلك الأساليب من أجل فهم أفضل للطلاب، وعكسها على البيئة التي يتعلمون بها.

وفي تعريف آخر: استخدام تقنيات التنقيب عن المعلومات لأنواع محددة من البيانات المستخرجة من البيئات التعليمية لمعالج وحل المسائل والمشاكل التعليمية الهامة.

كلا التعريفين يؤكدان على استخلاص المعرفة من بيانات تعليمية لتحسين الأنظمة التعليمية.

عادة ما يتم الخلط بين عملية التنقيب السابقة ‘EDM’ وتحليل التعلم ‘learning analytics’ والذي يعرّف بجمع وتحليل البيانات وإعداد التقارير حول المتعلمين بغرض فهم وتحسين العملية التعليمية والبيئة التي تحضنها. تُعنى العملية السابقة بالإجابة على أسئلة *محددة*. مع عدم وجود جدار فصل حقيقي بين المجالين ووجود العديد من الأهداف المشتركة بينهما تزداد صعوبة الفصل وتميز كون تطبيق يتبع لأحد المجالين أو للآخر.

يمكن اعتبار التنقيب عن المعرفة في المجال التعليمي مؤلّف من ثلاث مكونات رئيسية: علوم الحاسب computer science، التعليم education، والإحصائيات statistics. تقاطع المجالات الثلاثة السابقة يشكل مجالات أخرى فرعية مثل تحليل التعليم Learning Analytics، التعليم الإلكتروني E-Learning، والنقيب Datamining

يُعد التنقيب في البيانات التعليمية مجال متعدد الاختصاصات، إذ يستخدم طرق وتقنيات مختلفة مثل الإحصاء، التنقيب في المعطيات، نظم استرجاع المعلومات، نظم الاقتراح، تعلّم الآلة، علوم التربية وأصول التدريس وغيرها. إلا أن عملية اختيار التقنيات المستخدمة تعتمد بشكل أساسي على القضية التعليمية المطروحة.

يمكن تمثيل المراحل التي تمر فيها عملية التنقيب في الأنظمة التعليمية بالشكل التالي:

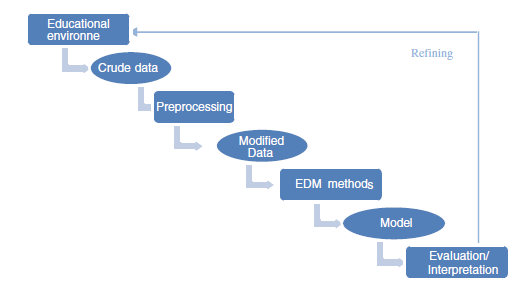


Figure : مراحل التنقيب في المعطيات التعليمية

على الرغم من أن الأساليب المستخدمة للتنقيب في المحتوى الناتج عن الأنظمة التعليمية يتشابه مع مجالات أخرى كالتجارة الإلكترونية إلا أن الأهداف تكون مختلفة. فمن وجهة نظر النظام، لا توجد فروق لأن الهدف من التنقيب على شبكة الإنترنت في كلا المجالين هو دراسة سلوك المستخدمين (العملاء في مجال التجارة الإلكترونية، والطلاب في أنظمة التعلم الإلكتروني)، تقييم هذا السلوك وتحسين الأنظمة لمساعدة المستخدمين. إلا أنه من وجهة المستخدم هناك اختلافات، إذ أن التجارة الإلكترونية تهدف إلى توجيه العملاء في عملية الشراء في حين أن هدف التعليم الإلكتروني هو توجيه الطلاب في التعلم. يمكن تلخيص ما سبق بأن الفروق الرئيسية بين التنقيب في البيانات التعليمية والمجالات الأخرى هو: نوعية البيانات، الهدف والتقنيات نظراً لخصوصية بنية الأنظمة التعليمية.

-هناك مجالات أخرى يتم تطبيق التنقيب عن المعطيات فيها مثل التحقيق بالجرائم والهندسة الصناعية واكتشاف الاحتيال وإدارة العلاقة بالعملاء وغيرها.

### أهداف التنقيب عن المعطيات في الأنظمة التعليمية:

تم تصنيف أهداف التنقيب في الأنظمة التعليمية بناء على أنواع المستخدمين النهائيين المستفيدين من المعرفة الناتجة عن التنقيب إلى:

* متعلمين: تقديم الملاحظات والنصائح للطلاب، الاستجابة لاحتياجات الطالب، العمل على رفع إنتاجية الطالب وسويته العملية.. إلخ
* معلمين: مساعدتهم على الفهم الأمثل لطبيعة تلاميذهم وكيفية سير العملية التعليمية ومستوى الطلاب وعكس ذلك على الطرق التي يتبعونها في التعليم وتحسين الأداء التدريسي.. إلخ
* باحثين: تطوير ومقارنة تقنيات استخراج البيانات لجعلها قادرة على تقديم الاقتراحات الأمثل لكل مشكلة أو مهمة تربوية محددة، تقييم فاعلية التعلم عند استخدام طرق وبيئات وأساليب مختلفة.. الخ.
* إداريين: إيجاد الطريقة الأفضل لتنظيم المصادر (البشرية والمادية) وتقييم مؤسساتهم التعليمية.

-إن عملية إشراك جميع أصحاب المصلحة وتقديم نتائج التحليل إليهم هو مهمة معقدة لابد من حلها، لأن بعض هذه المصالح قد تكون متناقضة. على سبيل المثال، استخدام LA من قبل موظفي الإدارة لإيجاد أفضل الطرق لدمج التكنولوجيا بطرق التدريس يمكن يشعر المدرسين بأنه يتم تقييمهم والتحكم بأساليبهم المتبعة في العملية التدريسية. كذلك الأمر بالنسبة للطلاب الذين قد يخشون من استخدام البيانات الشخصية لتقييمهم وتصنيفهم. وهذا يمكن أن يؤدي إلى أثر غير مقصود ألا وهو عدم ترحيب المدرسين والطلاب باستخدام هذه التكنولوجيات الجديدة والمشاركة في عمليات التحليل.

-في وجه النظر السابقة تم تصنيف الأهداف تبعاً لأنواع المستخدمين النهائيين، ولكنه من الصعب تصنيف جميع أهداف التنقيب عن المعرفة في الأنظمة التعلمية تبعاً لأنواع المستخدمين، وخاصة عند وجود أهداف ترتبط بأكثر من مستخدم، لذلك وجدت أبحاث أخرى قامت بتصنيف أهداف تنقيب البيانات التعليمية كالتالي:

* نمذجة الطالب: نمذجة المستخدم في النظام التعليمي تتطلب إعطاء معلومات مفصلة عن خصائص الطلاب وحالاتهم، كالمعرفة التي يمتلكها الطالب والمهارات والخبرات والدوافع للتعلم وصولاً إلى دراسة بعض أنواع المشاكل التي قد تؤثر سلباً على مخرجات العملية التعليمية كاستخدام مصادر تعليمية غير فعالة. الهدف هنا هو خلق أو تحسين نموذج الطالب انطلاقاً من المعلومات المتاحة.
* توقع المستوى التعليمي للطلاب ونتائجهم: الهدف هو توقع النتيجة النهائية، وتوقع نتيجة أي عملية تعليمية انطلاقاً من بيانات ناتجة عن تفاعل الطالب مع المادة الدراسية.
* إيجاد الاقتراحات للطلاب: فالهدف هو اقتراح المحتوى أو حتى المهام الأمثل لوضع الطالب في الوقت الحالي.
* تحليل سلوك المتعلم: وهذا الهدف يأخذ أشكال عدّة، بدء من الإجابة عن الأسئلة الثلاثة سابقة الذكر، وصولاً لتصنيف الطلاب تبعاً لعوامل عدة كملفهم الشخصي.
* التواصل مع الـ stockholders: الهدف هو مساعدة المعلمين ومديري المواد للتحليل أنشطة الطلاب.
* تحليل هيكلية المجال: والهدف هو تحديد بنية المجال لمعرفة المحتوى الأمثل الذي يمكن إضافته للمواد والتسلسل الأفضل للعملية التعليمية، وذلك باستخدام القدرة على التنبؤ بأداء الطلاب وتحديد المحتوى والبيئة التعليمية المثلى لهم.
* تحسين المحتوى التعليمي: وهذا ذو ارتباط وثيق بالهدفين السابقين. فالهدف هنا هو تحديد كيفية تحسين المواد التعليمية (محتويات، أنشطة وما إلى ذلك) باستخدام معلومات حول الطالب.
* دراسة تأثيرات الدعم التربوي pedagogical support الذي يتم تقديمه من قبل أنظمة التعليم.

من الجدير بالذكر أن أهداف تنقيب البيانات التعليمية سابقة الذكر تهدف إلى تحسين النظم التعليمية بشكل عام والنظم التعليمية المعتمدة على الحاسوب بشكل خاص، يمكن القول بأن نمذجة الطالب هو المفتاح الأساسي لتحقيق العديد من أهداف ومهام تنقيب البيانات التعليمية.

### الطرق المستخدمة في تنقيب البيانات التعليمية:

لتحقيق أهداف تنقيب البيانات التعليمية تم استخدام معظم طرائق تنقيب البيانات التقليدية (clustering, classification, association analysis..) والتي من الممكن أن تصنف كالتالي:

* **التوقع Prediction:** يكون هدف هذه الطريقة هو تطوير نموذج يؤدي إلى استدلال aspect معين من هذه البيانات (يطلق عليه Predicate Variable) من مجموعة من الـ aspects من هذه البيانات (يطلق عليها Predictor Variables). من بعض طرق التوقع هذهClassification, Regression Density Estimation,. مثال على تطبيقات تنقيب البيانات التعليمية باستخدام هذه الطريقة هي توقع نجاح الطالب أكاديمياً وتوقع سلوكه.
* **العنقدة Clustering**: تعني باختصار تصنيف العينات التي تحويها البيانات والتي تتشابه فيما بينها بشكل طبيعي (عمليّاً يتم استخدام بعض طرائق قياس البعد بين بعضها البعض لتحديد مقدار التشابه) وعند تحديد هذه التصنيفات من الممكن معرفة انتماء عينة جديدة ما وفقاً لهذه التصنيفات الأخيرة. تستخدم هذه الطريقة في تنقيب البيانات لتجميع الطلاب وفقاً لنمط تعلمهم أو استراتيجية cognitive.
* **تنقيب العلاقات Relationship Mining:** تستخدم لاكتشاف علاقات بين المتغيرات في عينة من البيانات، يوجد عدّة أنواع من تنقيب العلاقات مثل association rule mining, sequential pattern mining, correlation mining casual data mining,. في تنقيب البيانات التعليمية يستخدم تنقيب العلاقات لتحديد العلاقات بين نشاطات الطلاب على الشبكة وعلاماتهم مثلاً.
* **البيانات للتحكيم البشري Distillation:** تتضمن هذه الطريقة depicting للبيانات بشكل يسمح للمستخدم بالتعرف أو تصنيف ميزات البيانات بشكل سريع وسهل الاستخدام. تعتمد هذه التقنية على تلخيص البيانات Summarization وتصويرها Visualization واستخدام الواجهات التفاعلية Interactive Interfaces لتوضيح المعلومات المفيدة ودعم اتخاذ القرار.
* **الاستكشاف مع النماذج Discovery With Models:** تقوم هذه الطريقة باستخدام نموذج معرّف ومحقق مسبقاً.
* **اكتشاف العينات الشاذة Outliers Detection**: تعتمد هذه الطريقة على اكتشاف عينات معينة من البيانات ذات اختلاف ملحوظ عن باقي العينات. تستخدم هذه الطريقة في تنقيب البيانات التعليمية لكشف الاختلافات بأفعال المتعلم أو المدرس وسلوكه، عمليات التعليم غير النظامية أو تحديد المتعلمين الذين يعانون من صعوبات في عملية التعلم.
* **تحليل الشبكات الاجتماعية Social Network Analysis:** يهدف التحليل الهيكلي SNA إلى دراسة العلاقة بين الأفراد بدل من دراس سلوك وخصائص الأفراد أنفسهم، إذ يتم تمثيل العلاقات الاجتماعية بعقد تمثل الأفراد وروابط بين هذه العقد تمثل العلاقات بين هؤلاء الأفراد من صداقات وعلاقات وتعاونية وغيرها. يمكن استخدام ذلك في تنقيب البيانات التعليمية لتفسير وتحليل العلاقات في المهام التعاونية والتفاعل مع أدوات التواصل.
* **تنقيب العملية Process Mining**: تهدف إلى استخلاص العمليات المتعلقة بالمعرفة، وذلك باستخدام سجلات الأحداث التي سُجلت من قبل نظام المعلومات لتكوين تمثيل مرئي واضح لكامل العملية. تتكون من ثلاثة حقول فرعية: التوافق conformance، التحقق checking، واكتشاف النموذج model discovery. في نقيب البيانات التعلمية يمكن استخدام تنقيب العملية لعكس سلوك الطلاب من حيث تأثير دراستهم (المكونة من تسلسل صفوف، علامات... إلخ) لكل طالب.
* **تنقيب في النصوص Text Mining:** يركّز على إيجاد واستخراج أنماط مفيدة، نماذج، وقواعد من النصوص غير المهيكلة مثل ملفات HTML، رسائل الدردشة ورسائل البريد الإلكتروني. تشمل المهام تصنيف النص، تجميع النص، تحليل للمشاعر المضمنة في النص، تلخيص للوثائق، نمذجة علاقات الكائنات وغيرها. إذ يستخدم التنقيب في النصوص لتحليل مضمون المناقشات، المنتديات، الدردشات، صفحات الويب، الوثائق.. الخ

### تصنيف البيانات المحللة:

يوجد العديد من أنواع البيانات التي يتم استخدامها في عملية تنقيب البيانات التعليمية وذلك حسب الهدف من عملية التنقيب والتقنيات المستخدمة ويمكن تصنيف هذه البيانات وفق الخواص التالية:

* وفرة البيانات Data Availability:

مثل البيانات المسجلة على مدار السنين في قواعد معطيات المؤسسات التعليمية أو من سجلات النظم البرمجية المستخدمة.

* مصادر التجميع Collection Resources:

أي كون تجميع هذا البيانات قد تم بشكل يدوي من قبل موظفين ومراقبين في المؤسسة التعليمية أو أنها بيانات رقمية قد تم تجميعها على شكل أرقام (علامات) أو سجلات في قواعد معطيات أو ملفات صوتية أو أشرطة فيديو

* بيئة التعلم Learning Environment:

إذ يتم تجميع البيانات الناتجة عن عملية التعلم في البيئة الافتراضية Logging من الملاحة في النظام وعدد نقرات.. إلخ، والاستفادة من ذلك في الوصول إلى معرفة تساعد في اتخاذ قرار يصب في مصلحة الطالب.

### بعض الأدوات المستخدمة في عملية التحليل:

من أهم التطوّرات التي شهدها هذا المجال هو إيجاد طرق لإدارة الكميات الضخمة من البيانات مما يسهل إجراء عمليات التحليل عليها. فقط طورت غوغل MapReduce للحد من التحديات التي تظهر نتيجة إدارة البيانات على نطاق الإنترنت، بما في ذلك توزيع البيانات والتطبيقات التي تمتلك بيانات مرتبطة بشبكات الحواسيب. فقواعد البيانات السابقة كانت غير قادرة على إدارة البيانات على نطاق شبكة الإنترنت، مما أدى إلى تطوير Apache Hadoop الذي يستخدم الآن بشكل شائع في إدارة البيانات.

ظهر مؤخراً أدى جديدة تدعى Apache Spark الذي يتمتع بميزات عديدة تجعله يتفوق على الأدوات الأخرى كمعالجة البيانات المتدفقة real-time processing

بالإضافة للأدوات المستخدمة في إدارة البيانات ظهرت العديد من الأدوات التي تدعم عمليات التحليل وتؤمن عمليات التنقيب عن المعرفة دون الحاجة لمعرفة عميقة في علوم الحاسب، معظم هذه الأدوات مطور بما يتوافق مع مجال ذكاء الأعمال والذي يتجلى بظهور SAS وأدوات IBM في المجال التعليمي، والتي استخدمت في التحليلات التنبؤية وتحسين عملية صنع القرار من خلال تحليل كميات كبيرة من البيانات وعرضها بشكل تفاعلي. فيما سبق كان استخدام الأدوات التحليلية عملية معقدة تتطلب فهم عميق في التحليل والبرمجة، أما الآن فحتى الأدوات المعقدة كـ SAS, RapidMiner, DBMiner, DB2 Intellegent Miner, SPSS تتطلب معرفة تقنية أقل إلا أن معظمها تجاري(غير مجاني). وهنا لا بد من الإشارة أن كثير من برامج الحاسوب مثل Microsoft Excel شهد تطور ملحوظ في التحليل والإظهار. أدوات أخرى مثل Tableau Software تم تصميميها لدعم استخدام أدوات التحليل دون الحاجة إلى معرفة تقنية عالية، مما أدى إلى انتشار عمليات الـ EDM, LA بين عدد كبير من الباحثين في مجال علوم التعليم.

يوجد برمجيات أخرى مفتوحة المصدر تؤمن عمليات التنقيب مثل Weka, Keel. كما يوجد بعض البرمجيات المختصة بالتنقيب عن المعطيات التعليمية مثل: MultiStar الذي يعمل على التصنيف ودراسة علاقات الارتباط، EPRules لدراسة علاقات الارتباط، KAON للعنقدة والتنقيب في النصوص، Synergo/ColAT للحصول على الإحصائيات وعمليات عرض النتائج، GISMO للعرض، Listen tool للعرض، TADA-Ed للتنقيب والعرض، Sequential Mining tool لتنقيب الأنماط، MINEL لتنقيب طرق التعلم، CIECoF لتنقيب علاقات الارتباط، Simulog لكشف أنماط السلوك غير المتوقع.

## تحليل التعلّم Learning Analytics:

يوجد عدة تعريفات لتحليل التعلم، أبرزها ما تم توضيحه في المؤتمر الأول لتحليل التعلم والمعرفة LAK11 والذي ينص على تعريف تحليل التعلم على أنه قياس وتحليل واعداد التقارير في البيانات الناتجة عن العملية التعليمية والقائمين عليها والمشاركين بها.

### الفرق بين التنقيب في المعطيات الناتجة عن الأنظمة التعليمية وتحليل التعلم:

بالرغم من الأهداف الموحدة للمجالين السابقين، إلا أنه يمكن التمييز بينهما وفق النقاط التالية:

* يهتم التنقيب في البيانات التعليمية بتحويل عملية اكتشاف المعرفة إلى عملية آلية، بينما يكون الاعتماد الاساسي في عملية التحليل على العنصر البشري.
* تعمل أنظمة التحليل على تقسيم الأنظمة إلى واحدات مستقلة واكتشاف هيكلية كل واحدة والعلاقة بين العناصر السابقة بينما يقوم التحليل على فهم النظام كوحدة متكاملة.
* يعتمد التنقيب على تقنيات التصنيف والتجميع والنماذج البايزية وتنقيب العلاقات، في حين بركز التحليل على تحلي الشبكات الاجتماعية والمشاعر والخطاب والتنبؤ بنجاح المتعلم.

يوضح الشكل التالي تداخل التنقيب في المعطيات الناتجة عن الانظمة التعليمية وتحليل التعلم وغيره من المجالات ذات الصلة بعلوم الحاسب والتحليل والتعليم:

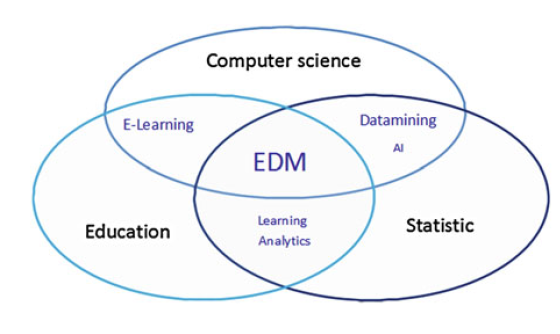


Figure : تداخل EDM مع LA

### النموذج المرجعي في تحليل التعلم:

قبل البدء بعملية تحليل البيانات في الأنظمة التعليمية يجب الإجابة عن أربعة أسئلة لتبيان هيكلية عملية التحليل كاملة.

* ماهي طبيعة البيانات المراد تحليلها والبيئة المولدة لها.
* ما هي الأهداف المرجوة من عملية التحليل.
* ما هي الطرق المتبعة في عملية التحليل.
* من هم المستفيدون النهائيون من عملية التحليل.
* هم المستفيدون النهائيون من عملية التحليل.

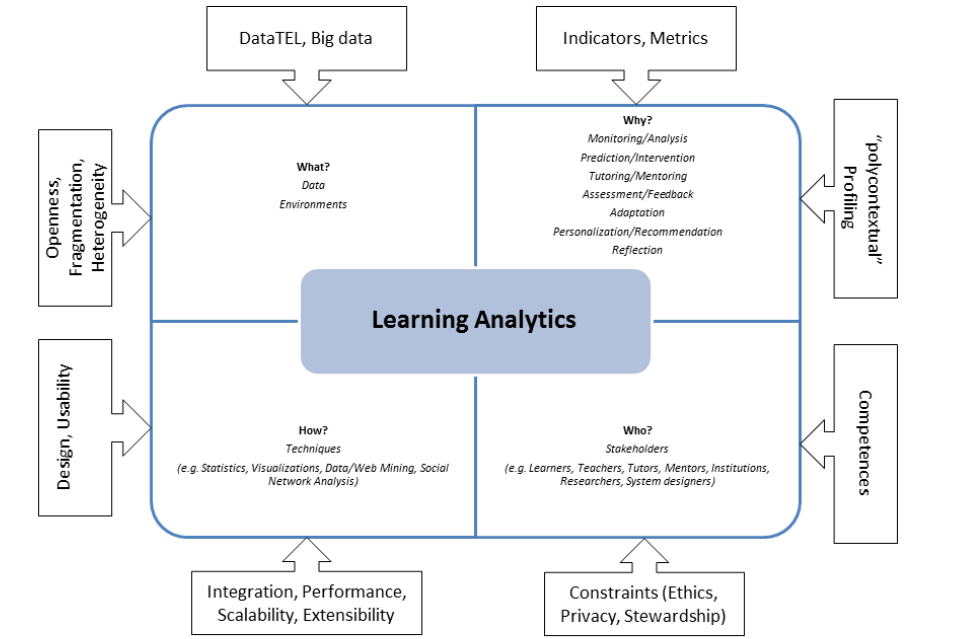


Figure : النموذج المرجعي في تحليل التعلم

نورد فيما يلي دراسة للجوانب الأربعة السابقة:

#### أنواع البيئات الافتراضية التعليمية:

فيما يلي ذكر لأهم البيئات الافتراضية التعليمية:

##### Massive Open Online Courses (MOOC):

إن الدورات الافتراضية على شبكة الإنترنت في تزايد مستمر، كما أن الاهتمام بها من قبل المجتمع العلمي يزداد أيضاً. يمكن تعريف الـ MOOC بأنها دورات على شبكة الإنترنت تهدف للتفاعل الكبير والواسع بين مرتادي هذه الدورات. إن إتاحة هذه الدورات على شبكة الإنترنت يجعل من الممكن وصول أي فرد يملك اتصال بشبكة الإنترنت التسجيل بهذه الدورات مجاناً. بعض الأمثلة عنها Udacity, MITx, EdX, Coursera, Udemy. إن هذه المنصات تشابه لحد كبير استخدام نظم إدارة التعلم LMS ولكن مع زيادة عدد المستخدمين بشكل كبير، وبذلك تكون كمية البيانات الناتجة عن هذه الأنظمة ضخمة جداً مما يجعل الحاجة لاستخدام التنقيب في البيانات والتحليل ماسة لمعالجة هذه البيانات.

##### Intelligent Tutoring Systems:

أنظمة الإرشاد الذكية هي أنظمة تقدم تعليمات وملاحظات **مباشرة** للطلاب (عادة ما تكون دون تدخل بشري من المعلم). ويهدف البرنامج إلى توفير فرص الحصول على تعليم عالي الجودة إلى كل طالب، وبالتالي إصلاح النظام التعليمي برمته. تتغير نماذج التفاعل بناءً على النماذج الفردية. يؤدي قدرة هذه الأنظمة على تسجيل وتجميع تفاصيل التفاعلات مع أعداد كبيرة من الطلاب إلى خلق مجموعة ضخمة من البيانات. على الرغم من أن هذه الأنظمة تسجل كل تفاعلات الطلاب في ملفات أو قواعد بيانات، هناك مخازن بيانات أخرى متاحة في هذه الأنظمة، على سبيل المثال، نموذج المجال الذي يشتمل على مجموعة من القيود ذات الصلة بنطاق المدرس، ومجموعة البيانات التربوية التي تحتوي على مجموعة من المشاكل وإجاباتها، ونموذج الطالب الذي يقوم بتخزين المعلومات حول كل طالب مع مراعات جميع القيود. هناك أبحاث جارية لتحسين فعالية هذه الأنظمة.

##### Adaptive and Intelligent Hypermedia Systems:

أنظمة الوسائط الفائقة المتكيفة والذكية توفر بديل لمنهج just-put-it-on-the-Web التقليدي في تطوير المناهج التعليمية التربوية. تحاول هذه الأنظمة أن تكون أكثر قابلية للتكيف من خلال بناء نموذج للأهداف، والأفضليات، ومعرفة كل طالب على حدة واستخدام هذا النموذج في التفاعل مع الطلاب من أجل التكيف مع احتياجات هؤلاء الطلاب. تكون البيانات القادمة من هذه الأنظمة غنية ويمكن أن تؤدي إلى تحليل أفضل من التحليل القائم على البيانات الناتجة عن أنظمة التعليم المعتمدة على الويب التقليدية. إن البيانات المتاحة في هذه الأنظمة مشابهة لبيانات أنظمة الإرشاد الذكية ITS.

##### Test and Quiz Systems:

أنظمة الاختبار هي من أكثر الأدوات انتشار واستخدام في مجال التعليم كما أنها يتم تطويرها بشكل فعال وجيد. يتكون الاختبار من سلسلة من الأسئلة بغرض جمع المعلومات من الممتَحنين. الهدف الرئيسي من هذه الأنظمة هو قياس المستوى المعرفي للطلاب بما له علاقة بموضوع أو أكثر. هناك أنواع مختلفة من الأسئلة مثل الإجابة بـ نعم/لا، أسئلة الاختيار من متعدد MCQ، ملأ الفراغات، الأسئلة الكتابية وغيرها. وتخزّن نظم الاختبار كمية كبيرة من البيانات، مثل الأسئلة، أجوبة الطلاب، نتائج الطلاب، وبعض الإحصائيات.

##### Learning Management Systems:

هي نوع خاص من أنظمة التعلم عبر الإنترنت تتيح إدارة، توثيق وتتبع البرامج التعليمية، الفصول الدراسية، الأحداث عبر الإنترنت والمحتوى التعليمي. كما أنها توفر مجموعة كبيرة ومتنوعة من القنوات ومساحات العمل لتسهيل تبادل المعلومات والتواصل بين جميع المشاركين في العملية التعليمية.

بعض الأمثلة عن نظم إدارة التعلم التجارية: Blackboard و Virtual-U في حين أن هناك أنظمة أخرى مجانية مثل Moodle, Ilias, Sakai and Claroline.

تُراكم هذه الأنظمة بيانات هائلة ناتجة عن تسجيل كافة أنشطة وعمليات تفاعل الطلاب مع الأنظمة التعليمية، وعادة ما يوجد أدوات تتبع مضمّنة في هذه الأنظمة تسمح للمدرس بعرض بيانات إحصائية للبيانات. يمكن أن يتم تسجيل أي نشاط يتم في النظام مثل القراءة، الكتابة، الاختبارات، استعرض محتوى، التعليق على الأحداث وغيرها. كما أن أنظمة التعلم هذه توفر قاعدة بيانات علائقية تخرن فيها معلومات الطالب في جداول مختلفة كمعلومات المستخدم الشخصية (الملف الشخصي)، النتائج الأكاديمية (الدرجات)، وبيانات تفاعل المستخدم (التقارير).

نتيجة لاحتضان هذه الأنظمة بيانات ضخمة التي يتم توليدها يومياً من بيانات شخصية عن الطلاب ونتائج الطلاب الأكاديمية وبيانات التفاعل، يصعب إدارة هذه البيانات بشكل يدوي. وبالرغم من أن بعض أنظمة إدارة المحتوى توفر أدوات إعداد تقارير إلا أنها لا تسمح للمعلمين بتتبع وتقييم بدقة جميع الأنشطة التي يقوم بها المتعلمون وتقييم هيكلية ومحتوى وفعالية المادة العلمية، كما أنه يصعب على المعلمين استخلاص معلومات مفيدة من هذه الأدوات مع وجود عدد كبير من الطلاب، وهنا تبرز أهمية التنقيب.

* **Moodle**: Modular Object Oriented Developmental Learning Environment

نظام إدارة تعلم مفتوح المصدر، يساعد المدرسين على خلق بيئة تعليمية تفاعلية فعالة على شبكة الإنترنت. يدعم حوالي 75 لغة فيما يزيد عن 160 دولة. ويعتمد في تصميمه على نمط تعلّم يدعى social constructionist pedagogy الذي يقوم على أن أفضل العمليات التعليمية تكون عند التفاعل مع: المحتوى التعليمي والأقران حول المحتوى التعليمي.

يحوي مودل على:

-خمس أصناف للأدوات التعليمية: الصفحات النصية، صفحات الويب، رابط لشيء آخر على الويب، view into one of the course’s directories وعرض لنص أو صورة.

-ست طرق للتفاعل مع المقرر: الوظائف، الخيارات، الصحف، الدروس، الاختبارات والدراسات الاستقصائية.

-خمس طرق للتفاعل مع الآخرين: الدردشة، المنتدى، glossary، الويكي وورشات العمل.

يخزن مودل كل تحركات الطلاب وكل النقرات التي يقومون بها في النظام، ويمكن فلترة هذه المعلومات على حسب المقرر، النشاط، الطالب، واليوم. كما يمكن للمعلمين استخدام هذه البيانات لمعرفة الطلاب الأكثر نشاط في المقرر. لا يتم تخرين البيانات السابقة على شكل نصوص وإنما يتم تخزينها في قاعدة معطيات علائقية. أفضل النظم إدارة قواعد المعطيات الداعمة لمودل MySQL, PostgreSQL إلا أنه يمكن استخدام نظم أخرى مثل Oracle, Access, Interbase. يحوي مودل حوالي 145 جدول مترابط، إلا أن عملية التنقيب لا تحتاج إلى كل هذه البيانات وهنا يبرز دور مرحلة ما قبل التنقيب. نتيجة لبنية مودل تكون كمية الجهد المبذول في مرحلة ما قبل المعالجة أقل من الأنظمة التعليمية الأخرى. إلا أن بعض العمليات لا تزال مطلوبة مثل اختيار البيانات التي سيتم عملية التنقيب عليها وتحويل البيانات الرقمية إلى وصفية ونقل وتحويل البيانات للصيغة الأمثل لعميات التنقيب كصيغة ARFF(Attribute-Relation File Format) مثلاً.

تطبيق تقنيات التنقيب على بيانات Moodle:

* الإحصائيات: لا يؤمن مودل إحصائيات تزود المدرسين بتقارير محددة عن تفاصيل إحصائية لأداء كل طالب. وإنما يزود بإحصائيات لها علاقة فقط بأداء الطلاب (كالعلامات والاختبارات)، فيمكن للمدرسة معرفة أداء الطلاب في كل سؤال في الاختبار، مما يساعد المعلمين على تحديد الأسئلة الواجب تعديلها أو حذفها نتيجة لصعوبتها أو سهولتها الفائقة.
* العرض: يتعلق برسومات الحاسوب وواجهات المستخدم ويهتم بطرق عرض البيانات بشكل رسومي أو متحرك ليسهل على المستخدم فهمها. تبرز فعالية هذه الطريقة عند التعامل مع كمية كبيرة من البيانات ولتمثيل البيانات متعددة الأبعاد. لا يؤمن مودل طرق مرئية لعرض هذه البيانات غير النصوص، إلا أنه يمكن استخدام بعض الأدوات مثل GISMO الذي يمثل البيانات الناتجة عن تتبع الطلاب في مودل بطرق متعددة مرئية ومفهومة.
* العنقدة: تتصف بأنها [*unsupervised learning*](http://en.wikipedia.org/wiki/Unsupervised_learning)، تعمل العنقدة على كشف العلاقات بين العناصر بهدف تجميع العناصر المتشابهة. المبدأ المتبع في العنقدة هو زيادة التشابه في العنقود الواحد وتقليص التشابه بين العناقيد المختلفة. في التعلم الافتراضي يمكن الاستفادة من العنقدة في إيجاد الطلاب ذوي الخصائص التعليمية المشتركة وتشجيع التعلم التعاوني. يمكن بعد تجميع الطلاب في مجموعات تحديد مرشد خاص بهم بناء على الخصائص التي يتمتع بها طلاب المجموعة ومهاراتهم. كما يمكن توجيه اختبارات خاصة بهم. تؤمن الأداة Weka عدة خوارزميات عنقدة. يمكن على سبيل المثال تجميع الطلاب بناء على نشاطاتهم في مودل ونتائج الاختبارات النهائية.
* التصنيف: يتصف بأنه[supervised learning](http://en.wikipedia.org/wiki/Supervised_learning)، يعتمد على معرفة انتماء عناصر جديدة لفئات موجودة ومحددة مسبقاً. يمكن الاستفادة من التصنيف في التعلم الافتراضي في كشف الطلاب ذوي الخصائص المتشابهة وردّات الفعل المتشابهة على الاستراتيجية التربوية المتبعة، وللتنبؤ بأداء الطلاب ونتيجتهم النهائية ولتحديد المتعلمين الذين لا يملكون الدوافع والمحفزات اتخاذ الإجراءات لتخفيض معدلات التسرب. تؤمن الأداة Keel عدة خوارزميات تصنيف. يمكن للمعلم استخدام المعرفة الناتجة عن التصنيف باتخاذ قرارات لها علاقة بالنشاطات المتعلقة بالمقرر ضمن مودل.
* علاقات الارتباط: تعتمد علاقات الارتباط على مدى ارتباط حدوث العنصر x بعنصر آخر y. يمكن الاستفادة من علاقات الارتباط بتحديد المواد التعليمية الأنسب لاقتراحها على المستخدم، ولاكتشاف العلاقات المثيرة للاهتمام من استخدام الطالب للمعلومات بهدف توفير التغذية الراجعة لمدرس المقرر، ولكشف العلاقات بين سلوكيات التعلم. كما يمكن استخدامها لكشف الأخطاء التي يشترك الطلاب بارتكابها. تؤمن الأداة Weka عدة خوارزميات لكشف علاقات الارتباط.
* تنقيب الأنماط التتابعية
* تنقيب النصوص

##### أنواع أخرى للأنظمة التعليمية:

إضافة لما سبق يوجد أنواع أخرى من الأنظمة التعليمية مثل بيئات الألعاب التعليمية، بيئات الواقع الافتراضي، مستودعات كائنات التعلم، الويكي، المنتديات، المدونات وغيرها.

وهنا يجب الإشارة إلى أن توجه أساليب ومصادر التعلم نحو الحوسبة السحابية، فالتحدي الأكبر هو كيفية مكاملة البيانات القادمة من مصادر مختلفة غير متجانسة وممثلة بصيغ مختلفة في بعض الأحيان لخلق بيانات تعليمية مفيدة تعكس الأنشطة الموزعة للمتعلم. بالإضافة لكمية البيانات الضخمة الواجب معالجتها والتي تشكل تحدي جديد إذ يجب تأمين نتيجة المعالجة في الوقت الأنسب للمستخدم دون تأخير.

يوضح الشكل التالي نسبة الأبحاث في مجال التحليل تبعاً لمصدر هذه البيانات.

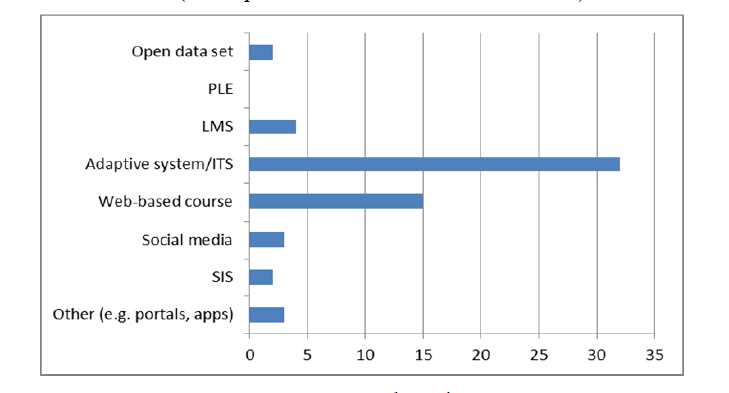


Figure : البيانات ومصادرها

#### أهداف عملية التحليل:

يمكن تحقيق [الأهداف المذكورة في الفصل الثاني](#_هدف_المشروع_العام:) من هذه الوثيقة بسبع طرق:

##### Monitoring and analysis:

يكون الهدف هو تتبع الأنشطة الطلابية وتوليد التقارير بهدف دعم عملية صنع القرار من قبل المعلم أو المؤسسة التعليمية. يرتبط الرصد أيضا بالتصميم التعليمي، ويشير إلى تقييم العملية التعليمية من قبل المعلم بهدف التحسين المستمر لبيئة التعلم. يمكن دراسة كيفية استخدام الطلاب لنظام التعلم وتحليل الإنجازات الطلابية تساعد المعلمين بكشف الأنماط واتخاذ قرارات بشأن تصميم نشاط التعلم في المستقبل.

##### Prediction and intervention:

الهدف في الأنظمة التي تقوم على التنبؤ هو تطوير نموذج يحاول التنبؤ بمعرفة المتعلم أداؤه في المستقبل، على أساس أنشطته وإنجازاته الحالية. يمكن فيما بعد استخدام هذا النموذج التنبؤي لتوفير التدخل الاستباقي الطلاب الذين قد يحتاجون إلى مساعدة إضافية. يمكن للتنبؤ بأداء الطلاب أن تدعم المعلم أو المؤسسة التعليمية في التدخل باقتراح الإجراءات التي ينبغي اتخاذها لمساعدة المتعلمين على تحسين أدائهم.

##### Tutoring and Mentoring:

تهتم بمساعدة الطلاب على التعلم، عادة ما تكون المساعدة ضمن نطاق معين ومحدود بسياق المقرر. فعلى سبيل المثال، يدعم المدرس المتعلمين في ميولهم ويقدم لهم وحدات تعلم جديدة، كما يعطيهم التعليمات المتعلقة بالمقرر. يكون التركيز في عملية الإرشاد على الطالب بدل من التركيز على عملية التعليم.

##### Assessment and feedback:

الهدف منها هو دعم عملية التقييم الذاتي لتحسن فعالية وكفاءة عملية التعلم. والحصول على تغذية راجعة مفيدة للطلاب والمدرسين استناداً على بيانات تتعلق باهتمامات المستخدم وسياق التعلم.

##### Adaptation:

يتم تشغيله من قبل المعلم أو المؤسسة التعليمية. يهدف تحليل التعلم هنا لإعلام المتعلمين بما يجب عليهم القيام به بتكييف مصادر التعلم والأنشطة التعليمية وفقاً لاحتياجات المتعلم.

##### Personalization and recommendation:

يكون التركيز هنا على المتعلم، وكيفية مساعدتهم في اتخاذهم قرارات بشأن عملية التعلم الخاصة بهم لتحقيق أهدافهم من التعلم.

##### Reflection:

يمكن للطلاب والمعلمين الاستفادة من نتائج مقارنة البيانات من أجل نفس المقرر أو بين الصفوف أو حتى بين الجامعات والمؤسسات التعليمية لاستخلاص استنتاجات عن فعالية التعلم والممارسات المتبعة في العملية التعليمية. تؤمن هذه الطريقة الفرصة لتقييم العمل الماضي والخبرات الشخصية بهدف تحسين الخبرات المستقبلية وتعزيز التعلم المستمر.

النقاط السابقة تشكل أهداف فرعية تصب في مصلحة الأهداف الأساسية، والتحدي يكمن باختيار الهدف الأكثر ملائمة للنظام التعليمي المستخدم قبل البدء بعملية التحليل.

الشكل التالي يوضح نسبة الأبحاث في مجال التحليل تبعاً للأهداف المرجوة من عملية التحليل.

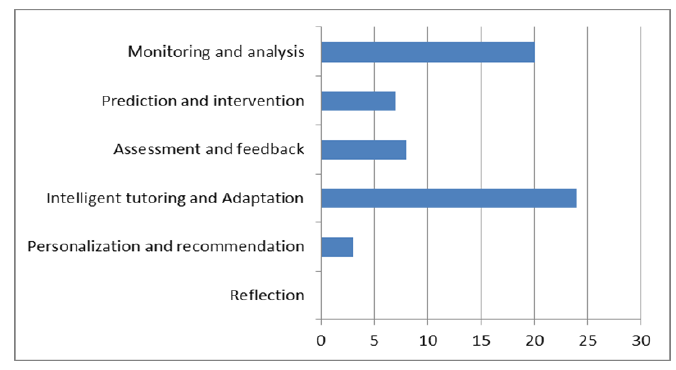


Figure : أهداف عملية التحليل

#### وسائل عملية التحليل:

ورد في الفصل الثاني الوسائل والطرق الممكن اتباعها للقيام بعملية التحليل.

قد تستخدم بعض الأبحاث طرق مختلفة في التحليل في سبيل الوصول للهدف المنشود. بين عامي 1995 و 2005 شكلت الأبحاث المتمحورة حول association rule mining and sequential pattern mining حوالي 43% متبوعة باستخدام classification and prediction التي شكلت حوالي 28% من الأوراق البحثية، وتغيرت الأخيرة لنسبة 42% بحلول عام 2008. بينما فقط 9% من الأوراق البحثية تمحورت حول قواعد الارتباط. الشكل التالي يوضح نسبة الأبحاث في مجال التحليل تبعاً للطرق المتبعة في عمليات التحليل.

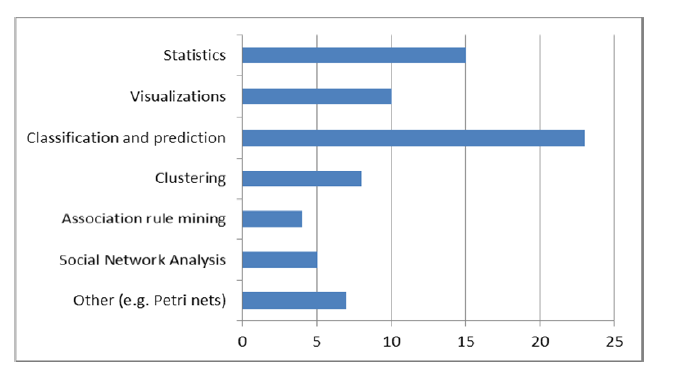


Figure : الطرق المتبعة في عمليات التحليل

#### المستفيدين النهائيين من عملية التحليل:

ورد في الفصل الثاني الأفراد التي بالإمكان توجيه نتائج عملية التحليل لها.

الشكل التالي يوضح نسب ارتباط عمليات التحليل بالمستخدمين النهائيين. حوالي 34% من الأبحاث في مجال التحليل يتم توجيها لخدمة المدرس الذكي intelligent tutor في أنظمة التعليم المتكيفة. وحوالي 30% إلى الباحثين ومصممي النظام التعليمي. بينما يكون التوجه نحو تقديم الخدمات في الأبحاث للمعلمين قليلة مقارنة مع المستخدمين الآخرين.

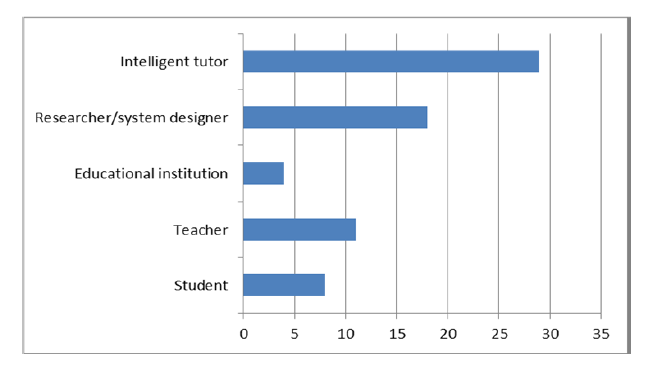


Figure : المستفيدين من عمليات التحليل

## بعض المعايير المستخدمة في الأنظمة التعليمية:

من أبرز المشكلات التي قد تقف عائق في عملية تحليل البيانات التعليمية هو تعدد المصادر المنتجة للبيانات واختلاف بنيتها، فمع تطور المجال التعليمي ظهرت أنواع كثير من الأنظمة التي يصعب مكاملة البيانات فيما بينها، لذلك تظافرت الجهود في الآونة الأخيرة لتسهيل عملية مكاملة هذه الأنظمة وتسهيل عملية تناقل وتبادل البيانات فيما بينها بالاعتماد على معايير موحدة. بالرغم من وجود عدد من المعايير المتبعة في الأنظمة التعليمية إلا أن خصوصية عملية تحليل التعلم فرضت قيود واعتبارات معينة على هذه المعايير. نورد فيما يلي بعض المعايير الموافقة لتحليل البيانات التعليمية:

### Xapi:

يشكل Experience API (xAPI) وثيقة مواصفات تم العمل عليها من قبل مبادرة Advanced Distributed Learning (ADL) من وزارة الدفاع الأمريكية وهي مفتوحة المصدر. النموذج المبدئي للمشروع كان تحت مسمى Tin Can. إلا أن الاسم المتعارف عليه حالياً هو xAPI. إنّShareable Content Object Reference Model (SCORM) و xAPI هما بروتوكولان مختلفان لكنهما يصبوان للأهداف نفسها. بالرغم من أن xAPI تم تطويره مؤخراً إلا أنه لم يحل محل SCORM تماماً بعد.

عند الحديث عن xAPI نتطرق إلى ثلاثة جوانب كما في الشكل التالي:

* Activity Providers: المرحلة التي تقوم بإنتاج البيانات بما يتوافق مع المعيار xAPI ليتم تخزينه ضمن الـ LRS. تتكون مزودات الخدمة من الأنظمة والتطبيقات التي تضم الأنشطة والأحداث التعليمية.
* Learning Record Stores (LRS): تشكل قاعدة البيانات التي يتم فيها تخزين البيانات الموافقة للمعيار xAPI بعد التأكد من صحتها.
* Activity Consumers: وهي أنظمة مشابهة للأنظمة المولدة للنشاطات (في الحقيقة من الممكن أن يكون النظام AC و AC في نفس الوقت) ، تقوم هذه الأنظمة بالعمليات على البيانات القادمة من الـ LRS.

الشكل التالي يوضح مسار العملية كاملة:

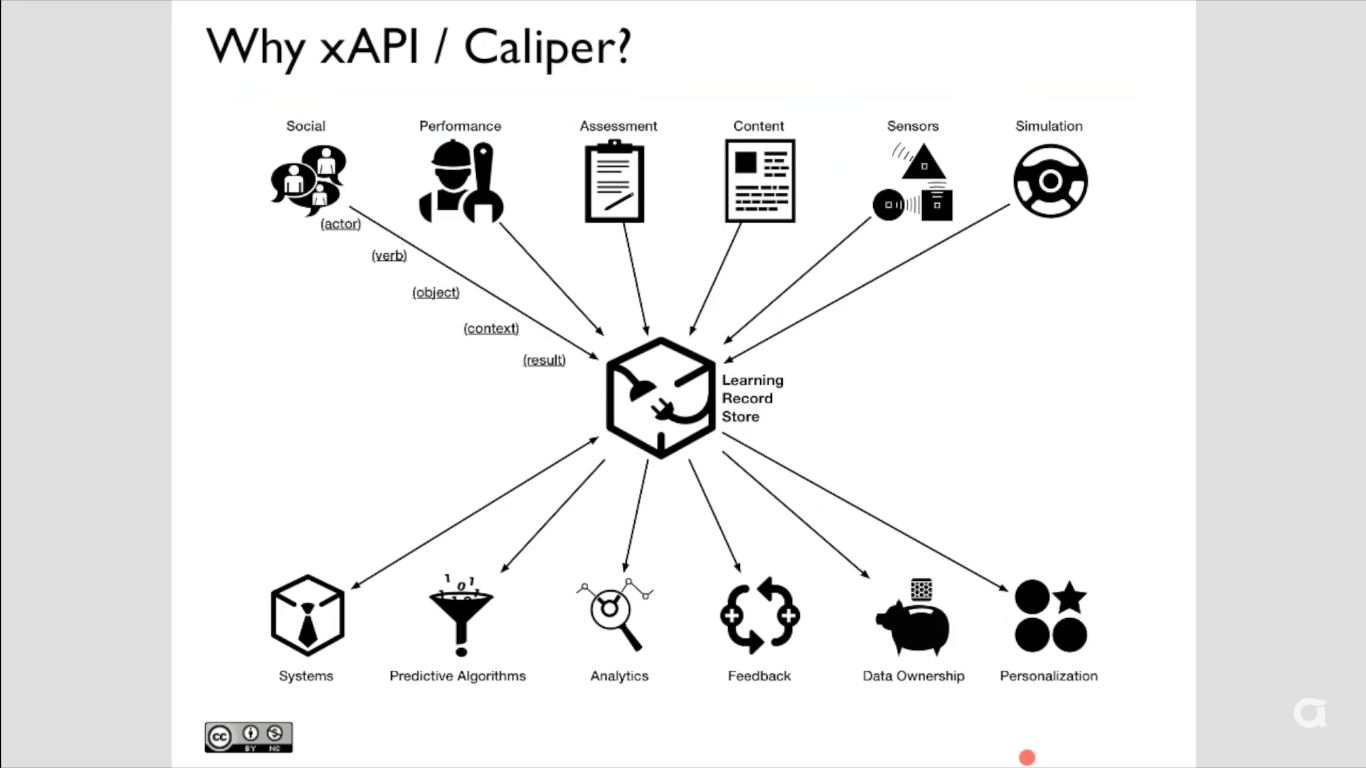


Figure : تدفق العمليات في الأنظمة التي تتبع معيار xAPI

عمل المعيار xAPI على حل مشكلة البينية Interoperability وتسهيل عملية تبادل البيانات المتعلقة بالمجال التعليمي، إلا أن توحيد بنية البيانات باستخدام الثلاثيات التي سيرد ذكرها لاحقاً هو أمر غير كافي. بل يجب التطرق إلى المعنى، فعلى سبيل المثال run تحمل معاني مختلفة تبعاً للسياق، run a workshop, run a business or marathon. لذلك يتم استخدام مجموعة محددة من المصطلحات المتعلقة بالأفعال القابلة للحدوث خلال العملية التعليمية وتدعى **Controlled vocabulary**. فعلى سبيل المثال، إذا أردنا وصف الأفعال المتعلقة بالتفاعل مع فيديو يمكن استخدام الأفعال play, pause, finish، ونفس الأفعال تماماً يمكن استخدامها في كل الأنظمة التعليمية. قام المجتمع المطور لـ xAPI بجمع مصطلحات العمليات المتعلقة بالمجال التعليمي [هنا](http://xapi.vocab.pub/datasets/adl/). تم الذهاب في الـ **recipe** إلى أبعد من ذلك، فبالإضافة للمعنى والبنية المحددة يتم إضافة معلومات إضافة للـ recipe لكل جزء من الـ xAPI Statement. على سبيل المثال: يمكن إضافة زمن الفيديو التي تم مشاهدته للفعل wached وبالتالي يمكن بسهولة حساب الزمن المنصرم على مشاهدة الفيديوهات من قبل شخص. بالإضافة لما سبق، يتم دمج الـ recipes السابقة للحصول على ما يدعى بالـ **profile**، إذ يحوي معلومات كاملة عن سياق معين.

تتصف مخازن البياناتLRS بأنها immutable فحالما يتم تخزين البيانات فيها لا يمكن التعديل عليها أو حذفها.

تكون بنية البيانات التي تتبع معيار xAPI على شكل ثلاثيات كما في الشكل التالي:

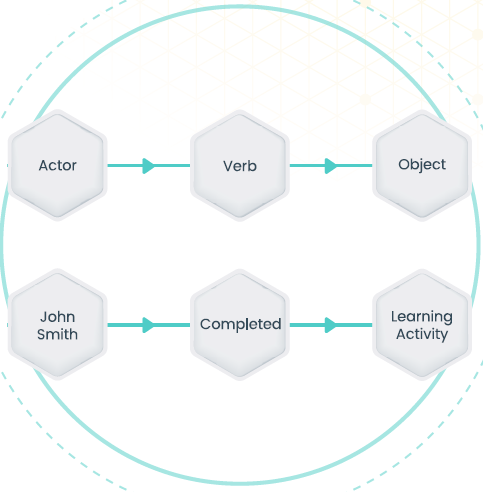


Figure : البنية الأساسية للبيانات الممثلة باستخدام المعيار xAPI

لتوضيح المعنى يتم إضافة معلومات أخرى، فمثلاً للفاعل John يمكن إضافة مكان تواجده عند القيام بالفعل. تقنياً يمكن تمثيل البيانات السابقة باستخدام التمثيل المعياري JSON .

#### بنية البيانات الممثلة باستخدام المعيارxAPI:

* Actor: يمثل من قام بالفعل، ويكون فرد (طالب أو مدرس) أو مجموعة، يتم تعريفه بأربعة طرق: mbox, mbox\_sha1sum, openid, or an account

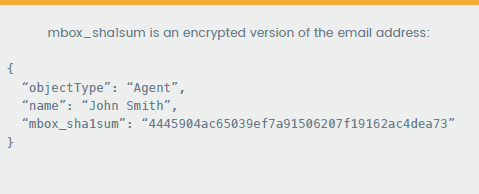


Figure : xAPI statement - mbox\_sha1sum

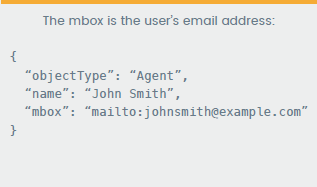


Figure : xAPI statement - mbox



Figure : xAPI statement – openID



Figure : xAPI statement - account

* Verb: تمثل العلاقة بين الـ actor والـ object، ويمكن للأفعال أن تحوي خاصتين: ID يعرف الفعل ولا يتكرر من أجل فعل آخر، و Display مسؤول عن كيفية عرض الفعل بشكل مقروء (كما يمكن ربطه بعدة لغات بشرية).

يُعرّف الـ ID طبيعة الفعل والمعنى منه، فعلى سبيل المثال هل المقصود John “fired” a gun or an employee

* Object: يمثل الشي الذي قام الفاعل بالتفاعل معه، يمكن أن يكون على أربعة أشكال:
* Activity: مثل ‘Jeff wrote an essay about hiking’
* Agent/group: مثل ‘Nellie interviewed Jeff’
* Sub-Statement: مثل ‘Nellie commented on 'Jeff wrote an essay about hiking’
* Statement Reference

بالإضافة للعناصر الثلاثة السابقة يتم إضافة بعض الخصائص الاختيارية لدعم وتوضيح خواص الـ statement، الجدول التالي يوضح خصائص الـ xAPI statement

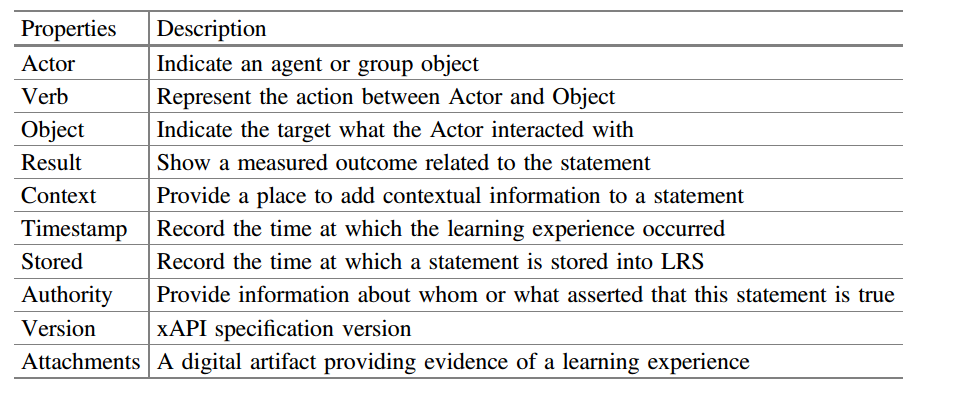


Table : خصائص الـ xAPI statement

يمكن أن يكون هناك بعض الخصائص التي توضح كل خاصية (ويكون الترتيب غير مهم)، مما يعطي البنية الهرمية للـ statement. الشكل التالي يوضح مثال للـ xAPI statement



Figure : مثال للـ xAPI statement

ملاحظة: لا يمكن أي خاصية من الخصائص السابقة أن تتكرر مرتين في نفس الـ statement، ويجب أن تحوي كل statement العناصر 「actor」, 「verb」,「object」

#### مصدر البيانات المتبعة للمعيار xAPI

*تطبيقات حالية:* يوجد أكثر من 150 تطبيق ونظام في العالم ينتج البيانات وفق للمعيار xAPI. هذا يشمل LMS مثل (Moodle, Blackboard, Totaro)، CMS مثل (Wordpress, Drupal)، Authoring Tools مثل (Storyline, Captivate, Lectora)، Social learning systems مثل (Curatr)، Mobile Apps مثل (Tappestry or Learning Locker’s Acorn app)، Performance Support tools مثل (Red Panda, the Trek Learning experience system, or xAPI Apps)، Assessment platforms مثل (TAO testing or Question mark)

*APIsو WeebHooks:* معظم التطبيقات الجديدة، وخاصة التي تعتمد مبدأ SaaS تقدم API كوسيلة للتعامل مع البيانات الناتجة من التطبيق. من الممكن تحويل الـ API المستخدم في التطبيقات إلى xAPI.

*قواعد المعطيات وCsv:* إذا كان من الممكن الوصول المباشر إلى البيانات، يمكن تحويل هذه البيانات إلى بيانات تتبع المعيار xAPI إلا أن عملية التحويل قد تستهلك بعض الوقت والمعالجة الخاصة. عادة ما يقوم المطورون باستخدام script لاستخلاص ونقل البيانات إلى مخزن البيانات LRS. وهنا يجب الإشارة أن هذه العملية لا تتم مباشرة real time، وإنما في أوقات محددة كعطلة نهاية الأسبوع أو في نهاية يوم العمل مثلاً.

### Caliper:

هو معيار مغلق المصدر، يتطلب العضوية فيه ما بين 1500$-55000$ تبعاً لحجم المؤسسة المشاركة. يتم تطويره من قبل IMS. يعمل Caliper على تحقيق نفس الأهداف التي يصبو إليها xAPI حتى أنه يبدو مشابه إلى حد كبير لـ xAPI إلا أنّ Caliper محدود بما يعرف بـ Metric Profiles. ويوجد أشكال محدودة للـ recipes ضمن التوصيف تغطي أحداث التعلم مثل القيام باختبار أو مشاهدة محتوى تعليمي وغيرها.

يحوي كل Metric Profile ثلاثة أنواع من البيانات:

* Entities: تشير إلى المشاركين في التفاعل مع التعلم مثل Person, Assessment, Video وغيرها.
* Actions: تشير إلى الأفعال التي ممكن أن تكون جزء من التفاعل مع التعلم مثل Searched, Viewed and Navigated To.
* Events: تشير إلى الكيانات Entities المشاركة في أحداث التعلم والإجراءات Actions التي يتم تنفيذها. فعلى سبيل المثال الحدث Event: Reading يتضمن من قام بفعل القراءة: Actor ومن وقع عليه الفعل: غالباً ما يكون المحتوى التعليمي.

ويتم تخزين البيانات في Event Storeبدلاً من الـ LRS*.*

بالرغم من استخدام كل منxAPI وCaliperثلاثيات: Actor/Verb/Object بحالة الـ xAPI و Actor/Action/Activity بحالة Caliper ومع اختلاف بينة الـ Action والـ Activity بينهما، إلا أنه يجب أن يكون من الإمكان لكل منهما استخدام الآخر.

عند تطوير كل منxAPIوCaliper يجب أخذ عملية تحليل البيانات الممكن إجراءها على هذه البيانات بعين الاعتبار وعدم الاهتمام فقط ببنية هذه البيانات وإنما تأمين خدمات تسهل عملية التحليل وتزيد فعاليتها وسهولتها.

## خاتمة:

تناول هذا الفصل دراسة للتقنيات والمجالات المتعلقة والمساعدة بالوصول إلى معرفة بناء على البيانات التعليمية والتي تساعد على تحسين أوضاع الطلاب. كما تم البحث في الطرق الممكنة لتقديم خدمة معمّمة لا تقتصر على نظام إدارة تعلم بعينه ومنه تم التوصل إلى ضرورة العمل على معيار موحد في هيكلة البيانات مهما كان مصدرها للتمكن من القيام بعمليات المعالجة المختلفة عليها دون قيود.

# الفصل الثالث: الدراسة المرجعية.

## مقدمة:

يتناول هذا الفصل دراسة لبعض الأنظمة التي عملت على الاستفادة من البيانات الناتجة عن الأنظمة التعليمية في القيام بعلميات تحليل ومعالجة عليها للوصول إلى معرفة تساعد بشكل فعال في تحسين أوضاع الطلاب، ثم المقارنة بينها وتبيان نقاط القوة والضعف فيها للوصول إلى قرار يحدد بنية النظام المقترح والخصائص والميزات التي يتمتع به.

تتصف معظم أنظمة التحليل والتنقيب الموجود حالياً لاحتياجها أشخاص ذوي خبرة للتعامل مع هذه الأنظمة نتيجة تعقيد هذه الأنظمة، إذ أن استخدام المدرسين مثلاً لمثل هذه الأنظمة دون الخضوع لتدريب خاص عملية شبه مستحيلة، بالإضافة إلى أنها تستهلك وقت كبير جداً. علاوة على ذلك، يمكن للمدرسين وحتى زملاء الطلاب في بعض الأحيان الاطلاع على معلومات شخصية غير ضرورية وتجاهل قضايا خصوصية البيانات في تصميم معظم بيئات التعلم الافتراضية. وهنا يجب الانتباه إلى الجانب القانوني، فعلى سبيل المثال في ألمانيا، لا يحق للمدرسين الوصول إلى *جميع* البيانات الناتجة عن أفعال الطالب خلال ارتياد دورات أونلاين، وإنما يمكنهم الوصول فقط إلى البيانات ذات الصلة بالتدريس بشكل يؤمن الشفافية للطلاب.

إلا أن عملية التحليل تهتم باكتساب معرفة واسعة حول بيانات الطلاب، لذلك ولضمان خصوصية الطلاب، يتم اللجوء إلى تزييف هوية الطلاب قبل البدء بالمعالجة، على سبيل المثال، يتم استخدام التجزئة hash مثلاً بدلاً من الهوية الحقيقية للطالبID ، وتقديم النتائج على شكل تصوّرات تشمل المجموعات وعدم التركيز على طالب واحد.

## بعض الأنظمة التي تُعنى بمعالجة المعطيات التعليمية:

### [Completing the Loop:](http://melbourne-cshe.unimelb.edu.au/research/edutech/completing-the-loop)

**هدف المشروع:** يعمل المشروع على **تحليل** عملية التعلم في بيئات التعلم الافتراضية وتقديم نتائج هذا التحليل *بالطرق الأمثل* بهدف الوصول إلى سُبل وإجراءات تفيد المعلمين والمدرسين في رفع سوية التعليم.

**البيئة الحاضنة:** تم تنفيذ المشروع على أربع مقررات ضمن ثلاث جامعات: مقررين في جامعة ملبورن Melbourne، مقرر في جامعة ماكواري Macquarie، وآخر في جامعة جنوب أستراليا the University of South Australia .

**عن المشروع:** تم التعاون في المشروع مع اثني عشر مدرس من ثلاث جامعات، امتد من بداية العام الدراسي 2014 إلى منصف العام الدراسي 2016. نوعية البيانات المقدمة من الجامعات كانت مختلفة لاختلاف طرق التعليم في الجامعات الثلاث. قدم بعض المشاركين اقتراح مصادر إضافية للبيانات غير نظم إدارة التعلم، كالبيانات الناتجة عن أنظمة التقييم، معلومات شخصية عن الطلاب مستقاة من نظم المعلومات الخاصة بالجامعة student information system، ومعلومات الحضور. من وجه نظر المعلمين كانت المشكلة الأساسية هي حاجة المعلمين للتدرَب على استخدام نتائج التحليل وفهمها نتيجة لكون نتائج التحليل غير مفهومة لغير المختصين، إضافة للوقت المطلوب لجمع وتحليل البيانات واتخاذ القرار بناء على نتائج عملية التحليل. أما عن الأداة المبنية فكانت باستخدام Python, Django, Pandas, MySQL للوصول إلى خدمة مرتكزة على الويب، تستخدم بيانات مستوردة من Blackboard وMoodle.

عادة ما يكون هناك فجوة معرفية في أنظمة التحليل ناتجة عن عدم توافر جسر يصل بين المعلومات الناتجة عن تحليل التعلم ونوع النظام والعمليات التعليمية المتبعة في المؤسسات التعليمية، لذلك يهتم المشروع بشقين أساسيين: عملية *تحليل التعلم Learning Analytics* التي تسعى لتحسين العملية التعليمية عن طريق كشف المشاكل الناتجة عن التعلم الإلكتروني وتقديم تغذية راجعة Feedback لكل من الطالب والمعلم، وعملية *تصميم التعلم Learning Design* التي تشمل التخطيط لكيفية سير العملية التعليمية والأنشطة التي تتضمنها. يجب المكاملة بين العمليتين السابقين للوصول إلى فهم أعمق لحالة الطلب، *فسلوك* الطالب في الأنظمة التعليمية لا يعبر بالضرورة عن حالته، على سبيل المثال: تحميل محاضرة لا يعني أن الطلب قد قرء وفهم المحاضرة، وبالتالي لا يجب الاعتماد بشكل كامل على نتائج التحليل بل يجب توخي الحذر عند القيام بالتحليل، والاستعانة بالـ *Learning Design* التي من شأنها أن تساعد في تحديد المعايير التي تجري من خلالها تقييم عملية التعلم.

##### Conversational Framework Laurillard’s:

تم استيحاء فكرة المشروع من Laurillard’s conversational framework والتي ترى أن عمليات التعلم لدى الطلاب بحاجة لأن تكون مدعومة بحلقة متكررة من التفاعل والحوار وتبادل المعلومات بين المعلمين والطلاب. يبدأ تفاعل التعليم عند تصميم النشاط التعليمي وتقديمه من قبل المعلمين إلى الطلاب. تلي هذه العملية انخراط الطلاب بهذا النشاط (كقراءة مصادر التعلم، المشاركة في المناقشة، إلخ) باستخدام فهمهم الحالي للموضوع. تقدم طرق انخراط الطلاب بالنشاط التعليمي شكلاً من أشكال التغذية الراجعة للمعلم. بناء على هذه التغذية الراجعة يمكن للمعلم إعادة تقديم النشاط، وتقديم بعض أشكال المعالجة وبذلك تبدأ حلقة جديدة.

**هيكلية المشروع:** يحوي المشروع مستويين من المستخدمين:

* مدير: يملك صلاحيات
* إضافة مستخدم جديد
* إضافة مجموعة مستخدمين يملكون أذونات وصول محددة.
* منح حق الوصول لمواد محددة
* إنشاء مقرر
* إنشاء حدث لمقرر، وتتراوح أنواع الأحداث بين أحداث تتكرر أسبوعياً (كمحاضرة كل يوم ثلاثاء)، أحداث خلال الفصل (كرحلة ميدانية في الأسبوع الرابع)، وأحداث تقديم (كالاختبارات الممتدة من الأسبوع السابع حتى التاسع من الفصل).
* مدرس: يمكنه الوصول إلى أداة الدعم التربوي وأداة تحليل التعلم من أجل جميع المواد التي يملك صلاحيات الوصول لها.

يقدم المشروع أداتين:

* Pedagogical Helper Tool: توفر للمعلمين فضاء لتبيان العلاقة بين أهداف التعلم، طرائق التعلم، والتكنولوجيا المستخدمة في المقرر. يقوم المدرس بإضافة أهداف التعلم (يمكن إضافة أكثر من هدف) ومن أجل كل هدف يتم إضافة النشاطات التي سيتم اتباعها لمساعدة الطلاب للوصول للهدف (كتحضير المحاضرة قبل ارتيادها). يتم بعد ذلك إضافة مصادر للتعلم من أجل كل نشاط. وبذلك يتم تمثل مرئي للعلاقات ما بين *تصميم* التعلم (الممثل بأهداف ونشاطات التعلم) *وتحليل* التعلم (الممثل بمصادر التعلم). تكون مخرجات هذه المرحلة كخريطة تساعد المعلمين بتفسير البيانات الناتجة عن عملية التحليل.
* Learning Analytics Tool: صُممت لعرض البيانات الناتجة عن التحليل للمعلمين بطريقة مفهومة. تقسم هذه الأداة لثلاث أقسام رئيسية: *course dashboard* تقدم ملخص لتفاعل جميع الطلاب مع نظام إدارة التعلم خلال أسابيع محددة أو خلال كامل الفصل، *course access* تشمل معلومات عن المحتوى، التفاعل والتقدير، *وstudents* التي تسمح بالوصول إلى طالب محدد واستعراض تفاعله مع نظام إدارة التعلم خلال مدة المقرر.

### [Open University Analyse:](https://analyse.kmi.open.ac.uk/)

**هدف المشروع:** يهدف المشروع للكشف المبكر عن الطلاب المعرضين لخطر الفشل والرسوب. احتمالات فشل أو تقصير جميع الطلاب تكون متاحة أسبوعياً لمدرسي المواد وفِرق دعم الطلاب لتقديم الدعم المناسب لهؤلاء الطلاب.

**البيئة الحاضنة:** تم دراسة المشروع على ضوء الجامعة الافتراضية Open university الواقعة في المملكة المتحدة.

**عن المشروع:** أُجري عدد كبير من الأبحاث في سبيل إنجاز المشروع، ومر المشروع بعدة أطوار، نذكر فيما يلي أهم سمات كل مرحلة:

* الطور عام 2013

كان تركيز هذه المرحلة على انعدام التواصل المباشر وجهاً لوجه في الجامعة الافتراضية OU الذي يؤدي لخلق صعوبات بمعرفة المشاكل والعقبات التي يواجهها الطلاب والتي تؤدي لفشلهم. بالإضافة لأعداد الطلاب الكبيرة وضيق وقت المعلمين وعدم إمكانهم ملاحقة حال جميع الطلاب. بالتنبؤ بالطلاب الذين هم بخطر الفشل يتم تقديم الدعم الكافي لهم لتجنب هذا الفشل. نتيجة لاحتواء البيئة التعليمية الافتراضية على معلومات كثيرة شخصية عن الطالب وسلوكه التعليمي ونتائج الوظائف والامتحانات وغيرها، طورت الجامعة الافتراضية OU نماذج توقع تأخذ بعين الاعتبار نوعين من البيانات: أداء الطالب بالامتحانات والمذاكرات الدورية، وتفاعله مع بيئة التعلم الافتراضية. تم استبعاد البيانات المتعلقة بالطلاب قليلي التفاعل مع بيئة التعلم الافتراضية، لأنه وبعد الأبحاث وُجد أن بعض الطلاب لا تُظهر تفاعل كبير مع بيئة التعلم بالرغم من النتائج الجيدة التي يحققوها في المقرر نتيجة لاستخدامهم مراجع أخرى offline دون استخدام بيئات التعلم هذه أو نتيجة كون الطالب يعيد المقرر نتيجة رسوبه به سابقاً. لحل هذه المشكلة تم اتباع طريقة تحتفظ بالسلوك العام للطالب باستخدام learning profile إذ أن سلوك الطالب بشكل عام يكون ثابت "من أجل مقرر ما، ويتغير هذا السلوك من مقرر لآخر"، لتجنب مقارنة سلوك الطالب مع زملائه نتيجة لاختلاف السلوك من فرد لآخر (فمن الممكن أن يحصل طالبان على نفس العلامة بالرغم من الاختلاف بينهما بكمية التفاعل مع بيئة التعلم) يتم مقارنة الفرد مع سلوكه السابق لاكتشاف تغير سلوك الطالب نتيجة لصعوبات يواجهها. كما أنه يمكن الاستفادة من تفاعل *المدرسين* أيضاً في تحليل وتخمين نتائج الطلاب.

* الطور عام 2014

نتيجة الأبحاث تبين أن الطالب الذي يفشل بأول تقييم له غالباً يكمل بالفشل بكامل المقرر، لذلك كان لا بد من اتخاذ إجراءات لمنع فشل الطلاب في أول تقييم لهم وذلك بالتنبؤ بحال الطلاب قبل أول تقييم. تم الاعتماد في عملية التنبؤ هذه على بيانات ديموغرافية شخصية، والبيانات المخزنة في بيئة التعلم الافتراضية وتم الاعتماد على عدة خوارزميات machine learning مثل k Nearest Neighbours, Classification and Regression Tree, Bayes network.

أظهرت الأبحاث أن نسبة الطلاب الذين يكملون في الدورات على شبكة الإنترنت (كالدورات على منصات MOOC) لا تتجاوز الـ 5%، إلا أن هذه الحالة تكون أفضل بكثير في الجامعات التقليدية والـ OU. عادة ما يتم تقسيم الطلاب لمجموعات صغيرة (عشرات الطلاب) حسب موقعهم الجغرافي، وتعيين مشرف يساعدهم ويجيب على أسئلتهم، إلا أن المشكلة الحقيقية تكمن بمساعدة الطلاب الذين بحالة خطرة واكتشاف هذا الشيء بمراحل مبكرة، خاصة وأن كل مقرر يحوي آلاف الطلاب، فمن الصعب توفير موارد كافية لكل هذه أعداد، وتحديد الأشخاص الذين بحالة حرجة للقيام بالتدخل بالوقت المناسب.

نتيجة لكمية البيانات الهائلة الناتجة عن بيئة التعلم الافتراضية تم غربلة هذه البيانات، إذ تم احتساب:

* عدد النقرات في الأسبوع.
* عدد النقرات بالأسبوع وحسب المحتوى الذي تم الضغط عليه.
* Flag يدل عن كون الطالب نشيط وكثير التفاعل في بيئة التعلم أم لا.

وتم تصنيف المحتوى وفق 11 نوع (اختبارات، ويكي، مراجع..) وتحديد نسبة مساهمة التقصير بكل محتوى بفشل الطالب، واعتمدوا على Bayes Theorem لتحديد احتمالية فشل الطالب. أما عن تحليل سلوك الطالب ببيئة التعلم الافتراضية فاستخدموا General Unary Hypotheses Automaton “GUHA” && modelling based on Markov chains. السيئة التي كانت بالـ GUHA هي خسارة المعلومات التي لها علاقة ببعد الزمن، لذلك تم العمل على نموذج آخر يقوم على التحليل بالاعتماد على سلاسل ماركوف.

* الطور عام 2015

اعتمد البحث على نوعين من البيانات: بيانات ديموغرافية وبيانات ناتجة عن تفاعل الطلاب مع بيئة التعليم الافتراضية، باستخدام Bayesian approach يتم استخراج أهم الأنشطة من التفاعلات بحيث يتم استخدام هذه الأنشطة مع البيانات الديموغرافية لبناء أربعة نماذج تنبؤيه:

* + Bayesian classifier
  + Classification and regression tree (CART)
  + k Nearest Neighbours (k-NN) with demographic/static data
  + k-NN with VLE data

يتم استخدام البيانات الديموغرافية (العمر، التعليم السابق، عدد المواد التي قام الطالب بالتسجيل بها...) للقيام بعملية التنبؤ في مرحلة لم يقم الطالب فيها بالتفاعل مع النظام ولم يتم القيام بأي تقييمات بعد، مع مرور الوقت وعندما يبدأ الطالب بالتفاعل مع البيئة الافتراضية يتم تقليل وزن المعطيات الديموغرافية ويصبح التركيز الأكبر على التفاعل مع بيئة التعلم الافتراضية. تكمن الصعوبة بإعطاء **الأوزان** للتفاعلات بين الطالب والنظام التعليمي وكيفية معرفة أيهما أكثر أهمية. صعوبة أخرى تكمن في عملية تحضير البيانات التي تتم بطريقة يدوية. والعمل على رفع معدل الطالب وعدم الاكتفاء بالتنبؤ بنجاح أو رسوب الطالب.

* الطور عام 2016

شمل التطبيق الذي تم تقديمه باستخدام لغة R التحليلية 32593 طالب تشمل بيانات عن تقييمات الطلاب وتفاعلهم مع البيئة التعليمية الافتراضية، وقد تم إخفاء هوية الطلاب باستخدام أداة إخفاء هوية البيانات ARX. جرى في ذلك العام إطلاق Open University Learning Analytics Dataset OULAD وإرفاق شرح عن الجداول والبيانات وكيفية المعالجة.

يشكل OULUD *جزء* من البيانات المجنية من الجامعة الافتراضية OU، إذ تحوي معلومات ديموغرافية ومعلومات عن أداء الطلاب خلال المقررات وتفاعلهم مع بيئة التعلم. تقدم هذه البيانات معلومات عن أداء الطلاب وتقدم فرصة لخلق جيل جديد من نظم إدارة التعلم. تتركز قاعدة المعطيات هذه حول الطالب بدل من المقرر. علاوة على ذلك فإن الجدول أو الملف الرئيسي بقاعدة المعطيات هي معلومات الطالب.

### Student Insight:

**هدف المشروع:** هي منصة لتحليل التعلم تعمل على تحسين أداء الطالب ونتائجهم. إذ تقوم هذه المنصة بتحديد الطلاب المعرضين للخطر والذين هم بحاجة للتدخل في وقت سابق باستخدام التحليلات التنبؤية المتقدمة التي يتم تصميمها لتلبية الخصائص الفريدة للمؤسسة وطلابها. تركيز هذه المنصة على زيادة معدل الاحتفاظ بالطلاب نتيجة لوجود حوالي 10% من الطلاب في الولايات المتحدة يقومون يترك موادهم.

**البيئة الحاضنة:** المشروع هو منتج مدفوع وليس بمفتوح المصدر ولكن التركيز الأساسي فيه على الجامعات في الولايات المتحدة.

**عن المشروع:** يتم الاستفادة في النظام من بيانات الطلاب ونشاطاتهم الموجودة في نظم معلومات الطالب والبيانات الناتجة عن التفاعل مع نظم التعلم الافتراضية لتوقع الطلاب الذين هم بحاجة لتدخل مبكر وخاص، والذين هم بخطر الرسوب. يتم فيما بعد تقييم فعالية التدخل وأخذ تغذية راجعة من الطلاب. يسمح النظام استخدام بيانات قادمة من أي مصدر بشرط احتوائها على بيانات عن الطلاب أو نشاطاتهم ونتائجهم.

كحد أدنى، يحتاج النظام البيانات عن كل طالب مسجل في مقرر في المؤسسة التعليمية، ويتطلب معلومات عن: خصائص الطلاب ومعلوماتهم الشخصية، بيانات حول تاريخ الطالب التعليمي، بيانات الاختبار التحصيلي للطلاب، معلومات التحاق الطلاب كالبرامج المسجل عليها الطلاب، معلومات حول كل مقرر، معلومات حول كل وحدة دراسية. وعادة ما يتم تخزين البيانات السابقة في نظم إدارة التعلم. كما يمكن الاستفادة من البيانات الناتجة عن تفاعل الطالب مع بيئة التعلم والناتجة عن تسجيل أفعاله logging لتدريب النماذج التنبؤية التي تستخدم مشاركة الطلاب مع موارد التعليم والتعلم للتنبؤ بالمخاطر المحيطة بالطالب.

يحتاج النظام إلى بيانات فصل دراسي واحد على الأقل للتمكن من التنبؤ بحالة الطلاب المستقبلية.

بعكس معظم أدوات التحليل التي لا تتيح للمستخدم بتغيير موديل التنبؤ المُستخدم، يتيح هذا النظام للمستخدم إمكانية تغيير موديل التنبؤ بما بتوافق مع المؤسسة التعليمية، كما يسمح للمستخدم بعرض العوامل المساهمة بعملية التنبؤ بحالة الطلاب الذين هم بحالة خطر بطريقة يسهل قراءتها وفهمها.

أما عن دقة عملية التنبؤ فتعتمد بشكل أساسي على العوامل المساهمة في عملية التنبؤ ككمية البيانات التاريخية للطالب ونوعيتها، وعادة ما تصل نسبة الدقة إلى 79%. من أهم ميزات هذا النظام إمكانية مكاملته مع أي نظام إدارة تعلم. في الوقت الحالي يمكن للطاقم الدراسي فقط الوصول لنتائج التحليل أما الطلاب فلا يمكنهم الوصول المباشرة، إلا أن عملية مناقشة حالة الطالب بين الأستاذ والطالب عملية مهمة، لذل يتم العمل حالياً على إشراك الطالب في العملية السابقة.

### PANDA:

هدف المشروع: يؤمن المشروع منصة مفتوحة لتحليل التعلم. ويهدف إلى تأمين منصة مركزية تسمح لنظم إدارة التعلم بالاستفادة منها بعد تمرير البيانات الخاصة بالمتعلمين وتجارب تعلمهم للحصول على معرفة تساعد بتحسين مستواهم التعليمي. يعتمد المشروع على طرق واستراتيجيات متنوعة لتحليل البيانات المجموعة واستكشاف أنماط مفيدة منها.

البيئة الحاضنة للمشروع: يُعد المشروع حصيلة أبحاث استمرت على مدار سنتين بتمويل من الحكومة الألمانية وبمشاركة عدد من الجامعات والمؤسسات مثل Ilmenau University of Technology, the Fraunhofer Institute for Digital Media Technology (IDMT) and the company Magh & Boppert.

تعتمد كل جهة من الجهات المشاركة منصة تعلم مختلفة مثل (Moodle, askMe!, Avendoo) كما أنها تختلف بأنواع احتياجاتها ونوع المستخدمين النهائيين.

عن المشروع: يؤمن المشروع خدمات مختلفة كالإحصائيات، وبطرق عرض مختلفة كالرسوميات والجداول والخرائط وغيرها. كما أن التقنيات التي يتم استخدامها لاستنباط المعرفة متنوعة ما بين تقنيات الذكاء الصنعي والتنقيب في المعطيات وغير ذلك.

يتكون النظام من وحدتين أساسيتين، الأولى لتخزين تجربة الطلاب، والثانية لتحليل هذه البيانات. ما يميز النظام هو **تخزين البيانات في LRS وفق المعيار xAPI** مما يسهل عملية مكاملة هذا النظام مع أنظمة التعلم. وبالتالي فإن عمليات التحليل تتم على الـ statements المخزنة ضمن الـ LRS.

تقدم المنصة أربع أنواع من الخدمات:

* إحصائيات**:** ويتم تقديم النتائج على شكل تقارير وجداول.
* Information Visualization**:** ويتم تقديم النتائج على شكل رسوم بيانية، خرائط، scatter-plots وغير ذلك.
* التنقيب في المعطيات: يستخدم كافة الطرق التي من شأنها الحصول عن الأنماط المخبئة في المعطيات.
* تحليل الشبكة الاجتماعية: وتضمن كافة وسائل تحليل العلاقات بين الأفراد والمؤسسات.

من أهم النواحي التي يهتم بها المشروع هو موضوع **الخصوصية**، ويتم ذلك عبر تجميع البيانات مجهولة الهوية في مخازن المعطيات LRS من نظم إدارة التعلم.

## مقارنة بين الأنظمة المشابهة:

|  |  |
| --- | --- |
| [Loop](https://analyse.kmi.open.ac.uk/) | |
| Loop is an open source analytics tool funded by the Australian Office for Learning and Teaching. The tool can be connected with Moodle or Blackboard to provide a tool for teachers to visualize student behaviors in their learning management system. The dashboard component demonstrates student log data through the learning management site, such as class materials accessed, discussion forum activity, and assessment performance. This data can be viewed at the classroom or individual student level. At the same time, the tool incorporates information about the course structure and schedule into its visualizations. In this sense, the project aims to also incorporate a ‘pedagogical helper tool’ to aid teachers in data interpretation that make sense in their specific context (i.e. ‘pedagogical intent’). In 2015, the tool was piloted with four courses between three Australian universities, with hopes of a wide-scale release following soon. | |
| Type: | Learning environment tool. |
| Role of analytics: | * Visualization. * Monitoring. |
| Data sources: | * Moodle LMS. * Blackboard LMS. |
| Standard: | No Standard. |
| Supply model: | Self-hosted server software. |
| Education Level: | Post-compulsory education. |
| Source Code: | Open source. |
| Compounds Integration: | Not available. |
| Drawbacks: | * Need technical skills. * Manual data feed. * No standard use logs. * No compounds integration. |

Table Existing systems-Loop

|  |  |
| --- | --- |
| [Student Insights](http://www.tribalgroup.com/higher-education/review-and-improvement-services/student-insight/) | |
| Tribal, based in the UK, is a global provider of software solutions specializing in products supporting the management of education.    Tribal's Student Insights is a piece of software that is currently developed to predict student performance and 'at-risk' students.    The software generates predictive models about a student's likelihood to pass a module. The software provides dashboards of this information at student and module level. Managers of universities can use this information, for example, to provide individual student support, or to monitor modules with regard to their predicted performance. | |
| Type: | * Learning support tool. * Analytics for assessment. |
| Role of analytics: | * Visualization. * Prediction. * Modeling. |
| Data sources: | * Management information systems. * Virtual learning environment. * Assessment system. |
| Standard: | No Standard. |
| Supply model: | Shared service Model. |
| Education Level: | Higher Education. |
| Source Code: | Closed source. |
| Compounds Integration: | Yes, but need support. |
| Drawbacks: | * Only prediction at-risk students. * No standards. * Need support to integrate new compounds & data sources. |

Table Existing systems-Student Insights

|  |  |
| --- | --- |
| [ASSISTments](http://www.aboutus.assistments.org/) | |
| ASSISTments is an intelligent tutoring system (ITS) developed by Neal Heffernan and colleagues and is researched at Worcester Polytechnic Institute (WPI) in collaboration with a variety of universities and organizations in the United States. The core system was designed to give progressive hints to students who answer a question incorrectly, in order to simulate the type of instantaneous directed feedback a tutor would provide. From this platform, there have been a variety of studies of the system focusing on how to use the student log data generated from the system effectively. For example, studies have been carried out to see how this data can influence parent engagement or predict performance on high stakes tests. | |
| Type: | * Learning support tool. * Analytics for assessment. |
| Role of analytics: | * Visualization. * Prediction. * Modeling. * Recommendation. |
| Data sources: | Use own Data. |
| Standard: | No Standard. |
| Supply model: | Privately-hosted software. |
| Education Level: | Mainly Secondary Education, but can be used for any level. |
| Source Code: | Closed source. |
| Compounds Integration: | Limited to only pre-defined LMS. |
| Drawbacks: | * Focus only on homework and quizzes. * You need to use the system to get the feedback and analysis. |

Table Existing systems-ASSISTments

|  |  |
| --- | --- |
| [X-Ray](https://uki.moodlerooms.com/) | |
| X-Ray Analytics is a predictive modeling tool linked with Moodle LMS & Moodlerooms framework, and was recently acquired by Blackboard in 2015. The dashboard provides teachers with visualizations of past behaviors in their learning management system at multiple levels: course, multiple course and intuitional. Its algorithms then make predictions on future performance and behaviors to identify ‘at-risk’ students potentially in need of an intervention. The tool also considers student engagement by analyzing contributions to online collaborative tools, such as discussion forums, using social network analysis. A student can be identified as ‘at-risk’ according to the time they have spent in the course, their grades and their discussion forum engagements. X-Ray Analytics uses a cloud-based model and analyses pre-existing data in the learning management system. The tool is expected to be available for all Blackboard courses in the near future. | |
| Type: | Learning environment tool. |
| Role of analytics: | Summarization & description.  Visualization.  Prediction. |
| Data sources: | Moodlerooms LMS. |
| Standard: | No Standard. |
| Supply model: | Self-hosted server software. |
| Education Level: | Higher & Post-compulsory education. |
| Source Code: | Open source. |
| Compounds Integration: | Not available. |
| Drawbacks: | * Limited use with Moodle and Moodlerooms. * No Standards. * No compounds integration. * Need Purchased license of X-Ray Learning Analytics service. |

Table Existing systems- X Ray

|  |  |
| --- | --- |
| [SmartKlass™ | The Learning Analytics Plugin](http://klassdata.com/smartklass-learning-analytics-plugin/) | |
| SmartKlass™ is Learning Analytics Moodle dashboard that should be included as a part of the Moodle virtual learning platform to empower teachers to manage the learning journey of their students. We use xAPI 1.0 specification (aka Tin Can API) to collect user interaction with the platform. Our technology analyzes the statements collected using machine learning algorithms and builds dashboards analysis with consolidated user information. | |
| Type: | Learning environment tool. |
| Role of analytics: | * Summarization & description. * Visualization. |
| Data sources: | Moodlerooms LMS. |
| Standard: | xAPI Standards. |
| Supply model: | Plugin. |
| Education Level: | Higher & Post-compulsory education. |
| Source Code: | Open source. |
| Compounds Integration: | Not available only with Moodle. |
| Drawbacks: | Limited use with Moodle LMS.  No compounds integration. |

Table Existing systems-SmartKlass™

|  |  |
| --- | --- |
| [Skillaware](http://skillaware.com/) | |
| Skillaware is a company base in Italy that designs learning environment software for workplace learning and training. The program is used simultaneously with pre-existing company software or procedures to determine worker effectiveness and areas where training may be useful. Through a variety of tools, Skillaware captures user activities and behaviors within existing software.    The SkillEditor function captures user behaviors and automatically suggests trainings to make their use of various forms of software more productive. The SkillAgent function provides suggestions for next steps in a task when a user seems to need assistance.    Additionally, the SkillAnalyzer tool allows company analysts to watch real-time user activity and provide data visualization for management staff, monitors user performance and learning analytics during the rollout of new software and software processes across your organization. | |
| Type: | Learning support tool.  Design and planning tool. |
| Role of analytics: | Summarization & description.  Visualization.  Recommendation. |
| Data sources: | Uses own data. |
| Standard: | xAPI Standards. |
| Supply model: | Unknown. |
| Education Level: | Workplace. |
| Source Code: | Closed source. |
| Compounds Integration: | Yes, with compounds provide by the same company. |
| Drawbacks: | Directed to Business Process Improvement and Guidance not learning.  Limited compounds integration. |

Table Existing systems-Skillaware

يمكن تلخيص ما سبق بالجدول التالي:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Loop | Student Insights | ASSISTments | X-Ray | SmartKlass | Skillaware | OU |
| Type | Learning environment tool |  |  |  |  |  |  |  |
| Learning support tool |  |  |  |  |  |  |  |
| Analytics for assessment |  |  |  |  |  |  |  |
| Design and planning tool |  |  |  |  |  |  |  |
| Standard | |  |  |  |  |  |  |  |
| Data  Source | LMS | Moodle  Blackboard |  |  | Moodle rooms | Moodle rooms |  |  |
| Management Information system |  |  |  |  |  |  |  |
| VLE |  |  |  |  |  |  |  |
| Assessment system |  |  |  |  |  |  |  |
| Owned data |  |  |  |  |  |  |  |
| Integration | |  |  | limited |  | limited |  |  |
| End  User | Teacher |  |  |  |  |  |  |  |
| Students |  |  |  |  |  |  |  |
| administrators |  |  |  |  |  |  |  |

Table : مقارنة بين الأنظمة المشابهة

## نقاط الضعف في الأنظمة السابقة:

يمكن تصنيف الحلول المقدمة في مجال تحليل البيانات التعليمية على ثلاثة أشكال:

* Tightly coupled LA Solution: مثل نظامي .ASSISTments, X-Rayتقوم بتقديم خدمة تحليل البيانات الناتجة عن التفاعل مع المحتوى التعليمي إلى جانب تقديم خدمة تعليمية أو تساهم في العملية التعليمية، فعلى سبيل المثال: يعمل نظام ASSISTments على تقديم خدمات تنبؤ، اقتراح، تمثيل مرئي وغير ذلك للبيانات الناتجة عن الاختبارات المتواجدة في **النظام نفسه**.

سيئة هذه الأنظمة أنه لا يمكن استخدام النظام كخدمة في الأنظمة الأخرى، كما لا يمكن الاستفادة من ميزة التحليل إلا بعد القيام بالاستفادة من ميزة الاختبارات التي يقدمها النظام.

* Pluggable LA Solution: مثل نظام .SmartKlass تقوم بتقديم النظام على شكل إضافة تضاف للنظام التعليمي، كالإضافة الممكن استخدامها مع نظام إدارة التعلم Moodle.

سيئة هذه الأنظمة محدودية الخدمات التي تقدمها والأنظمة التي تلائمها.

* Standalone LA Solution: مثل نظامي Loop, Student Insight. يتم تقديم خدمات التحليل بنظام منفصل عن النظم المولّدة للبيانات مثل نظاميLoop, Student Insight إلا أن معظم هذه الأنظمة تتصف بعدم اتباعها معيار موحد وبالتالي وملاءمتها لأنظمة محدودة.

يكون شكل النظام المقترح قائم بحد ذاته Standalone وبالتالي فهو قادر على تقديم خدمة تحليل البيانات بغض النظر عن نوع النظام المولد للبيانات والخدمات التي يقدمها لتفادي محدودية عملية التحليل والمؤسسات والجهات المستفيدة بشرط اعتماد معيار موحد لعدم حد نوع المؤسسات التعليمية المستفيدة من خدمات هذا النظام ودعم أنظمة ومؤسسات مختلفة ومتنوعة.

## النظام المقترح:

بناء على الدراسة التي أُجريت للأنظمة التي تقوم بتحليل المعطيات الناتجة عن الأنظمة التعليمية المتوافرة حالياً وبالمقارنة بينها ودراسة نقاط ضعفها، تم وضع صورة مبدئية للنظام الذي سيتم بناءه:

يتصف النظام المقترح بأنه نظام قائم بحد ذاته Standalone Learning Analytics System يقوم باستقبال البيانات من مصادر مختلفة ومتنوعة شرط أن تكون مهيكلة وفق معيار معين ليتم فيما بعد تخزينها في مخزن معلومات بعد أن يتم التحقق من صحة هيكليتها. كما يتمتع النظام بالمرونة الكافية لإمكانية إضافة مصدر جديد للمعطيات. ولتحقيق هذه الميزة يجبر النظام الأنظمة الراغبة بالاستفادة من خدمات هذا النظام بهيكلة بياناتها وفق نموذج معين يتم الاتفاق عليه بشكل مسبق وبالتالي عدم الاقتصار على دعم نظام إدارة تعلم بعينه. وتقديم خدمات تحليل التعلم بشكل ديناميكي وسهل دون الحاجة للاستغناء عن النظم السابقة المسؤولة عن إدارة عملية التعلم المستخدمة في المؤسسات التعليمية.

تتنوع الخدمات التي يقدمها النظام تبعاً لنوع المستخدم النهائي، ويمكن لكل من يشارك في العملية التعليمية الاستفادة من هذه الميزات. فعلى سبيل المثال لا الحصر: يقدّم النظام للمعلمين معلومات مهمة تساعدهم في اتخاذ قرارات لتغيير الطرق المتبعة في العملية التعليمية تبعاً لدراسات إحصائية يقدمها النظام للمعلمين وتدل على فعالية الطرق الحالية المتبعة. ويمكن للمرشدين تلقي التنبيهات من النظام عند وصول أحد الطلاب لوضع حرج يستدعي التدخل المباشر لتفادي فشل هؤلاء الطلاب. كما يمكن للطلاب الاطلاع على دراسات إحصائية حول سلوكه ونشاطاته المنفّذة في نظم التعلم الافتراضية وتلقي الإرشادات والاقتراحات ببعض الأحيان.

## خاتمة:

تناول هذا الفصل دراسة لأهم الأنظمة المتواجدة والتي تهتم بتحليل البيانات التعليمية، مقارنة للميزات والمساؤى التي يتمتع بها كل نظام لتحديد أولويتها وتجنب نقاط الضعف في النظام المقترح. كما تم ووضع تصور مبدأي للنظام الذي سيتم بناءه.

# الفصل الرابع: الدراسة التحليلية.

## مقدمة:

يعرض هذا الفصل جوانب المشروع كافة بشكل عملي وعلمي دقيق إذ أنه سيتم تفصيل متطلبات المشروع الوظيفية وغير الوظيفية، وآلية تحقيقها.

## دراسة جدوى المشروع:

### الجدوى التقنية:

بعد تحديد أهداف المشروع، وطبيعة الوظائف التي يقدمها وأساليب تقديمها، وبمراعاة الإمكانيات المتاحة في  
عملية تطوير وانجاز المشروع على كل من الصعيدين المالي والبشري، تم الاتفاق على استخدام اللغات البرمجية التالية لإنجاز المشروع:

### الجدوى المالية:

المشروع المقدم عبارة عن مشروع طلابي يقوم على تطويره طلاب في كلية الهندسة المعلوماتية، لذلك فإن عملية التطوير والبرمجة تتم على الأجهزة والمخدّمات الافتراضية المتوفرة على الحواسيب الشخصية  
لأعضاء الفريق المطور، والتي لا يترتب على استخدامها أي كلفة تذكر. كما أن جهد أعضاء الفريق المطور  
يتم تقديمه مجاناً في سبيل التعلم، لذلك لا يوجد أي عائد مالي لقاء تنفيذ هذا المشروع.  
أما عن طبيعة تقديم المشروع لمستخدميه من طلاب ومؤسسات تعليمية فسيكون مجاني نتيجة لتقديم المشروع بشكل مفتوح المصدر open source ودعم لخاصية الـ openness وتحقيقاً للمرونة المطلوبة في النظام والتي تميزه عن غيره من أنظمة التحليل.

### الجدوى التشغيلية:

الإطار القانوني: النظام المقترح موافق للإطار القانوني في سوريا، وطرحه في متناول المواطنين لا يتطلب إصدار قوانين جديدة، كما أنه لا يتعارض مع القوانين الموجودة حالياً.

الإطار التنظيمي: تحتاج المؤسسة المستخِدمة لهذا النظام إلى استخدام معيار xAPI وهيكلة بياناتها وفق هذا المعيار للتمكن من استخدام النظام.

### الجدوى التنظيمية:

يقوم على المشروع فريق مطِور من أربع طلاب يملكون خبرات متعددة ومختلفة، ووفقاً لهذه الخبرات ولأهداف أعضاء الفريق في التعلم وتطوير الذات، وللحدود الزمنية للمشروع، تم تقسيم المهام وتنظيمها بين أعضاء الفريق.

### جدوى الموارد البشرية:

ترتب على أعضاء الفريق المطِور للمشروع القيام بتنسيق العمل فيما بينهم، وامتلاك الخبرة والمهارة البرمجية في التعامل مع اللغات والتقنيات البرمجية المذكورة سابقاً في فقرة الجدوى التقنية.

## الفاعلون في النظام:

* مطور المؤسسة التعليمية: هو الشخص الذي يقوم بإعداد وتهيئة أنظمة المؤسسة التعليمية لتستخدم الخدمة التي نقدمها في نظامنا.
* مدير النظام: وهم المسؤولين عن إدارة نظام التحليل المطروح.
* الطالب: المنضم إلى مؤسسة تعليمية ويمتلك حساب فيها.
* الأستاذ: وهو المسؤول عن إعطاء مقرر أو أكثر في المؤسسة التعليمية ومتابعة طلابه. وهنا يجب الأخذ بعين الاعتبار إمكانية استعرض بيانات تخص طلابه فقط.
* إداري: يشكل العنصر الإداري في المؤسسة التعليمية، كالعميد مثلاً، أو منسق المواد والمحتوى أو أي شخص يمتلك صلاحيات الاطلاع على العلامات والطلاب والمعلمين، كالمسؤول عن المعلمين في المؤسسات التعليمية أومندوب الوزارات وغير ذلك.

## المتطلبات الوظيفية:

|  |  |
| --- | --- |
| الاسم | مطور المؤسسة التعليمية |
| الصلاحيات | * تسجيل المؤسسة التعليمية كزبون جديد * إضافة Data Source وحذفها. * تعديل إعدادات التطبيق. * تسجيل الدخول والخروج على النظام. |

|  |  |
| --- | --- |
| الاسم | مدير النظام |
| الصلاحيات | * الموافقة على إضافة مؤسسة تعليمية جديدة. * تعديل محتويات نظام التحليل. * حظر تطبيق مؤسسة وتفعيلها. * حظر مستخدم وتفعيله. * تسجيل الدخول والخروج على النظام. |

|  |  |
| --- | --- |
| الاسم | طالب، أستاذ، إداري |
| الصلاحيات | * إضافة مؤشر indicator جديد. * استعراض مؤشر indicator جديد. * استخدام مؤشر indicator جديد وتحديد اعداداته وطرق عرض نتائجه. * حذف مؤشر indicator تم إضافته مسبقاً. * تسجيل الدخول والخروج على النظام. |

## المتطلبات غير الوظيفية:

الجدول التالي يعرض أهم المتطلبات غير الوظيفية التي يجب تحقيقها في النظام.

|  |  |
| --- | --- |
| Data privacy | يجب أن يأخذ النظام بعين الاعتبار مبدأ خصوصية البيانات وعدم عرض البيانات الشخصية على الآخرين دون إذن صاحب هذه البيانات. |
| Standardization | يجب على النظام اتباع معايير موحدة في هيكلة البيانات وتقديم الخدمات وتخزين النتائج وغير ذلك. |
| Interoperability | يجب أن يتمتع النظام بالقدرة على تبادل البيانات مع عدد كبير من الأنظمة ودعم أنظمة متنوعة، ويتم تحقيق قابلية التشغيل البيني بتقليل عدم التجانس بين البيانات التي تولدها أنظمة التعلم الافتراضي. |
| Transparency | يجب أن يتمتع النظام بالشفافية وأن يكون المستخدم على إطلاع على نوعية العمليات التي سيتم تنفيذ البيانات عليها وعلى من سيتم عرضها. |
| Modularity | يجب أن تتمتع معيارية النظام بقابلية استيعاب مكونات جديدة بهدف الاستجابة للتغيرات مع مرور الوقت. |
| Usability | على النظام أن يكون سهل الاستخدام والاهتمام بطريقة عرض النتائج واستخدام رسومات مرئية Visualizations سهلة القراءة للمستخدم النهائي. |

Table : متطلبات النظام غير الوظيفية

## بنية النظام:

الهدف الرئيسي من النظام المقترح هو تطوير أداة تسمح لنظم التعلم المختلفة بالاستفادة من بيانات طلابهم في الحصول على معلومات مفيدة تعزز تعلمهم. يعتمد النظام على أساليب تحليل وتنقيب متنوعة لتحقيق تلك الأهداف. الشكل التالي يوضح بنية النظام المقترح.

The Developed System

Prediction

Prediction

Monitoring

reflection

Statements

Data Source

LMS, SIS, MOOC, etc.

Analytics Engine

LRS

Visualization

Personalization

Assessment

Adaption

Tutoring

Figure : رسم توضيحي للنظام المقترح

* Data Source: تشكل المولد للبيانات التي تمتلكها المؤسسات التعليمية. ولا يهم مصدر البيانات طالما أنها تقوم بإرسال البيانات وفق المعيار xAPI.
* LRS: يتم فيه تخزين البيانات الهيكلية ليتم لاحقاً معالجتها وإجراء العمليات عليها بعد التأكد من صحة بنيتها.
* Learning Analytics & Data Mining: وهي الجزئية في النظام المسؤولة عن معالجة البيانات التي يتم قراءتها من مخزن البيانات LRS. تحوي جزئيات components عديدة، تهدف لتأدية أهداف متنوعة تصب في مصلحة الطالب والمؤسسة التعليمية. وتكون كل جزئية مسؤولة عن تحقيق عملية معينة. يتمتع النظام بقابلية التوسع وإضافة الجزئيات المختلفة. أما عن الجزئيات التي سيتم تطويرها فهي:
  + Monitoring Student Engagement: تؤمن هذه الجزئية دراسات إحصائية لتفاعل الطلاب مع المحتوى وبيئات التعلم الافتراضيات وغير ذلك. ويتم تقديم الخدمات بالاعتماد على مؤشرات indicators عديدة تختلف باختلاف المستخدم النهائي، ويكون بعضها معرف بشكل مسبق في النظام مثل:
    - Top n elements: تشمل عرض للعناصر مرتبة (الممتدة على فترة زمنية معينة) بالاعتماد على تقييمها وفعاليتها. مثلاً يمكن عرض المحاضرات الأكثر مشاهدة وتحميلاً خلال الأسبوع الماضي، أو عرض للأستاذ أكثر الطلاب تجاوباً وعرضاً للمحتوى.
    - Activities: تشمل عرض لنشاطات الطالب (الممتدة على فترة زمنية معينة) من أجل عدة أنشطة، مثل استعراض المحتوى، التواصل مع الآخرين، تقديم الاختبارات، وغيرها من أجل مقرر معين.
    - Content Metrics: تشمل معلومات عن محتوى معين، كمحاضرة من طول المحاضرة وعدد حضورها وغير ذلك.
  + Predict at risk student: تتم بالاعتماد على خوارزميات التنقيب، إذ يتم التنبؤ بنتائج الطلاب وبالتالي تحديد الطلاب الذين هم بخطر الفشل.
* Visualization: تؤمن هذه الجزئية عرض نتائج التحليل بطريقة سهلة وقابلة للقراءة لبعض النتائج صعبة الفهم والاحاطة. ويتم العرض بواجهة منفصلة عن نظام إدارة التعلم.

## الخاتمة:

# الفصل الخامس: الدراسة التصميمية.

## مقدمة:

## عملية ما قبل المعالجة:

كلما كانت عملية تحضير البيانات الخام قبل البدء بمعالجتها أكثر، كلما أعطت عملية المعالجة نتائج أفضل. ولكن عملية التحضير هذه تتطلب جهد يدوي كبير، إذ أن هذه العملية تستهلك حوالي 60-90% من الوقت والجهد والموارد. فيما يتعلق بالبيانات الناتجة عن الأنظمة التعليمية نجد أن كمية البيانات ضخمة جداً وربما لا تكون بالشكل الأنسب وبالتالي تكون عملية التحليل غير مجدية. لذلك يجب العمل على تهيئة هذه البيانات للحصول على النتائج الأمثل من عمليات التحليل. هناك أمور أخرى يجب الاهتمام بها كعملية المكاملة بين البيانات الناتجة عن أنظمة مختلفة.

يمكن تلخيص أهم المهام التي يتم إنجازها في عملية ما قبل المعالجة بما يلي:

data cleaning, user identification, session identification, path completion, transaction identification, data transformation and enrichment, data integration, data reduction.

نتيجة لاعتماد معيار موحد في هيكلة البيانات تتقلص كمية العمل الواجب القيام به على البيانات بغية تجهيزها قبل البدء بمعالجتها.

## حالات الاستخدام: