

التنقيب في البيانات واكتشاف المعرفة كآلية لتحسين جودة القرار في

المؤسسة الاقتصادية

الاسم واللقب: د. بلعباس رابح	الاسم واللقب: د. صلاح محمد	الاسم واللقب: د. طلال زغبة
الرتبة العلمية: أستاذ محاضر (أ)	الرتبة العلمية: أستاذ محاضر (أ)	الرتبة العلمية: أستاذ محاضر (أ)
المؤسسة: جامعة المسيلة	المؤسسة: جامعة المسيلة	المؤسسة: جامعة المسيلة
الهاتف: 0549681117	الهاتف: 0776853763	الهاتف: 0663 92 64 92
البريد الإلكتروني:	البريد الإلكتروني:	البريد الإلكتروني:
rabah_belabbas@yahoo.fr	slahmohamed83@gmail.com	zaghbatalal@gmail.com

محور المشاركة: جمع، تنقيب و تمايل المعطيات.

ملخص:

تتيح التغيرات المتسارعة في مجال تكنولوجيا المعلومات والاتصالات والكم الهائل من البيانات الضخمة التي نعيشها اليوم والتزايد في أعداد برامج التنقيب في البيانات، اختصار كثير من الوقت للحصول على المعلومات التي تساعد متخذي القرار على اتخاذ القرار الأنسب، وتهدف هذه الورقة البحثية إلى عرض مفهوم وأهمية التنقيب في البيانات من أجل اكتشاف المعرفة، وذلك بتحويل البيانات من مجرد معلومات متراكمة وغير مفهومة إلى معلومات قيمة يمكن استغلالها والاستفادة منها في تحسين جودة القرار.

كلمات مفتاحية:

تكنولوجيا المعلومات والاتصال، التنقيب في البيانات، اتخاذ القرار، اكتشاف المعرفة، تحسين جودة القرار.

Résumé :

Les changements accélérés dans le domaine des technologies de l'information et de la communication, le volume gigantesque des données disponibles et le numéro croissant des programmes d'exploration de ces dernières, que nous vivons actuellement, ont permis d'écourter considérablement le temps nécessaire pour l'obtention de l'information indispensable à la prise de décision efficace. Dans cet ordre d'idées, le présent article prétend exposer le concept et l'importance de l'exploration des données dans la découverte des connaissances, et ce en transformant les données du stade de simples informations accumulées et incompréhensibles en informations précieuses pouvant être exploitées dans l'amélioration de la qualité de la décision.

Mots clés :

Technologies de l'information et de la communication, exploration des données, prise de décision, découverte des connaissances, amélioration de la qualité de la décision.

المحور الأول: التنقيب في البيانات

من المعروف أن عملية تجميع البيانات في سجلات بأحجام كبيرة قد تحتوي على معارف كبيرة، والتي قد تكون مفيدة لأصحابها من خلال معرفة العديد من الإحصاءات المطلوبة مستقبلاً، فالتقنيات العادية قد لا تكون مفيدة وغير عملية في مثل هذه الأحجام الكبيرة من البيانات، لذلك تواجدت تقنيات التنقيب في البيانات لتسهيل وتحسن عملية البحث والاستنباط.

I- ماهية التنقيب في البيانات

أولاً: نشأت التنقيب في البيانات

إن التطور في العلم والاقتصاد وتكنولوجيا المعلومات أدى إلى زيادة كمية البيانات الرقمية، هذه الكميات الهائلة من البيانات لم تعد وسائل التحليل التقليدية (الإحصائية مثلاً) قادرة على التعامل معها، فظهر التنقيب عن البيانات (Data mining) منذ أواخر الثمانيات واثبت وجوده كأحد الحلول الناجحة لتحليل كميات ضخمة من البيانات، وذلك بتحويلها من مجرد معلومات متراكمة وغير مفهومة (بيانات) إلى معلومات قيمة يمكن استغلالها والاستفادة منها بعد ذلك (معرفة)، حيث ظهر مصطلح التنقيب في البيانات في منتصف التسعينات في الولايات المتحدة الأمريكية بإسهام من عدة تخصصات من بينها الإحصاء وقواعد البيانات والذكاء الصناعي والتعلم الآلي، فقد بدأ الاهتمام بالتنقيب عن البيانات عام 1989 أثناء انعقاد ورشة عمل حول اكتشاف المعرفة في قواعد البيانات، ومن ذلك الحين تم عقد هذه الورشة بصفة مستمرة سنوياً حتى عام 1994، أما في 1995 أصبح المؤتمر الدولي لاكتشاف المعرفة والتنقيب عن البيانات من أهم الأحداث السنوية¹.

ثانياً: تعريف التنقيب في البيانات

يوجد عدة تعاريف لمفهوم التنقيب في البيانات منها²:

أ. عملية اكتشاف المعرفة في قواعد البيانات عن طريق تحديد الأنماط والاتجاهات في البيانات التي تم جمعها

باستخدام طرق مختلفة مثل التصنيف Classification أو التحليل التسلسلي Sequential

analysis أو العنقدة (التجميع) clustering أو قواعد الارتباط Association Rule؛

ب. تحليل البيانات المتواجدة في قواعد البيانات باستخدام الأدوات التي تبحث عن الاتجاهات أو البيانات

التي لا معنى لها، واستخراج معلومات ضمنية، لم تكن معروفة سابقاً، ويمكن أن تكون مفيدة.

من خلال التعريفات السابقة يمكن القول أن التنقيب في البيانات عبارة عن سيرورة لاستخراج أو

اكتشاف معرفة مفيدة و قابلة للاستغلال من خلال مجموعة كبيرة من البيانات.

ثالثاً: أسباب نمو تطبيقات التنقيب في البيانات

تطبيقات التنقيب في البيانات بدأت تنمو بصورة كبيرة للأسباب التالية:

أ. كمية البيانات الموجودة في مخزن البيانات وسوق البيانات تنمو بصورة أسية، ومن اجل ذلك فإن المستخدم يحتاج إلى أدوات متطورة من مثل التنقيب في البيانات من اجل استخلاص الفائدة والمعرفة من هذه البيانات؛

ب. الكثير من أدوات التنقيب عن البيانات بدأت تظهر مؤخراً وكل أداة أفضل من الأخرى؛

ج. المنافسة الشديدة الموجودة في السوق تدفع المؤسسات إلى الاستفادة القصوى من البيانات التي بيدها، وعمليات التنقيب في البيانات تساعد في ذلك.

رابعاً: مزايا استخدام التنقيب في البيانات

التنقيب في البيانات توفر معلومات دقيقة وصحيحة وبشكل سريع ومزايا أخرى نستعرضها في مايلي³:

د. تساعد متخذ القرار على تفعيل الترابط بين الأقسام والأعمال المختلفة، لأن ذلك ينعكس إيجاباً على مصلحة العمل في المنشأة؛

هـ. تسهل التعامل مع تقنيات المعلومات المتطورة، وتساعد على قياس فاعلية نظم المعلومات الفرعية المختلفة وإنتاجيتها، من خلال توفير المعلومات الدقيقة والصحيحة؛

و. تساعد على الاستخدام الفعال لمصادر البيانات والموارد المتاحة، وفي التخطيط لتحسين نظم المعلومات المحاسبية والمصرفية المستخدمة وتطويرها؛

ز. تساعد على الكشف عن قدرة المؤسسة على النمو ومواكبة التطور، وحاجتها إلى تطوير المصادر الفنية والبشرية لنظم المعلومات المستخدمة، ومن ثم فإن المعلومات تساعد على زيادة المعرفة، والحد من البدائل، وهذا الأمر يؤدي إلى التخلص من حالة عدم التأكد التي تتمثل في هذه البدائل؛

ح. تمكن متخذ القرار في أن يخطط لاستمرارية إجراء الفحص وتحديد المشكلة وعناصرها وإتمام الرقابة على ملفات المعلومات والبيانات اللازمة لقراراته.

II- مهام وعناصر التنقيب في البيانات ومجالات وتحديات تطبيقها

أولاً: عناصر التنقيب في البيانات

يمكن القول إن التنقيب عن البيانات يعتمد على أربعة عناصر أساسية وهي⁴:

أ. **البيانات**: هي عبارة عن الحقائق والأرقام والنصوص التي يمكن أن تعالج من قبل الحاسب.

ب. **المعلومات**: النماذج والعلاقات بين تلك البيانات التي تشكل معلومات مفيدة.

ج. **المعرفة**: المعلومات السابقة يمكن أن تحول إلى معرفة حول الأنماط التاريخية أو التوقعات المستقبلية، مثال

معلومات عن حركة المبيعات والمشتريات للزبائن يمكن أن تزودنا بمعرفة عن سلوكهم الشرائي، فيساعدنا ذلك في معرفة أي من المواد تحتاج إلى ترويج أكثر.

د. **مستودعات البيانات**: المستخدمة في التحليلات الزمنية واكتشاف المعرفة واتخاذ القرارات، فهي مصممة

خصيصاً لاستخلاص البيانات ومعالجتها وتمثيلها وتقديمها بصورة مناسبة لهذه الأغراض، وتخزن كمية ضخمة من البيانات قد تكون من مصادر مختلفة، مثلاً عدة قواعد بيانات من عدة نماذج.

ثانياً: مهام التنقيب في البيانات

تصنف مهام التنقيب في البيانات إلى مهمتين أساسيتين هما⁵:

أ. **مهمة التنبؤ Predictive**: هناك العديد من الطرق التي تعمل بهذا الخصوص وفكرتها تقوم على

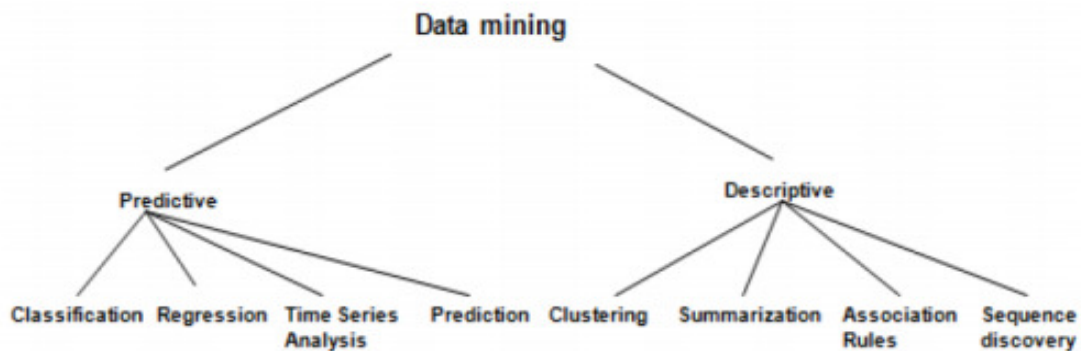
استخدام بعض المتغيرات للتنبؤ بالقيم المستقبلية وغير المعروفة لمتغيرات أخرى، أي تحاول إيجاد أفضل التنبؤات اعتماداً على المعطيات ويعتمد هذا التنقيب على استخدام المعلومات القديمة لتوقع ما سيحدث في المستقبل.

ب. **مهمة الوصف Descriptive**: قائمة على إيجاد أنماط قابلة للتفسير بواسطة الإنسان والتي يمكن أن

تستخدم لوصف البيانات أو ظاهرة معينة، أي تعتمد على إعادة تنظيم البيانات والتنقيب في أعماقها لاستخراج المؤشرات الموجودة فيها.

أ. هناك العديد من الطرق للتنقيب في البيانات إما طرق تنبؤية أو طرق وصفية كما في الشكل التالي:

الشكل رقم (1) تصنيف مهمات التنقيب في البيانات



- أ. التصنيف **Classification**: يتمثل في تفسير أو التنبؤ بخاصية فرد ما من خلال خصائص أخرى.
- ب. الانحدار **Regression**: يستخدم في التنبؤ بالقيم الجديدة بالاعتماد على القيم الموجودة.
- ج. تحليل السلاسل الزمنية **Time Series Analysis**: تتنبأ بالقيم غير المعروفة بالاعتماد على سلاسل تأخذ بالحسبان الخوص المميزة للزمن.
- د. التنبؤ **Prediction**: يشبه التنبؤ التصنيف أو التقدير، ما عدا أن البيانات تصنف على أساس التنبؤ بسلوكها المستقبلي أو تقدير قيمتها المستقبلية.
- هـ. التجميع العنقودي **Clustering**: البحث عن مجموعات متجانسة في مجتمع من الأفراد.
- و. التلخيص **Summarization**: تفتيت كتل البيانات وتقديم وصفا عاما للمتغيرات وعلاقاتها.
- ز. قواعد الارتباط **Association Rules**: البحث عن علاقات موجودة بين عدة خصائص.
- ح. اكتشاف التسلسل **Sequence Discovery**: تستخدم لتحديد نماذج متسلسلة في البيانات.

ثالثاً: تحديات التنقيب في البيانات

تواجه عملية التنقيب عدة تحديات نوجزها فيما يلي⁶:

- أ. القدرة علي الاستيعاب والتوسع **Scalability** : تقيس مدي قدرة الخوارزميات الموجودة حالياً على معالجة القدر الضخم والهائل من البيانات.
- ب. البعدية **Dimensionality**: تحتوي البيانات الآن علي آلاف الخصائص المختلفة بالتالي لن تستطيع الطرق التقليدية للتحليل معالجتها بشكل سليم.
- ج. البيانات المعقدة وغير المتجانسة **Heterogeneous and Complex Data**: تتعامل الطرق التقليدية للتحليل عادةً مع بيانات ذات سمات موحدة.
- د. ملكية البيانات **Data Ownership and Distriution**: أحيانا نحتاج إلى بيانات غير مخزنة في مكان ثابت ولا تنتمي لمؤسسة بمفردها مما يتطلب منا تطوير تقنيات لتوزيع تلك البيانات بأمان.
- هـ. جودة البيانات **Data Quality**:
- و. حماية الخصوصية أثناء عملية التنقيب في البيانات **Privacy Preservation**:
- ز. البيانات المتدفقة التي يتم توليدها بشكل متسلسل ومستمر **Streaming Data**:
- ح. التحدي الأهم هو المهارات والخبرات التي لا بد أن تتوافر فيمن سيستخدم هذه التقنيات، حيث لا بد أن يتمتع بخبرة كبيرة في تحليل البيانات والتعامل مع المصادر المفتوحة.

رابعاً: مجالات تطبيق التنقيب في البيانات في المؤسسات الاقتصادية

لقد عرف التنقيب في البيانات نجاحاً كبيراً وازدادت تطبيقاته في المؤسسات الاقتصادية، وذلك راجع إلى النمو الكبير في البيانات، وقد بدأ استخدام التنقيب في البيانات في مؤسسات التوزيع، ثم انتقلت بعدها سريعاً إلى البنوك، ومؤسسات التأمين، ومؤسسات الهاتف (خاصة الهاتف النقال)، و مؤسسات المياه والكهرباء، و مؤخراً مؤسسات النقل الجوي والنقل بالسكك الحديدية... الخ، وكانت تطبيقاتها الأولى في ميدان إدارة العلاقة مع الزبون من خلال تحليل سلوك الزبائن من أجل الوصول إلى جعلهم أكثر وفاء و اقتراح منتجات وفق رغبتهم، وعموماً هناك عدة ميادين لتطبيق التنقيب في البيانات في منظمات الأعمال نذكر منها⁷:

- أ. **التسويق**: استخدمت الشبكات العصبية الاصطناعية في دراسات التسويق المستهدف بما في ذلك الحصص السوقية، وقد ساعدت هذه الأساليب التسويق على استخدام نهج تخصيص الزبائن وفقاً إلى الحقائق الديمغرافية (السكانية) الأساسية مثل الجنس والعمر والمجموعات وكذلك أنماطهم الشرائية؛
- ب. **التجزئة**: لقد استخدمت أساليب التنقيب في البيانات بصورة فعالة في التنبؤ بالمبيعات حيث أخذت العديد من المتغيرات في الدراسات مثل متغيرات السوق المتعددة، قدرات الزبائن المستندة على العادات المتبعة في الشراء، كما ساعدت أساليب مثل تحليل السلة الشرائية أو السلة السوقية كثيراً على إيجاد أي من المنتجات التي يمكن أن تشتري سوية من قبل الزبائن؛
- ج. **البنوك**: لقد أثبتت تنبؤات الأعمال والمالية على أنها الأساليب الممتازة في تطبيقات أساليب التنقيب في البيانات، وقد استخدمت هذه الأساليب في تنبؤات السعر المستقبلية وأداء الأسهم، كما حقق استخدام مثل هذه الأساليب النجاح في تطوير أنظمة القياس الرقمية في تحديد مخاطر القروض والاحتيالات المالية؛
- د. **التأمين**: لقد استخدمت أساليب التنقيب في البيانات كذلك وبصورة واسعة في مجال شركات التأمين وخاصة في تخصيص مجاميع الزبائن بهدف تحديد أسعار البوليصة وذبذبات المطالبات المتوقعة مستقبلاً، وتحديد المطالبات المزيفة؛
- هـ. **الاتصالات**: حيث استخدمت وبصورة فعالة أساليب التنقيب في البيانات مثل الشبكات العصبية في محاولة تقليل الاضطرابات التي تظهر عندما تفقد المؤسسة هؤلاء الزبائن خلال المنافسة وذهابهم إلى المنافسين الآخرين؛
- و. **إدارة العمليات**: حيث استخدمت الشبكات العصبية في عمليات التخطيط والجدولة وإدارة المشاريع بالإضافة إلى إدارة الجودة.

المحور الثاني: مدخل مفاهيمي للمعرفة واكتشافها

I: مفهوم المعرفة ومصادرها وتميزها عن المفاهيم المشابهة

أولاً: تعريف المعرفة

✓ المعرفة عبارة عن رأس مال فكري وقيمة مضافة، ولا تعد كذلك إلى إذا اكتشفت واستثمرت من

المؤسسة، وتم تحويلها إلى قيمة لخلق الثروة من خلال التطبيق.

✓ المعرفة هي معلومات بالإمكان استخدامها واستثمارها للوصول إلى نتائج مفيدة⁸.

من خلال ما سبق فإن المعرفة هي مزيج من المعلومات والخبرة، أي هي الفهم المكتسب للمعلومات من

خلال الخبرات والدراسة.

ثانياً: مصادر المعرفة

تنقسم مصادر المعرفة إلى قسمين⁹:

✓ **المصادر الخارجية:** وهي تلك المصادر التي تظهر في بيئة المؤسسة المحيطة، كالمكتبات والانترنت

والانترانت، والقطاع الذي تعمل فيه المؤسسة، والمنافسون لها، والموردون والزبائن والجامعات ومراكز

البحث العلمي... الخ.

✓ **المصادر الداخلية:** هي تلك المصادر التي تنشأ من خبرات أفراد المؤسسة المتراكمة حول مختلف

الموضوعات وقدرتها على الاستفادة من تعلم الأفراد والجامعات والمؤسسة من المؤتمرات الداخلية،

والمكتبات الالكترونية، الحوار، أو من خلال التعلم بالعمل والبحوث الداخلية.

ثالثاً: التمييز بين البيانات والمعلومات والمعارف والحكمة

يوماً بعد يوم يزداد النتاج الفكري زيادة هائلة تجعل من العسير على المتلقي متابعة كل ما

يصدر وينشر من معلومات، الأمر الذي يؤدي إلى إغفاله عن الحصول على معلومات مهمة

تفيد في دراساته أو أعماله على اختلاف أنواعها، ولمواكبة ما ينشر ظهرت الحاجة إلى إيجاد أجهزة وبرمجيات

لحصر هذا الكم الهائل من البيانات ومعالجتها، منذ كونها بيانات وحقائق خام أوليه ليست ذات قيمة بشكلها

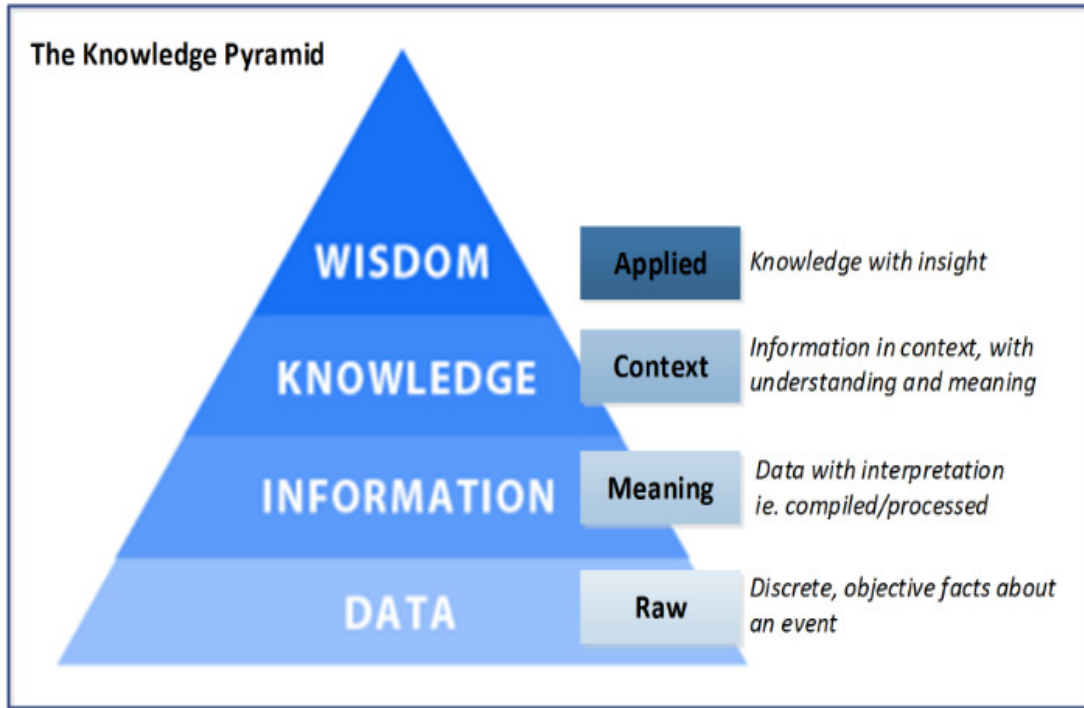
الأول متناثرة هنا وهناك، إلى أن تصبح معلومات متجانسة تغطي معنى خاص، لتتحول إلى معارف التي تمثل

حصيلة الامتزاج الخفي بين المعلومات والخبرة والمدرجات الحسية، وصولاً إلى قمة الهرم المتمثلة بالحكمة التي تعني

إتقان الفعل والقول والنفاد إلى باطن الأمور وحسن تدبرها والمعرفة الأتم، والشكل الآتي يوضح العلاقة التي تتجسد

بين البيانات والمعلومات والمعارف والحكمة.

الشكل رقم (2) الشكل الهرمي للبيانات والمعلومات والمعارف والذكاء



Source: Division Of Information Technology, 21/10/2017, site: <https://www.csu.edu.au/division/dit/eal/portfolios/information>

- ✓ **البيانات Data**: هي حقائق غير منظمة تؤخذ من الملاحظات والسجلات ولا تعطي معنى.
- ✓ **المعلومات Information**: بيانات تمت معالجتها (تصنيفها ، تحليلها ، تنظيمها وتلخيصها) بحيث أصبحت ذات معنى وتوفر معرفة حول مسألة معينة.
- ✓ **المعرفة Knowledge**: المعرفة هي نتاج معالجة المعلومات مع الخبرة الشخصية .
- ✓ **الحكمة Wisdom**: هي حسن استخدام المعرفة المتاحة.

II: مشاركتها المعرفة، معوقاتها ودورة تكوينها

أولاً: مفهوم مشاركة المعرفة

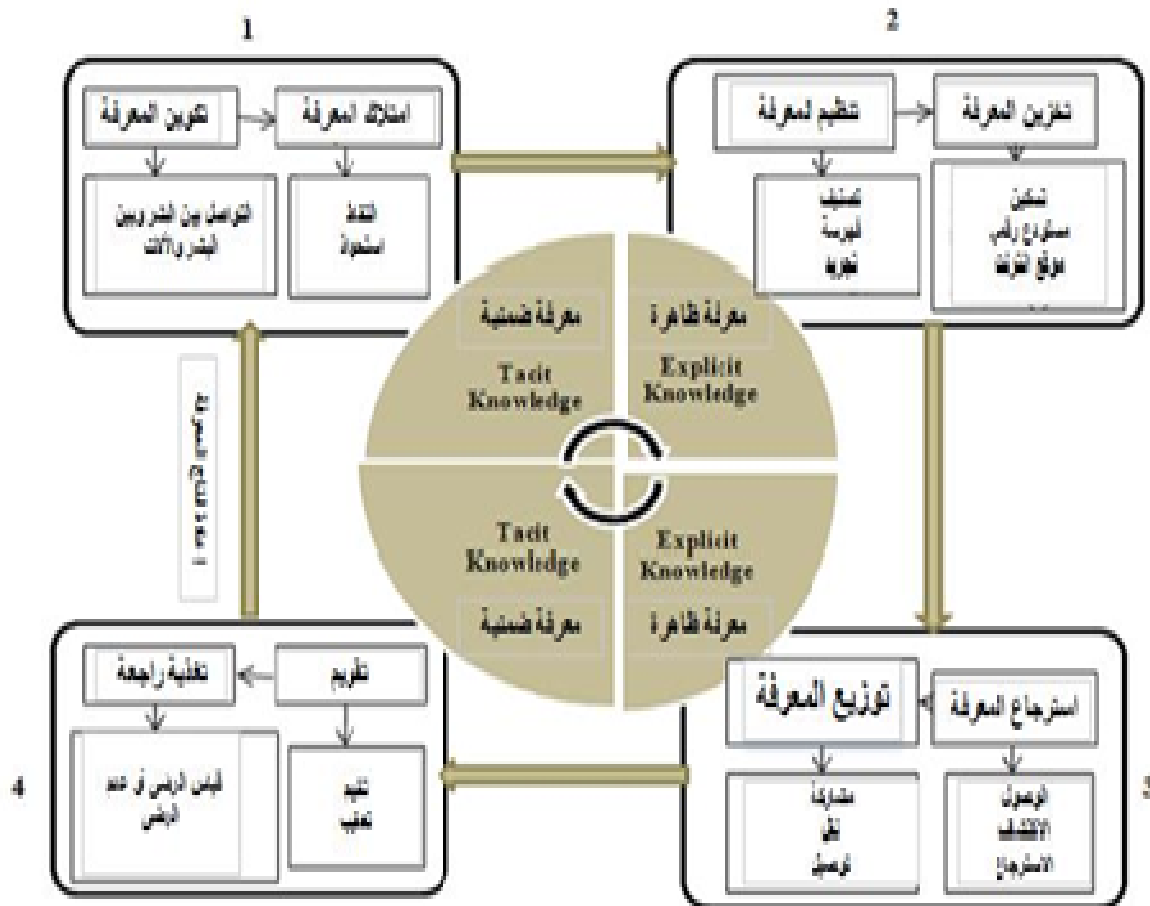
- علمية نشر المعرفة أو تشارك المعرفة هي احد العمليات الأساسية لإدارة المعرفة تعرف كمايلي:
- أ. مشاركة المعرفة تتمثل في عملية الاتصال بين اثنين أو أكثر من المشاركين بحيث يكونوا مشمولين في عملية نقل المعرفة هذه العملية تتضمن مخزوناً معرفياً متاحاً لدى مصدر المعرفة ومن خلال تفسير الاتصالات بواسطة واحداً أو أكثر من متلقي المعرفة فإن مخرجات هذه العملية تمثل خلقاً لمعرفة جديدة¹⁰؛
 - ب. إن التشارك في المعرفة مطلب أساسي للمؤسسة هدفه إدارة العمليات داخل المؤسسة، بهدف نقل المعرفة الملائمة للشخص الملائم في الوقت الملائم، كما أنها تمكن الأفراد داخل المؤسسة من استخدام المعرفة التي تم التشارك بها في اتخاذ القرارات المناسبة¹¹؛

مما سبق فان تشارك المعرفة هي العلمية التي يتم من خلالها نقل المعرفة الصريحة أو الضمنية إلى الأفراد الآخرين، من خلال الاتصالات التي تتم بين هؤلاء الأفراد بواسطة مختلف وسائل الاتصال.

ثالثاً: دورة تكوين المعرفة

يقترح نوناكا وتاكيوتشي Nonaka et Takeuchi نموذجاً يعبراً عنه بأربعة حروف (SECI)، وهي تمثل أربع عمليات فرعية في دورة تكوين المعرفة، والتي تتكون من خلال دعم ثقافة تشارك ونقل المعرفة وذلك عبر توافر الأنظمة التقنية التي تدعم التشارك والبحث والتنقيب، ووجود قنوات حرة لنقل المعلومات والمعرفة كما هو موضح في الشكل التالي¹²:

الشكل رقم (3) مراحل بناء المعرفة حسب نموذج Nonaka et Takeuchi



المصدر: عماد سرحان، إدارة المعرفة.... الطريق نحو التميز، 2017/10/28، موقع:

<https://taelum.org/%D8%A5%D8%AF%D8%A7%D8%B1%D8%A9-%D8%A7%D9%84%D9%85%D8%B9%D8%B1%D9%81%D8%A9/>

أ. **التنشئة الاجتماعية Socialization**: وتمثل بتحويل المعرفة الضمنية إلى معرفة ضمنية، وتحدث عندما يتشارك الفرد مع فرد آخر فيما يمتلكه كل منهما من معرفة ضمنية عن طريق تبادل الأفكار والمهارات والخبرات، ومن ثم تتطور لتصل إلى المرحلة الثانية؛

ب. **التجسيد Externalization**: من ضمنية إلى صريحة وتحدث عندما يكون الفرد قادراً على أن يبين أو أن يلفظ بوضوح ما يمتلكه من معرفة ضمنية، وهو بذلك يقوم بتحويل المعرفة الضمنية التي يمتلكها إلى معرفة معلنة، أي هي عملية إظهار المعرفة الضمنية وتجسيدها لتتحول إلى معرفة صريحة من خلال عملية الاتصال واعتماد لغة الحوار والتفكير الجماعي.

ج. **التركيب Combination**: من صريحة إلى صريحة وفي هذه المرحلة يكون الفرد قادراً على دمج وتركيب أجزاء غير مترابطة من المعرفة المعلنة، وبالتالي فهي عملية دمج وتصنيف المعرفة الصريحة لتحويلها إلى معرفة صريحة جديدة.

د. **الاستيعاب الداخلي Internalization**: تمثل هذه المرحلة ظاهرة انتقال المعرفة الصريحة إلى ضمنية تضاف إلى الخزين المعرفي للأفراد بالذاكرة الإنسانية، وفيها يدخل الأفراد الذاتية على تجاربهم وخبراتهم من خلال التعلم واستخدام الأدلة والإرشادات.

ثانياً: معوقات تشارك المعرفة

هناك معوقات تحد من عملية مشاركة المعرفة نذكر منها¹³:

- أ. الرغبة في الاحتفاظ بالمعرفة والخوف من مشاركتها الآخرين للمحافظة على المكاسب المادية؛
- ب. مخاوف من احتمالية المنافسة من الغير؛
- ج. مخاوف أصحاب المعرفة من أن يفقدوا سلطتهم وقوتهم نتيجة مشاركة المعرفة، وبالذات ما يمتلكونه من معرفة ضمنية؛
- د. الخوف من مشاركة المعرفة الخطأ و تعريض الآخرين للضرر؛
- هـ. عدم وجود حوافز لمشاركة المعرفة، فالأفراد الذين هم على استعداد لإشراك الآخرين فيما لديهم من معرفة يرغبون بالحصول علي عائد، وقد يحجموا عن إشراك الآخرين إذا شعروا بأنه ليس هناك فائدة أو تعويضاً ينتظرهم.

المحور الثالث: اكتشاف المعرفة من البيانات وتحسين جودة القرار في المؤسسة الاقتصادية

على الرغم من التراجع في تجارة السلع والخدمات في كثير من أنحاء العالم، إلا أن التدفق الدولي للبيانات يواصل نموه، فقد بلغ حجم البيانات العابرة للحدود 211 تيرابايت في الثانية في عام 2014، ما يُعادل 1.3 حجم بيانات مكتبة الكونغرس في كل ثانية، ويبلغ أكثر من 45 مرة مقدار البيانات المتبادلة في عام 2005.

أولاً: تنامي حجم البيانات المخزنة في قواعد البيانات

هناك عبارة شهيرة للكاتب John Naisbitt سنة 1982 تقول "نحن غارقين في البيانات ولكن جائعين للمعرفة We are drowning in information but starved for knowledge"، وهذه العبارة تعني تضاعف كمية البيانات المخزنة في وسائل الإعلام المختلفة خصوصاً خلال ثلاث سنوات من الفترة 1999 إلى 2002، حيث أن كمية البيانات التي وضعت للتخزين في تلك الفترة بلغت حوالي 5 Exabytes، وهي كمية كبيرة وهائلة من البيانات مساوية تقريباً في حينها لمحتويات نصف مليون مكتبة جديدة تحتوي كل منها على النسخ الالكترونية لجميع المطبوعات في مكتبة الكونغرس في م¹⁴، إذا هذه العبارة تشير إلى أن هناك كميات كبيرة من البيانات المولدة بواسطة أدوات ووسائل مختلفة، ولكن الجدل الذي يمكن أن يثار هو ما مقدار الاستفادة من هذه البيانات؟، وما مقدار المعرفة التي يمكن أن تكتشف من هذا الكم الهائل من البيانات؟، بحيث تفيد في تحسين عملية اتخاذ القرار وإيجاد حلول لمشاكل عملية، والشكل التالي يوضح مدى تنامي البيانات.

الجدول رقم (1) تطور حجم البيانات خلال الفترة 1999-2002

Organization	Scale of Data
Walmart	~ 20 million transactions/day
Google	> 4.2 billion Web pages
Yahoo	~10 GB Web data/hr
NASA	satellites ~ 1.2 TB/day
NCBI GenBank	~ 22 million genetic sequences
France Telecom	29.2 TB
UK Land Registry	18.3 TB
AT&T Corp	26.2 T

Source : Stephen Cass, "A Fountain of Knowledge," IEEE Spectrum, January 2004

من خلال الجدول يتضح انه هناك تنامي سريع للبيانات من حيث النوع والحجم لمختلف المؤسسات العالمية، والميزة الايجابية التي جعلت من الحواسيب أداة للتفاعل الالكتروني تكمن في قدرتها على معالجة كميات

هائلة من البيانات بموثوقية عالية وسرعة كبيرة، لكن نقطة الضعف الكبيرة لأجهزة الحواسيب تكمن في انه ليس لديها أدنى فكرة عن ما تعنيه أي من تلك البيانات في الواقع، وهو ما يشير إلى الحاجة للمعالجة التحليلية لهذه البيانات حتى يظهر معنى لها وهو ما يسمى بالمعرفة.

ثانياً: خطوات اكتشاف المعرفة من البيانات لصنع القرار

اكتشاف المعرفة طريقة تتضمن أدوات ومنهجيات لتحليل البيانات النصية والرقمية بآن واحد، وتنقيب البيانات تشكل أحد خطوات عملية اكتشاف المعرفة، وتتضمن عملية اكتشاف المعرفة الخطوات الموضحة في الشكل التالي¹⁵:

الشكل رقم (4) خطوات عملية اكتشاف المعرفة

The Knowledge Discovery Process

- Data Mining v. Knowledge Discovery in Databases (KDD)

- ▶ DM and KDD are often used interchangeably
- ▶ actually, DM is only part of the KDD process



Source: Luis Goldster, Data mining and knowledge discovery, 31/10/2017,
<https://www.slideshare.net/LuisGoldster/data-mining-and-knowledge-discovery-61715851>

✓ الاختيار Selection: اختيار البيانات التي سنطبق عليها تقنيات التنقيب في البيانات.

✓ **المعالجة الأولية Preprocessing:** البيانات التي تم اختيارها تحتاج إلى معالجة أولية لكي نتغلب على المشاكل الخاصة بالبيانات ، كونها جاءت من مصادر مختلفة وغير متجانسة، لذا يتم في هذه الخطوة تصحيح الأخطاء وتعبئة البيانات الناقصة، وهي تأخذ جهد ووقت أكبر.

✓ **التحويل Transformation:** يتم تحويل البيانات إلى صيغة مناسبة لتقنيات تنقيب البيانات .

✓ **تنقيب البيانات Data Mining:** استخدام تقنيات وطرق التنقيب في البيانات لاكتشاف المعرفة وهناك العديد من الخوارزميات التي يمكن أن تستخدم في هذا المجال.

✓ **التقييم Interpretation/Evaluation:** المعرفة التي نحصل عليها تسمى أنماط Patterns تحتاج إلى تفسير وتقييم لأخذ المفيد منها لصانع القرار.

ثالثاً: مساهمة المعرفة في صنع القرار وتحسين جودته

بناءً على ما سبق فإن اكتشاف وتشارك المعرفة تساهم في تحسين جودة القرارات المتخذة وذلك كما يلي:

أ. إمكانية اتخاذ أفضل القرارات

تتم عملية اتخاذ القرارات على أساس أفضل المعلومات التي يتشاركها الفرد مع الآخرين، فالأفراد عادة بحاجة إلى الدعم عند اتخاذ القرارات ويتم ذلك من خلال تقديم المعلومات حول الخيارات والمخاطر، والمعلومات المطلوبة تختلف باختلاف نوع الاستخدام فبعض القرارات تكون من البساطة بحيث تكفي معلومات بسيطة لأخذ القرار، لكن عندما يتعلق الأمر بقرارات حاسمة فهنا تختلف طبيعة المعلومات المستخدمة لاتخاذ القرار،

ب. تقليص دورة القرار

يساعد توفر كم هائل من المعارف على تقليص دورة عملية صنع القرار وذلك راجع إلى:

✓ كنتيجة حتمية فإن دورة اتخاذ القرار تقلص عن طريق المعارف الجديدة، خصوصاً مع استخدام الحاسوب وما صاحبه من طفرة كبيرة في تكنولوجيا المعلومات، ومن ثم الاستفادة من سرعة ودقة الحاسوب في تنفيذ التطبيقات المختلفة وبالتالي السرعة في اتخاذ القرار؛

✓ توفر المعرفة بسهولة ويسر يؤدي بم اتخاذ القرار إلى عدم تحمل أي تكلفة بما يزيد عن منافعها؛

✓ تتجسد أهمية التنقيب في البيانات في أنها تقود إلى تراكم المعارف وزيادتها عبر الزمن، ومعنى ذلك أن الفرد الأكثر خبرة يحتاج إلى معلومات أقل ومن ثم تكلفة القرار تكون أقل؛

✓ الأهمية النفسية التي تحملها المعلومات والمعرفة غير المستخدمة المتراكمة من الشبكات العنكبوتية، والإحساس بأنها موجودة وجاهزة للاستخدام مهما استدعت الحاجة إليها.

النتائج والتوصيات:

إن هدف البقاء والنمو والاستمرار يتطلب من المؤسسة العمل الجاد والدائم من أجل تحقيق مزايا تنافسية، من خلال الاتجاه نحو الذكاء الاقتصادي والتنقيب في البيانات وذلك لاكتشاف المعرفة التي تدعم جودة قراراتها.

• النتائج:

- ✓ التنقيب في البيانات عبارة عن سيرورة لاستخراج أو اكتشاف معرفة مفيدة و قابلة للاستغلال من خلال مجموعة كبيرة من البيانات؛
- ✓ هناك تنامي سريع للبيانات من حيث النوع والحجم لمختلف المؤسسات العالمية،
- ✓ اكتشاف المعرفة طريقة تتضمن أدوات ومنهجيات لتحليل البيانات النصية والرقمية بآن واحد، وتنقيب البيانات تشكل أحد خطوات عملية اكتشاف المعرفة؛
- ✓ تعد المعرفة ركيزة الأساسية لتحقيق جودة القرار فوجود معلومات واضحة يكسب المدراء وأصحاب القرار فهما متزايدا لبيئة العمل؛
- ✓ وجود المعلومات المتعمقة من مجموعة متنوعة من المصادر يؤدي إلى اتخاذ القرارات بشكل أفضل.

• والتوصيات

- ✓ ضرورة استخدام تكنولوجيا التنقيب عن البيانات بمفاهيمها المعرفية والتقنية مع ما يلزم لذلك من أساليب وأدوات لتعزيز ثقافة المؤسسة المعرفية، وتوفير كل ما يلزم من مخازن ومستودعات البيانات في تكوين معارف جديدة من شأنها أن تساعد في تطوير أداء المؤسسة؛
- ✓ ضرورة استخدام التنقيب عن البيانات لسهولة تنظيم المعلومات والوصول إلى المعرفة بسهولة في ظل التضخم الهائل في البيانات؛
- ✓ عقد العديد من الندوات والمؤتمرات فيما يتعلق بالتنقيب عن البيانات، مع توضيح أوجه الاستفادة طبقا للتطورات التكنولوجية الحادثة في المجال؛
- ✓ الاهتمام بموضوع تبني المفاهيم الحديثة في نجاح نظم دعم القرار في المؤسسة، ولغايات توفير المعلومات المناسبة في الوقت المناسب، للمساعدة في اتخاذ القرارات السليمة؛
- ✓ التركيز على تدريب وتطوير متخذي القرار على مثل هذه التقنيات الحديثة لكي يسهل عليهم استخدام أو تطبيق البرنامج والتوصل إلى المعلومات المطلوبة في اتخاذ قراراتهم؛
- ✓ تشجيع و تنمية عملية التكوين في مجال تكنولوجيات المعلومات وإدارة المعرفة.

المصادر والمراجع:

- ¹: احمد فايز احمد سيد، أدوات التنقيب عن البيانات مفتوحة المصدر، مجلة الآداب والعلوم الإنسانية، العدد 10، جامعة طيبة، السعودية، 2016، ص: 808
- ²: إياد زعرور، التنقيب في البيانات والتوقعات الاقتصادية، المعهد الوطني للإدارة، لبنان، 18/10/2017، موقع: std.ena.gov.lb/Files/Samples/DataMiningAndForecasting.pdf
- ³: حسين أحمد حسين علي، نظم المعلومات المحاسبية الإطار الفكري والنظم التطبيقية، الدار الجامعية للنشر والتوزيع، مصر، 2004، ص: 24
- ⁴: احمد فايز احمد سيد، مرجع سبق ذكره، ص: 807
- ⁵: سيف الدين عثمان فتوح، الشفيع جعفر محمود، التنقيب في البيانات واتخاذ القرار، مجلة النيل الأبيض للدراسات والبحوث، العدد 3، كلية النيل الأبيض، السودان، 2014، ص ص: 6-8
- ⁶: أسامة ربيع، ماذا تعرف عن التنقيب في البيانات، مجلة AraGeek، 2017/07/20، موقع: <https://www.arageek.com/tech/2015/07/02/telepathy-is-the-future-of-facebook.html>
- ⁷: عبد الستار العلي، عامر إبراهيم قنديلجي، غسان العمري، المدخل إلى إدارة المعرفة، دار المسيرة للنشر و التوزيع و الطباعة، الطبعة الأولى، عمان، 2006، ص: 99
- ⁸: وهيبه داسي، دور إدارة المعرفة في تحقيق ميزة تنافسية، مجلة الباحث، العدد 11، جامعة ورقلة، الجزائر، 2012، ص: 169
- ⁹: محمد احمد إسماعيل، مفهوم إدارة المعرفة في العصر الحديث، المنتدى العربي لإدارة الموارد البشرية، 2017/07/22، موقع: <https://hrdiscussion.com/hr4521.html>
- ¹⁰: Usoro A, Sharratt MW, Tsui E, Shekhar S, Trust as an antecedent to knowledge sharing in virtual communities of practice, Knowledge Management Research & Practice, V5, NO3, 2007.
- ¹¹: PAN Xing, WANG Jun, LIU Lu, Knowledge sharing model based on concept clustering, Systems Engineering-Theory & Practice, V27, NO2, 2007.
- ¹²: حسين على عبد الرسول، العلاقات بين تقنية المعلومات والاتصالات ومستوى نقل واكتساب المعرفة، مجلة الغري للعلوم الاقتصادية، المجلد 02، عدد 14، جامعة الكوفة، 2009.
- ¹³: عبد السلام نجادات، واقع نقل المعرفة والتشارك بها في ظل العولمة، المؤتمر العلمي الدولي حول عولمة الإدارة في عصر المعرفة، أيام 15 – 17 ديسمبر 2012، جامعة الجنان، لبنان، 2012.
- ¹⁴: Lyman P, Varian H, How Much Information? 2003, UC Berkeley, 2003, p :1
- ¹⁵: خضر مصباح إسماعيل طيطي، إدارة المعرفة التحديات والتقنيات والحلول، دار الحاد للنشر والتوزيع، الأردن، 2010، ص: 184