تكنولوجيا المعلومات و البيانات الضخمة (تطبيقات نظم المعلومات، أدوات ذكاء الأعمال BI و الحوسبة السحابية) دراسة تجارب.

Information technology and Big Data (Application of information systems, business intelligence and cloud computing)

ط د فريحة سلمى

جامعة الجزائر 3 friha.salma@univ-alger3.dz / frihasalma2017@gmail.com

تاريخ الإستلام: .../.../..... تاريخ قبول النشر: .../.../..... تاريخ النشر: ../.../.....

الملخص:

تعتبر البيانات الضخمة من أهم المواضيع البحثية في عالم اليوم، حيث تساهم في جودة اتخاذ القرار وتشكيل المعرفة في مختلف مجالات الحياة، إذ تمثل مرحلة هامة من مراحل تطور نظم وتقنية المعلومات والاتصالات التي من شأنها تسهم في التنمية الاقتصادية والاجتماعية، ومما لا شك فيه أن إمكانية الاستفادة من بيانات موجودة بالفعل في مؤسسة ما على نحو أفضل يمثل ميزة كبيرة لأي مؤسسة، إلا أن تقنيات معالجة البيانات الضخمة تصحبها تحديات تتعلق بالأمن والخصوصية، ورغم ذلك هناك العديد من الآليات التي تساعد البيانات الضخمة في تطويرها وتمكينها من توفير خدمات أفضل في المستقبل وتحسين عمليات اليومية ومواجهة التحديات والكوارث الطبيعية وخلق فرص عمل بشكل أفضل المتمثلة في تطبيقات نظم المعلومات وأدوات ذكاء الأعمال BI و الحوسبة الحسابية التي توفر العديد من التقنيات المرنة و تقنية معاجلة كمية هائلة من البيانات.

الكلمات المفتاحية: البيانات الضخمة، معالجة البيانات، تطبيقات نظم المعلومات، ذكاء الأعمال BI، الحوسبة السحابية.

Abstract:

Big data is one of the most important research topics in the world nowadays, It contribute to the quality of decision making and the development of knowledge in the various fields of life, as it represents an important stage in the development of information technology and communications systems that would contribute to the economic and social development, There is no doubt that the possibility of making use data already exit in an institution better represents a big advantage to any institution, But the big data processing techniques accompanied by challenges related to security and privacy, security and privacy, Although, there are many areas that Big data helps in their development and it enable them to provide better services in the future to improve the daily operations and challenges, nature disasters and the creation of employment opportunities for the better represented

by: Information systems applications, business intelligence tools, and computational computing that provide many flexible techniques and technology for processing a huge amount of data.
Keywords: Big data, Information systems applications, business intelligence tools, and computational computing.

1-المقدمة:

لقد أصبح الكم الهائل من البيانات التي يجري إنتاجها وتخزينها والعمل على إتاحتها من مواقع متعددة مصدر قوة رئيسي لأي مجتمع قائم على المعرفة، فهذه البيانات الضخمة تمثل مرحلة هامة من مراحل تطور نظم وتقنية المعلومات والاتصالات التي من شأنها في حال إدارتها على نحو صحيح أن تسهم إسهاما مؤثرا في التنمية الاقتصادية والاجتماعية المستمرة.

تعبر البيانات الضخمة في مفهومها البسيط عن كمية هائلة من البيانات المعقدة التي يفوق حجمها قدرة البرمجيات والآليات الحاسوبية التقليدية على فرزها ومعالجتها وتوزيعها، الأمر الذي حذا بالأخصائيين إلى وضع حلول بديلة متطورة تمكن من التحكم في تدفقها والسيطرة عليها، كما تمتلك تقنية البيانات الضخمة إمكانية تحليل بيانات مواقع الانترنت وأجهزة الاستشعار وبيانات شبكات التواصل الاجتماعي حيث أن تحليل هذه البيانات يسمح باستكشاف ارتباطات بين مجموعة من البيانات المستقلة كشف جوانب عديدة منها على سبيل المثال الاتجاهات التجارية للشركات ومكافحة الجريمة، كما توفر هذه التنبؤات لصانع القرار أدوات مبتكرة لفهم أفضل للظروف والمعطيات وبالتالي اتخاذ قرارات صحيحة تحقق الأهداف المطلوبة.

لتبرز الإشكالية التالية: **ماهي البيانات الضخمة؟ وماهي آليات استغلالها وتطبيقاتها؟**

منهجية الدراسة:

. للإجابة على الإشكالية سيتم تقسيم البحث إلى محاور متمثلة في:

- إبراز ماهية البيانات الضخمة،
- إبراز مختلف آليات تكنولوجيا المعلومات لاستغلال البيانات الضخمة المتمثلة في: تطبيقات نظم المعلومات ذكاء الأعمال وأدواته و الحوسبة الحسابية،
- دراسة تجارب تحليلات البيانات الضخمة في عمليات التصنيع

أولا: ماهية البيانات الضخمة:

1- إطار مفاهيمي حول البيانات الضخمة

قبل التعرض إلى مفهوم البيانات الضخمة وما يتعلق بها لابد من التطرق إلى مفهوم البيانات و تصنيفها وتوضيح اختلافها عن المعلومات.

1-1مفهوم البيانات:

هي صفات وأرقام مشوشة ومزدحمة وغير مرئية ولا يمكن استخراج أي حكمة منها أو قاعدة قبل أن يتم معالجتها، وبالتالي هي تشكل الصورة الخام للمعلومات قبل عمليات الفرز والترتيب والمعالجة، فمثلا لو كان لديك عشرة أشخاص وقمت بقياس أطوالهم وسجلتها على ورقة فهذه الورقة تحوي بيانات، كما تعبر البيانات عن المدخلات إلى جهاز الكمبيوتر بهدف تشغيلها ومعالجتها داخل الجهاز والحصول على المخرجات في صورة معلومات والمعلومات هي مخرجات أي عملية معالجة للبيانات الخام بمعنى لو أخذت أطوال هؤلاء الأشخاص العشرة وقمت بحساب المتوسط الحسابي فهذا الأخير يمثل معلومة لأنه يعطي مقياس مفيد بينما البيانات مجرد أرقام مسجلة على ورقة. [1]

– **تصنيف البيانات** :تصنف البيانات الخام إلى ثلاثة أنواع كما يلي[2]:

–البيانات المهيكلة: وهي البيانات المنظمة في جداول أو قواعد بيانات حيث تتبع مخطط قاعدة بيانات محددة مسبقا، وتكون على شكل حقول مثبتة ومنسقة في سجل أو ملف ضمن قواعد البيانات العلائقية التي يتم فيها استخدام الجداول لتنظيم المعلومات في صفوف وأعمدة، وتحتوي هذه الجداول على أنواع البيانات كالأرقام والتواريخ والعناوين وذلك على نحو سلسلة من الأعمدة، والمخطط يحدد كيفية تخزين البيانات والتحكم في الوصول إليها وسهولة معالجتها، وهذا يكون استنادا إلى طريقة المخطط عند الكتابة

–البيانات غير المهيكلة: تمثل النسبة الأكبر من البيانات، وهي البيانات التي يتم الحصول عليها يوميا من كتابات نصية وصور وفيديو ورسائل، كما تعتبر المشاركات المكتوبة في مواقع التواصل الاجتماعي والمدونات الإلكترونية ورسائل البريد الإلكتروني مصدرا غنيا بالمعلومات التي تنمو بوتيرة متسارعة وغير مسبوقة، مما يعني أنه لا يمكن استخدام أدوات تحليل البيانات القديمة وإدارة قواعد البيانات التقليدية مع البيانات الحديثة التي تأتي من تحديثات الحالة على فيسبوك مثلا أو تغريدات على تويتر أو منصات التدوين لأنها ليست منتظمة وفق الهيكل الذي يتعامل معها كالجداول، مما يجعل من البحث والتحليل أكثر صعوبة بكثير من البيانات المهيكلة، إلا أنه لابد من الاهتمام بهذه البيانات بالرغم من انها غير منظمة

[1]Cloud computing ksu.word press.com/05/03/2012

[2] شلاح رامي ، نبذة عن Big Data موقع على : -https://3alam.pro/rami

–البيانات شبه المهيكلة: تعد نوعا من البيانات المهيكلة إلا أنها لا تكون في صورة جداول أو قواعد بيانات، حيث تكون على شكل نص في صفحة ويب كالملصقات التي تحتوي على معلومات حول عروض فنية و مكانها و زمانها وسعر التذاكر الخاص بالإضافة إلى مكان بيع التذاكر وغيرها، حيث يكون لهذه البيانات هيكل إلا انه قابل للتغيير بشكل سريع وغير متوقع

(Akkok, Keller 2019)هذا يطلق عليها اسم البيانات شبه المهيكلة

2–1 مفهوم البيانات الضخمة وتقنياتها:

يعرف الخبراء البيانات الضخمة بأنها أي مجموعة من البيانات التي هي بحجم يفوق قدرة معالجتها باستخدام أدوات قواعد البيانات التقليدية من التقاط ومشاركة ونقل وتخزين وإدارة وتحليل في غضون فترة زمنية مقبولة لتلك البيانات، ومن وجهة نظر مقدمي الخدمات هي الأدوات والعمليات التي تحتاجها المنظمات للتعامل مع كمية كبيرة من البيانات لغرض التحليل [1].

كما تعرفها المنظمة الدولية للمعايير بأنها مجموعة أو مجموعات من البيانات التي لها خصائصها الفريدة كالحجم، السرعة، التنوع، التباين، صحة البيانات ... وغيرها، والتي لايمكن معالجتها بكفاءة باستخدام التكنولوجيا الحالية والتقليدية للتحقيق الاستفادة منها.

2– خصائص البيانات الضخمة:

الكثير يعتقد أن خصائص البيانات الضخمة تصنف وفقا للحجم فقط لكنها في الحقيقة تصنف وفقا لمبدأ (4V's) ويتكون من velocity، variety، volume ،veracity بمعنى الحجم، التنوع، السرعة والموثوقية[2]

1– 2 الحجم: هي حجم البيانات المستخرجة من مصدر ما، وهو ما يحدد قيمة وحجم البيانات لكي تصنف من ضمن البيانات الضخمة، وبحلول العام 2020 سيحتوي الفضاء الالكتروني على مايقرب 40.000 ميتابايت من البيانات الجاهزة للتحليل واستخلاص المعلومات.

2–2 التنوع: يقصد تنوع البيانات المستخرجة والتي تساعد المستخدمين سواء كانوا باحثين أو محللين على اختيار البيانات المناسبة لمجال بحثهم، وتتضمن بيانات مهيكلة في قواعد بيانات، وبيانات غير مهيكلة كالصور ومقاطع وتسجيلات الصوت –والفيديو والرسائل القصيرة وسجلات المكالمات وبيانات الخرائط، وتتطلب وقتا وجهدا لتهيئتها في شكل مناسب للتجهيز والتحليل.

[1]Dinesh priyankara , Processing unstructured data, 2015. Web site: http://devday.lk/2015/assets/ProcessingUnstructuredData.pdf

[2] هيام حايك، تحليل البيانات الكبيرة في مؤسسات التعليم العالي، على موقع: http://blog.naseej.com

2-3 السرعة: يقصد بها سرعة إنتاج واستخراج البيانات لتغطية الطلب عليها حيث تعتبر السرعة عنصرا حاسما في اتخاذ القرار بناءا على هذه البيانات، وهو الوقت الذي نستغرقه من لحظة وصول هذه البيانات إلى لحظة الخروج بقرار بناءا عليها.

2-4 الموثوقية: و يقصدبها موثوقية مصدر البيانات، ومدى دقتها وصحتها وحداثتها، حيث أن هناك مدير تنفيذي من بين كل ثلاثة مدراء لا يثقون في البيانات التي تعرض عليهم لاتخاذ القرار.

3- مصادر البيانات الضخمة: تتمثل فيما يلي[1]:

3-1 المصادر الناشئة عن إدارة أحد البرامج: سواء كان برنامج حكومي أو غير حكومي، كالسجلات الطبية الالكترونية وزيارة المستشفيات وسجلات التأمين والسجلات المصرفية وبنوك الطعام.

3-2 المصادر التجارية أو ذات الصلة بالمعاملات: الناشئة عن معاملات بين كيانين على سبيل المثال معاملات البطاقات الائتمانية والمعاملات التي تجري عن طريق الانترنت بوسائل منها الأجهزة المحمولة.

3-3 مصادر شبكات أجهزة الاستشعار: على سبيل المثال التصوير بالأقمار الصناعية وأجهزة استشعار الطرق وأجهزة استشعار المناخ.

3-4 مصادر أجهزة التتبع: على سبيل المثال تتبع البيانات المستمدة من الهواتف المحمولة والنظام العالمي لتحديد المواقع.

3-5 مصادر البيانات السلوكية :على سبيل المثال مرات البحث على الانترنت عن منتج أو خدمة ما أو أي نوع آخر من المعلومات، ومرات مشاهدة إحدى الصفحات على الانترنت.

3-6 مصادر البيانات المتعلقة بالآراء: على سبيل المثال التعليقات على وسائل التواصل الاجتماعي

4- الأطراف في منظومة البيانات الضخمة:

لكي يتم تنظيم أي خدمة يجب تحديد الأطراف التي تتعامل مع هذه الخدمة وتحديد واجبات وحقوق كل طرف، تتكون منظومة البيانات الضخمة من عدة جهات تتفاعل فيما بينها، وهذه المنظومة مكونة من موفر البيانات الضخمة، ومقدم خدمة البيانات الضخمة، وعميل خدمة البيانات الضخمة يمكن توضيحها في ما يلي[2]:

[1] Global Pulse, Big Data for Development Challenges and Opportunities, New York, May 2012, p24.

[2] محمد حبش، لمحة عن البيانات الضخمة، على موقع: https://www.tech-

4-1 موفر البيانات الضخمة: يعمل موفر البيانات الضخمة على توفير البيانات من مصادر مختلفة إلى مقدم الخدمة، وتشمل أنشطة موفري البيانات على سبيل المثال إنشاء البيانات والمعلومات الوصفية التي تصف مصدر البيانات وإيجاد مصادر البيانات المفتوحة على الانترنت وتوفير بيان الخدمات إلى مقدم الخدمة عن البيانات القابلة للاستخدام.

4-2 مقدم خدمة البيانات الضخمة : يقوم مقدم الخدمة بتحليل البيانات الضخمة وتوفير البنية التحتية اللازمة لها، وتشمل أنشطة مقدم الخدمة على سبيل المثال البحث في مصادر البيانات وجمع البيانات عن طريق الطلب المباشر من موفر البيانات أو البحث في الانترنت، كما تشمل أنشطة مقدم الخدمة تخزين البيانات ودمجها وتوفير الأدوات لتحليلها ودعم إدارتها مثل خصوصية البيانات وأمنها وملكيتها وغيرها...

4-3 عميل خدمة البيانات الضخمة: وهو المستخدم النهائي لمنظومة البيانات الضخمة أو هو نظاميستخدم النتائج أو الخدمات التي يقدمها مقدم خدمة البيانات الضخمة، كما يمكن للعميل أن ينتج خدمات جديدة أو معرفة اعتمادا على نتائج تحليل البيانات الضخمة، تشمل أنشطة العميل على سبيل المثال طلب خدمة البيانات الضخمة من مزود الخدمة واستخدام مخرجات خدمة البيانات الضخمة في النشاط الذي يرغبه ويتخصص فيه.

5- تطبيقات البيانات الضخمة:

يمكن الاستفادة من البيانات الضخمة في العديد من مجالات الحياة اليومية[1]:

1-5المجال التقني: حيث عبرت العديد من الشركات التقنية عن رضاها نتيجة استخدام واستغلال البيانات الضخمة، فإلى جانب الشركات التقنية مثل غوغل، تويتر، فيسبوك، أمازون هناك العديد من الشركات أو المؤسسات التي تستخدم هذه التكنولوجيا الجديدة مثل المؤسسات التي تعنى بالسفر والفنادق فالحصول على رضا العميل أمر مهم لكن قياسه صعب إلا أن تحليل البيانات الضخمة يعطي هذه الشركات القدرة على جمع بيانات العملاء وتطبيقها على الفور فمن بين أولى الشركات التي كانت لها تجربة مباشرة في معالجة البيانات الضخمة هي شركة غوغل، حيث بدأت في عام 2007 بجمع البيانات من مئات المصادر بما في ذلك البيانات الموجودة في قواعد البيانات والتنبؤات الجوية والبيانات التاريخية والمالية وبيانات السفر وأسعار الصرف، وتطور البحث عن البيانات عام 2012ليشمل الرسم البياني الذي يعرض معلومات حول موضوع البحث إضافة إلى البيانات الشخصية

[1] Naci Akkok, Pal Halvorsen,Arthur M.Keller, Vera Geobel, Semistructured Data and XML, union 2006.

لكل مستخدم Gmail و +Google أجل التوصل إلى أفضل تخمين فيما كان يبحث عنه المستخدم وفي عام 2010 أطلقت الشركة إعلانا شاملا عن الخدمة التجارية من أجل السماح للشركات بتخزين وتحليل مجموعات كبيرة من البيانات على منصات الحوسبة السحابية، ومشروع آخر إلا وهو السيارة ذاتية القيادة التي تولد كميات هائلة من البيانات من أجهزة الاستشعار والكاميرات واستخدامها وتحليلها من أجل التنبؤات المستقبلية.

2-5مجال الخدمات الحكومية: حيث أصبح بإمكان المختصين بتطوير الخدمات الحكومية رصد مدى رضا المواطنين عن الخدمات الحكومية المقدمة لهم، وعلى ضوء النتائج المحللة يمكن استنتاج ما يلزم عمله للتطوير والتحسين، حيث بات مسح آراء الجمهور عن طريق الاستبيانات التقليدية مكلفا وغير مجد في كثير من الأحيان، وذلك نظرا لتنوع البيانات الديموغرافية وثقافات المتعاملين وحتى المساحات الشاسعة المراد مسحها واستقصاؤها، ومن أكبر المصادر لتلك البيانات الضخمة هي البيانات المسجلة من خلال
عمليات التعداد السكاني والتسجيل في قواعد البيانات الحكومية حيث يمكن استنتاج معلومات مهمة جدا من خلال تحليل تلك البيانات المخزنة.

3-5المجال الاقتصادي: أتاحت البيانات الضخمة لمؤسسات الأعمال زيادة أرباحها بشكل كبير وتعزيز وضعها التنافسي،فمعرفة رغبات الزبائن وميولهم ونفسياتهم يتيح للشركات الربحية توفير منتجات وخدمات بناءا على ذلك، وعليه تضمن تلك الشركات رضا عملائها مما يؤدي إلى زيادة مبيعاتها. ومن أبرز الأدلة على استخدام شركات التجزئة والخدمات لتلك البيانات الضخمة لزيادة أرباحها هو توجيه الإعلانات الترويجية التي تلاحق متصفحي الانترنت والمواقع الالكترونية، فهناك تطبيقات ومواقع التي تحلل رغبات المستخدم من خلال AdSense متخصصة بتحليل تلك البيانات مثل خدمة غوغل المدخلات التي يقوم بإدخالها سابقا في محركات البحث، فتقوم تلك المواقع رصد الكلمات وتخزينها في قواعد بيانات خاصة ليتم تحليلها وتوجيه الدعايات التي تتناسب مع هذه الاهتمامات والرغبات.

4-5مجال عملية صنع القرار:إن نجاح المؤسسة أو القطاع الحكومي يتوقف إلى حد كبير على قدرة وكفاءة القيادة الإدارية على اتخاذ القرارات الإدارية المناسبة، فعملية صنع القرار تبدأ بتجميع البيانات ومعالجتها واستخلاص المعلومات التي يتم اتخاذ القرار بناءا عليها، إذ تعتمد العديد من الشركات الكبيرة والقطاعات الحكومية على سياسة تحليل البيانات الضخمة والمعقدة التي تحتاج إلى البرمجيات المتخصصة في مجال إدارة البيانات وتحليلها من أجل تحسين العمليات الداخلية مثل إدارة المخاطر وإدارة علاقات العملاء

والخدمات اللوجستية، وتحسين المنتجات والخدمات القائمة والاستفادة من المعلومات وتقديم العروض المناسبة للعملاء في الوقت المناسب .

5-5المجال الطبي: أتاح التطور التكنولوجي المتزايد إمكانية تحليل الكم الهائل من المعلومات المتوفرة على الانترنت ما أدى إلى ابتكار تقنيات حيوية خاصة في مجال الطب والرعاية الصحية، فقد تحول البشر من طب التعامل مع الأمراض إلى طب يسعى للتنبؤ بالأمراض ومنعها وتقديم العلاج الملائم لكل شخص بمساعدة هاتفه الذكي ما يفتح الباب لعصر جديد من الطب تلعب فيه البيانات الصحية الضخمة و تحليلاتها دورا بارزا.

ثانيا: مختلف آليات استغلال و تحليل البيانات

1- تطبيقات نظم المعلومات القائمة على تكنولوجيا المعلومات:

المؤسسة و في إطار نشاطاتها تتعامل بالمعلومة ، ومع بروز أهميتها اوجب عليها إيجاد سبل لإدارتها بواسطة نظم تستخدم البرمجيات و التطبيقات الخاصة بنظم المعلومات في تحليل و ترتيب و إدارة البيانات الضخمة و معالجتها بواسطة سيرورات تم تحديدها مسبقا و آلية سير مبنية على مداخلات و مخرجات و تفاعلات و ترابطات و معايير و مؤشرات للتحكم في المعلومة الضرورية و التي تستخدم تطبيقات قائمة على تكنولوجيا المعلومات لتدير البنية التقنية للمعلومة [1]

و بالتالي تطبيقات نظم المعلومات هي عبارة عن برمجيات و سيرورات تكنولوجية لإدارة المعلومة يمكن تقسيمها إلى تطبيقات داخلية و تطبيقات ممتدة ، حيت تتكون التطبيقات الداخلية من :

1-2 برمجية إدارة الموارد المالية FRM[2] : و التي تعد تطبيق داعم للإدارة المالية للمؤسسة و تهدف إلى تسيير و متابعة مواردها و تحقيق الامتلية في استغلالها.

1-3 برمجية إدارة الموارد البشرية HRM[3]:و هي برمجية تهدف إلى التسيير الإداري للمعلومات حول المرد البشري و تقييم أداءه و تدريبه و مساره المهني .

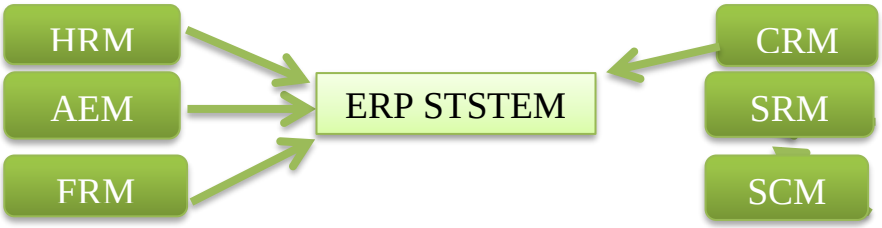
1-4 برمجية دورة حياة المنتج PLM[4].

1-5 برمجية إدارة أصول المؤسسة AEM[5].

[1] د .كاريش صليحة ، محاضرات مقياس تطبيقات نظم المعلومات ، تخصص ادارة اعمال و انظمة المعلومات كلية العلوم الاقتصادية و العلوم التجارية و علوم التسيير جامعة الجزائر3 ، 2016

[2] FRM : FINANCIAL RELATIONSHIP MANAGEMENT.
[3] HRM : HUMAN RESOURCES MANAGEMENT.
[4]PLM : PRODUCT LIFECYCLE MANAGEMENT.
[5] AEM : ENTRERISE ASSET MANAGEMENT.

1-6 إدارة تخطيط موارد المؤسسة ERP[1] : و الذي يهدف إلى الانتقال من التسيير الجزئي إلى التسيير المتكامل للمؤسسة و هو عبارة عن نظام جزئي من نظام المعلومات قادر على تحقيق التسيير الكلي للمؤسسة و تجميع وظائف التسيير في مرجعية واحدة و تحقيق التنسيق بين الوظائف و البعد الأفقي لها و مع بروز ذكاء الأعمال أصبح بالإمكان طرحه في الشبكة على شكل سحابة الكترونية



التطبيقات الممتدة والمتمثلة في:

- إدارة العلاقة مع المورد SRM[2] : و هي عبارة عن برمجية تستخدم من طرف المؤسسة و التي تهدف من خلالها الى تحسين ميكانيزمات التموين و مختلف العلاقات مع الموردين بتحسين عملية الاتصال بهم و اختيار الأفضل بينهم .

- إدارة العلاقة مع العملاء CRM : هو عبارة عن نظام برمجي يدير العلاقات مع العملاء لإدارة تفاعل المؤسسة مع عملائها الحاليين و المستقبليين و يعمل على تحليل بيانات تاريخ العملاء مع المؤسسة من اجل تحسين أفضل علاقات تجارية مع العملاء، الذي يعد على انه إدماج تكنولوجي لسيرورات مرتبطة بالبيع و التسويق تبدأ من الزبون وصولا إلى خدمات ما بعد البيع ، و التي تعد على أنها برمجية و استعمال تكنولوجي من اجل التحليل الدقيق للأسواق و توقعات الزبائن و شخصنة التفاعلات معهم personalization وصولا إلى كسب ولائهم و زيادة عدد الزبائن المحتملين procpect و جعلهم زبائن فعليين و دائمين.حيث تعمل البرمجية على تحليل بياناتهم و تجميعها من مختلف المنافذ من اجل تحسين العلاقات معهم و بلوغ رضاهم.

- إدارة علاقات العملاء الالكترونية CRM-E : لقد تطور CRM ليشمل[3] E-CRM و الذي يمثل قفزة نوعية في مجال التسويق و برزت الحاجة إليه نظرا لتنافسية الأسواق و انفتاحها و السلوك الجديد للمستهلك الذي أصبح واعيا بسبب انتشار المعلومة و توفرها في العالم الافتراضي ، و بالتالي أصبح موضع اهتمام و تم تطوير تقنيات و أنظمة لجذبه و المحافظة عليه و شخصنة العلاقة معه و اعتباره شريكا في تعزيز الموقع

[1] ERP : ENTREPRISE RESOURCES MANAGEMENT.
[2]SRM : SUPPLIER RELATIONSHIP MANAGEMENT.
[3] ICM : INTERN CHAIN MANAGEMENT

التنافسي خاصة مع انتشار ثقافة التسويق الرقمي ، و بالتالي وجب توفير وسائل قادرة على إدارة المعلومة الخاصة بالزبائن انطلاقا من عدة منافذ و تخزينها بشكل مرتب الأمر الذي يوفره ذكاء الأعمال BI[1].

هي تطبيق للتقنيات المستندة إلى الانترنت مثل رسائل البريد الالكتروني و المواقع الالكترونية و غرف الدردشة و المنتديات و القنوات الأخرى لتحقيق أهداف CRM، و هي عملية منظمة و منسقة بشكل جيد لإدارة علاقات العملاء تقوم بأتمتة العمليات في التسويق و المبيعات و خدمة العملاء

– إدارة سلسلة التوريد SCM[2] : و هي عبارة عن تكامل تكنولوجي و إدماج لمجموعة من السيرورات انطلاقا من مرحلة التوريد وصولا إلى الزبون النهائي و التي تتكون من ثلاث سيرورات و هي :CRM ، SRM ، CRM-E.

2– ذكاء الأعمال مع الوقوف على أهمية تقنية التنقيب عن البيانات

هذه المعلومات التي تتميز بالضخامة كونها تترتب عن شبكة عالمية و عن زبائن مختلفين باختلاف أذواقهم و رغباتهم و تعقيدها ، و كون المؤسسات تسعى إلى شخصنة تعاملاتها كسبيل لكسب الولاء استوجب توفر أنظمة لتخزين المعلومات و معالجتها بطريقة تكنولوجية خاصة في عصر البيانات الضحمة[3] بتوفير حل للتخزين و الاسترجاع و بسرعة و امن و ذلك عن طريق قاعدة تسمى بمستودع البيانات DATAWERHOUSEو هي عبارة عن قاعدة بيانات ضخمة ترحل إليها البيانات من القواعد التشغيلية على غرار CRM بواسطة ETL[4] و هي أداة تسمح باستخراج و تحويل و شحن البيانات من القواعد التشغيلية إلى مستودع البياناتDW[5] الذي يحتوي على معلومات معالجة، مؤرخة ،غير قابلة للتعديل ،مرتبة حسب محاور التحليل ، موجهة و مساندة لاتخاذ القرار و الذي يقوم بتحليلها على عدة اتجاهات أي فهمها بشكل متعدد الإبعاد على شكل مكعب.

ليتعدى الأمر الى مرحلة الاستغلال و تحديد أدوات ذكاء الأعمال حيث يعد ذكاء الأعمال على انه مجموعة من الطرق و الوسائل و الأدوات تساعد على اتخاذ القرار و يستخدم في ذلك أدوات متمثلة في:

. المعالجة الفورية التحليلية للبياناتOLAP: و التي تمثل تقنية للتمثيل المجرد للمعلومة في شكل مكعب متعدد الاتجاهات و الأبعاد لتنظيم و تحليل المعلومات.

. **تقنية التنقيب عن البيانات:** تعد أهم تقنية موجهة لتحليل و اكتشاف البيانات الضمنية غير الظاهرة المخزنة في قواعد البيانات الضخمة الهدف منها اكتشاف و تحليل القواعد و العلاقات و النماذج و الاتجاهات الغير

[1] BI : BUSINESS INTELLIGENCE.
[2] SCM : SUPPLY CHAIN MANAGEMENT.
[3] BD : BIG DATA.
[4] ETL : EXTRACT , TRANSFORM , LOUAD.
[5] DW : DATA WERHOUSE.

مضبوطة مسبقا و وضعها في هيكل يضم معلومات ضرورية حيت يسمح بتحليل كميات ضخمة من البيانات و تحويلها من مجرد معلومات متراكمة و غير مفهومة إلى معلومات قيمة يمكن استغلالها لاتخاذ القرار حيت تعد أداة ابتكاريه كونها تسمح باكتشاف معارف جديدة و توليد المعرفة و بناء توقعات مستقبلية ، و تقوم على مراحل متمثلة في:

. التصنيف classification: اي تحديد المعلومات المستهدفة التي لها أفضل عوائد .

. التقدير: نمدجة السلوك المدروس اي سلوكيات المستهلكين من خلال علامات و نقاط .

. التجزئة:**segmentation** وضع العملاء في مجموعات متجانسة و كأداء الأسواق و المنافسين عن طريق مثلا cluster.

. الترابط : تحديد و تنشيط خطوط الإنتاج.

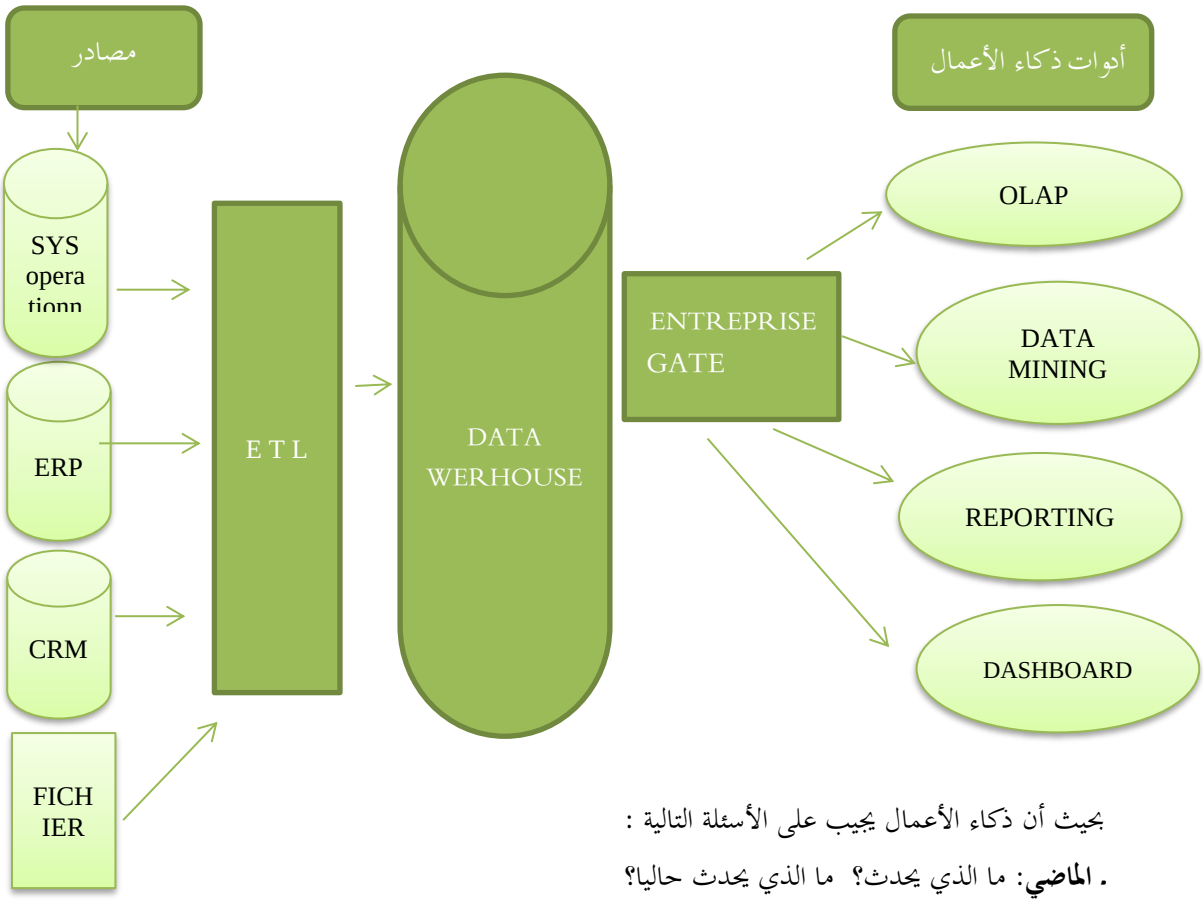
. التوقع :توقع حجم المتغيرات على المدى البعيد.

. التقاريرreporting: تعتمد التقنية على أشكال معدة مسبقا و تنبؤات عن طريق SQM .

. لوحة القيادةdashboard : تحتوي على معلومات و مؤشرات تستجيب لأهداف المؤسسة.

و يمكن تلخيص سيرورة ذكاء الأعمال في الشكل الأتي[1]:

[1] نسيم قائد عبده ، إطار مقترح لتطبيق نظام دكاء الأعمال في التسويق جامعة تعز الجمهورية اليمنية الجمعية العلمية للمعلوماتية.



بحيث أن ذكاء الأعمال يجيب على الأسئلة التالية :

. الماضي: ما الذي يحدث؟ ما الذي يحدث حاليا؟

. الحاضر: ما الذي قد يحدث؟ لمادا يحدث دلك ؟

. المستقبل: ما الذي أريده فعلا أن يحدث؟

و منه يمكن القول إن ذكاء العمال يوفر خاصية هائلة في إدارة المعلومة وتوليد معارف و بناء توقعات مستقبلية قادرة على استغلال الكم الهائل من المعلومات التي تتيحها السوق الفعلي و السوق الافتراضي.

3– الحوسبة السحابية و تحليل البيانات الضخمة

3–1 مفهوم الحوسبة السحابية:

هي تكنولوجيا تعتمد على نقل المعالجة ومساحة التخزين الخاصة بالحاسوب إلى ما يسمى السحابة و هي جهاز خادم يتم الوصول إليه عن طريق الانترنت. بهذا تتحول برامج تكنولوجيا المعلومات من منتجات إلى خدمات، و تعتمد البنية التحتية للحوسبة السحابية على مراكز البيانات المتطورة والتي تقدم مساحات تخزين كبيرة للمستخدمين كما أنها توفر بعض البرامج كخدمات للمستخدمين. و هي تعتمد في ذلك على الإمكانيات التي وفرتها تقنيات ويب 2.0.

لم تتطور تكنولوجيا الحوسبة السحابية في آن واحد، فقد تطورت من خلال عدد من التقنيات المتقدمة مع مرور الوقت، مثل الحوسبة الشبكية، الحوسبة الخدمية المقترحة في s1990، وتكنولوجيا البرمجيات كخدمة Software as Serviceالمقترحة في عام 2001. و أخيرا تم اعتبار تكنولوجيا الحوسبة السحابية لتكون الجيل القادم من حوسبة الانترنت[1].

2-3 تصنيف الحوسبة السحابية:

يمكن تصنيف تطبيقات الحوسبة السحابية وفقا لنوع الخدمة أو أسلوب النشر[2]

- التصنيف حسب نوع الخدمة:

•**البنية التحتية كخدمة Infrastructure as a Service IaaS** تزود الدعم المادي مع التخزين الظاهري و الخوادم. كما تدعم خدمة الشبكة أيضاً. يمكن للمستخدمين استخدام هذه التطبيقات وفقا لاحتياجاتهم. أمثلة Amazon Elastic Cloud Computing (EC2), Google Compute Engine (GCE), Microsoft Azure

•**المنصات كخدمة Platform as a Service PaaS** تدعم الاستضافة حيث يمكن للمستخدمين تطوير تطبيقاتهم بناءً على قواعد معينة. الأمثلة على ذلك محرك تطبيقات جوجل Google's Apps) . Engine.)

•**البرمجيات كخدمة Software as a Service SaaS** توفر التطبيقات التي تسخدم عبر التصفح او التطبيقات التي يمكن استخدامها مباشرة. من الامثلة على ذلك نظام دارة العلاقات مع الزبائن CRM Customer Relationship Manag ent التي توفرها شركة Salesforceو Google Docsو Gmailو. cisco WebEx

- التصنيف حسب طريقة النشر:

لحل مشاكل الأمن والموثوقية والتنظيمية في البيئات السحابية تم تصنيفها إلى ثلاث فئات لتحديد العلاقة بين مقدمي الخدمة والمستخدمين.

•**السحابة العامة** يتم تقاسم السحابة بين مقدمي الخدمات والمستخدمين.

---

[1] شاهين,أحمد(2013).محرك Google Drive واستخداماته التعليمية.مدونة مصمم تعليمي، على الرابط: http://www.id4arab.com/2013/04/google-drive.html

[2] حايك, هيام.(2013) الحوسبة السحابية تغزو مؤسسات التعليم العالي. مدونة نسيج الإلكترونة. تم استرجاعه في تاريخ 25-3-2014على الرابط:/http://blog.naseej.com

•**سحابة خاصة** يتم استخدام السحابة من جانب مقدمي الخدمات. يمكن للمستخدمين الاعضاء فقط تبادل الموارد السحابية.

•**سحابة الهجينة** هو مزيج من الطريقتين السابقة.

3-3 خصائص الحوسبة السحابية:

الحوسبة السحابية تختلف عن الحوسبة التقليدية في حجمها الكبير والعديد من الميزات الأخرى[1] الافتراضية (Virtualization) توفر منصات سحابية بناء على نسخ افتراضية من التطبيقات بدلا من تلك المادية. وهذا يساعد في نقلها عبر البيئات السحابية ومراكز البيانات.

•**الديناميكية (المرونة ((flexibility**) **Dynamic**) توفر المنصات السحابية إمكانية إجراء تغييرات في الحجم بناء على احتياجات المستخدمين.

•**مقياس اقتصادي Economies of scale** نطاق واسع من موارد منصات الحوسبة السحابية يمكنها السيطرة على رسوم الإيجار أو استخدام هذه الموارد مما يؤدي إلى زيادة أو تقليل عدد المستخدمين.

•**الخدمة على الطلب (On d and service**) المنصات السحابية تتيح للمستخدمين تحديد حجم الموارد بناء على احتياجاتهم.

•**امكانية عالية (High reliability**) المنصات السحابية عادة تتيح استخدام النسخ الاحتياطية لزيادة الموثوقية وتتيح استخدام نظم إدارة الشبكات الديناميكية لزيادة الكفاءة والأداء.

3-4 العلاقة بين الحوسبة السحابية و البيانات الضخمة:

إنّ العلاقة بين البيانات الضخمة والحوسبة السحابية هي علاقة تكاملية. الحوسبة السحابية هي نهج التكنولوجيا الجديدة في حين أن البيانات الضخمة هي ظاهرة البيانات ا المتزايدة بطريقة سريعة. تتيح تكنولوجيا البيانات الضخمة للمستخدمين معالجة الاستعلامات الموزعة من مجموعات مختلفة من البيانات والحصول على النتائج في الوقت المناسب . بينما توفر الحوسبة السحابية المحرك الأساسي للقيام بذلك . تم تصميم نظم إدارة قواعد البيانات التقليدية للبنية التحتية للمؤسسات المحلية، وهي تفتقر الى الكثير من الميزات السحابية[2]

[1] سيد،رحاب فايز أحمد.(2013).نظم الحوسبة السحابية مفتوحة المصدر:دراسة تحليلية مقارنة. المجلة العراقية لتكنولوجيا المعلومات ، 5(2) ، 17-41.

[2] Saif and S wazir,Performance analysis of big data and cloud computing technique Asurvery science direct May13 ;2020.

5-3 معالجة البيانات الضخمة في البيئة السحابة

تتطلب تقنيات خاصة حيث إنها تتطلب بنية تحتية مناسبة، وأنظمة تخزين واستعلام خاصة وطرق تحليل خاصة أيضاً تتمثل فيما يلي[1]:

– استخراج البيانات الضخمة في البيئات السحابية:

الخطوة الأولى في تحليل البيانات هو الحصول على البيانات أو استخراجها. بل هو شرط أساسي لمعالجة البيانات. ويشمل الحصول على البيانات خطوات تحضيرية مثل استخراج البيانات وتحويلها وتحميلها (ETL Extraction, Transformation, Loading)). ويشمل تحضيرها أيضاً تصفية البيانات وتنظيفها. في البيئات السحابية هناك أنواع مختلفة من المعلومات مثل **النصوص والصور و URL** التي تأتي من مصادر مختلفة مثل مواقع وسائل الإعلام الاجتماعية، ومحركات البحث والمواقع التجارية، كلها تدخل في عملية استخراج البيانات. **أدوات ETL** تستخدم في تحويل وتنظيف أنواع غير المتجانسة والموزعة وتحميلها على نظام تخزين مناسب (مثل مستودعات البيانات). أدوات ETL للبيانات الضخمة تختلف عن عنها في ETL التقليدية في سرعة إنتاج البيانات وحجم البيانات. يجب أن تتضمن أدوات ETL المستخدمة للبيانات الضخمة اختيار التكنولوجيا الحديثة مثل قواعد البيانات الموزعة والمعالجة في وقت مناسب.

– تخزين البيانات الضخمة مختلفة من البيانات في البيئات السحابية

في الوقت الحاضر هناك الكثير من تطبيقات الويب الفردية التي يمكن أن تولد كمية كبيرة من البيانات كل ثانية، على سبيل المثال في كل يوم يتم تخزين 500 تيرابايت من البيانات في قواعد بيانات الفيسبوك و أكثر من 360 تيرابايت من البيانات يتم تحميلها على موقع يوتيوب. إنّ نمو البيانات في البيئات السحابية يتزايد بشكل كبير و سريع مع ازدياد عد المستخدمين المتصلين بالانترنت حول العالم .مع هذا النمو السريع فإن السؤال الذي يتبادر إلى الذهن هو كيف يمكن تخزين هذه الكميات الهائلة من البيانات في البيئات السحابية؟ نحن بحاجة إلى تقنية التخزين التي تلبي احتياجات النمو السريع للبيانات على السحابة، نحن بحاجة إلى تقنية تخزين مع تكلفة منخفضة ، وموثوقية عالية وامكانية عالية. يمكن تصنيف طرق تخزين البيانات الضخمة في البيئات السحابية الى نوعين:

1– أنظمة الملفات الموزعة:.

تستخدم أنظمة الملفات الموزعة لدعم تخزين كميات كبيرة من الملفات. لحل مشكلة تخزين عدد كبير من صفحات الويب وضعت جوجل نظام تخزين الملفات جوجل (GFS) ولكن مشكلته أنه ليس مفتوح المصدر. ياهو

[1] تي سي ريدمان،تحسين جودة البيانات الضخمة لميزة تنافسية، معهد ماساتشوستس للتكنولوجيا، المراجعة الادارية، المجلد 36الرقم 2 ص 99

والمجتمعات الأخرى وضعت نظام مفتوح المصدر مبني على GFS أسمته Hadoop Distributed File Syst s (HDFS)نظام تخزين الملفات الموزعة HDFS يقوم بتخزين كمية كبيرة من الملفات حيث يتم تقسيمها الى عدة كتل.clusters حيث تتكون كل كتلة من مجموعة من العقد . nodes كتلة Hadoop هي مجموعة من الخوادم و مجموعة HDFS's تقوم بتوزيع و تخزين البيانات عبر كافة العقد في هذه المجموعة. كتلة HDFS تشمل ثلاث عقد

- **عقدة الاسم NameNode** تعمل كعقدة رئيسية و هي التي تسيطر على تخزين المجموعات. إنها تقسم البيانات إلى مجموعات، كل مجموعة يتم نسخها ثلاث مرات وتخزينها في الكتلة.
- **عقدة البيانات DataNode** تعمل على عقد الخوادم و توفر إمكانية السماح بالخطأ. إذا فشلت تضمن النسخ القادمة من NameNode التأكد من أن النظام لن يتوقف او يفشل في هذه الحلة
- **جهاز العميل Client Machine** يمكن أن يكون DataNode أوNameNode ، يكون نظام Hadoop مثبتاً عليه. يقوم بتحميل البيانات في الكتلة، يقدم وظائف الى MapReduceويظهر نتائج المهمة عندما تكتمل.

2- قواعد البيانات الموزعة NoSQL

إنها تُسخدم لتلبية احتياجات تحليل ومعالجة كميات كبيرة من البيانات غير المهيكلة او غير المنظمة .وقد حلت قواعد بيانات NoSQL مشاكل التزامن، والتدرجية، والتسامح مع الخطأ التي واجهت قواعد البيانات العلائقية التقليدية. و مثل قاعدة بيانات SQL التقليدية (لغة الاستعلام الهيكلية) يستخدم NoSQL أيضا الاستعلامات ولكنها لا تتقييد باستخدامها وحدها فقد تستخدم لغات اخرى غير.SQL

هناك أربعة أنواع من NoSQL**مخزن مفتاح القيم Key-Value store**يحتوي على جدول من المفاتيح والقيم (على سبيل المثال(,(Riak, Amazon S3)، و **المخزن القئم على الوثائق document based store**حيث تتكون الوثائق من عناصر مختارة. **و المخزن القائم على العمود Column based store**الذي كل كتلة فيه تحتوي بيانات من عمود واحد فقط (مثل HBase)والمخزن القئم على الرسم البياني Graph based storeحيث تستخدم قاعدة بيانات الشبكة الحواف والعقد لتمثيل وتخزين البيانات مثلNeo4J

لحل مشكلة التعامل مع مجموعة متنوعة من البيانات الضخمة، وضعت جوجل نظام قاعدة بيانات وزعت على أساس المخزن القائم على العمود يسمى BigTable نظام التخزين الموزع للبيانات المنظمة ، معاملات ACMعلى نظم الحاسب الآلي(TOCS) ، يعمل على تقديم تحليل للبيانات الضخمة غير المتناسقة. نموذج

البيانات من BigTable يتكون من الصفوف والأعمدة، والطوابع الزمنية. يتم تخزين البيانات في الخلايا. يتكون القرص من العديد من الصفوف ويتم حفظها إلى عقدة الخادم.

3-6 تحليل البيانات الضخمة في الحوسبة السحابية

ينقسم تحليل البيانات الضخمة في البيئات السحابية الى نوعين رئيسيين[1]:

- النوع الأول batch offlineتتم معالجة مجموعات كبيرة من البيانات تلقائيا دون أي تبادل أو مراقبة من قبل المستخدم. بعد جمع البيانات، تتم معالجة جميع البيانات في وقت واحد. و من الأمثة عليها Hadoop و MapReduce . MapReduce.هو نموذج البرمجة التي وضعتها جوجل على أساس GFS لتجهيز عدد كبير من مجموعات البيانات. ويستخدم مجموعة واسعة النطاق لأداء المهام والعمليات الموازية تلقائيا باستخدام برنامج مثل .Hadoop ويمكن تطبيقها على البحث في مواقع الويب وتطبيقات الويب الأخرى. الخوارزمية MapReduce تشمل وظيفتين وظيفةMap لتحويل العناصر في نوع من قائمة إلى نوع آخر و وضعها مرة أخرى في نفس النوع من القائمة. ثم يتم تنفيذ التصفية وعمليات الفرز. وظيفة Reduceلجمع العناصر الموجودة في القوائم المخرجات من Map وأداء بعض الحسابات على كل منهم للحد منها في قيمة واحدة Hadoop .هو برنامج مفتوح المصدر يعتمد على HDFS و هو يعمل على معالجة البيانات الموزعة معالجة فعالة وموثوق بها لكميات هائلة من البيانات. وقد وضعت Hadoop أساسا عن طريق Yahoo لحل مشكلة البيانات الضخمة عن طريق كسر البيانات إلى قطع صغيرة يمكن معالجتها بشكل متواز. يضم Hadoopاثنين من الأنظمة الرئيسية تخزين HDFS ومحرك.MapReduce

Hadoop هو يستخدم لإدارة والاستعلام عن مجموعات كبيرة من البيانات في تخزين الموزعة. مثل لغة SQL،

- النوع الثاني الوقت الحقيقي real time analysis عندما يتم إدخال البيانات من قبل المستخدم يتم معالجتها فورا ويحتاج المستخدم إلى الانتظار لفترة وجيزة لا تتجاوز ميلي ثانية (مثل نظم الحجز). من الأمثلة على الحوسبة التفاعلية دريميل جوجل.

ثالثا: دراسة تجارب تحليلات البيانات الضخمة في عمليات التصنيع

1- دراسة لويس كولومبوس التحديات العشرة للبيانات الضخمة في التصنيع[2]:

[1] م،نيسر، ذكاء الأعمال كتقنية دعم عملية صنع القرار في المؤسسة الحديثة، مؤتمر ادارة المعرفة يومي 26و28 يونيو 2013

[2] ل.كولومبوس، الطرق العشر للبيانات الضخمة و هو ثروة التصنيع، نوفمبر 2014

ركزت الدراسة على نظام ماكينزي , وتركزت على منشئتين مصنعة تقوم باستخدام البيانات في التحليل لزيادة الإنتاجية وتخفيض التكاليف، كانت نتائج الدراسة بعد أن قام بتحليل بيانات الدراسة على زيادة دقة وجودة الإنتاج.

- الإسراع في تكامل تكنولوجيا المعلومات مع البيانات الضخمة والأنظمة.
- عمليات التنبؤ أصبحت أفضل بنسبة تتعدى (46) للمنتجات من خلال البيانات التي تم توفيرها مثل العرض والطلب.
- أصبحت خدمة ودعم العملاء أسرع
- وضع مستويات الجودة للموردين ودقة في توقع أداء الموردين مع مرور الوقت.
- القدرة على تتبع الأنظمة والأجهزة والأخطاء الناتجة عنها.
- إنتاج وبيع المنتجات التي تحقق أكثر ربحية. اضافه إلى تقديم الخدمة بشكل استراتيجي.

2- دراسة شركة ماكينزي[1]

عملت مع العديد من المنظمات لمساعدتها في تحسين الأداء وتحقيق الأهداف الهامة عن طريق بناء تطبيق الإحصاءات مع الأدوات الرياضية(EXHIBIT) كيف يمكن للبيانات الضخمة تحسين الصناعات التحويلية،يعمل التطبيق من خلال الوصول إلى البيانات الكبيرة والعثور على نمط للعلاقات بينها وباستخدام خطوات عملية منفصلة ومدخلات الأعمال من أجل تقييم وتحسين إجراءات العمل وتعيينها لتصبح متاحة وبذلك يمكن قياسها ومقارنة تأثيرها النسبي لمدخلات الإنتاج المختلفة

يعمل نظام (EXHIBIT) مع أنظمة البيانات الضخمة بشكل متزايد. ويقوم بتحسين العمليات الحالية للأنظمة والمعلومات المتاحة عن طريق مركزة وفهرسة البيانات من مصادر متعددة وإنشاء تقارير للأنشطة عمليات التصنيع وسلسلة التوريد في مراحل الإنتاج، ويطبق تحليل متطورة كجزء من التحسينات العملية وتبسيط المراحل المعقدة بشكل أفضل, وإبراز نقاط القوة للشركة على المنافسة والنجاح والاستناد الى هذه المعلومات لتصبح احد المنافسين بالسوق.

3- دراسة ستيف.ديرتين البنية التحتية للتحليلات البيانات الضخمة[2]

على بناء نظام جديد يدعم البنية التحتية لتحليل البيانات, وقام بشرح البنية التحتية وأهميتها، وكانت النتيجة هي أداة بحث تقوم بدورها بتحليل ومنطق الأعمال، والقدرة على تحليل البيانات الضخمة وخلق قيمة من هذه

[1] ك.هامر، راجاجوبيول كيف يمكن للبيانات الضخمة تحسين الصناعات التحويلية ، ماكينزيو الشركات، 2014

[2] ستيف. ديرتين، تعريف البنية التحتية للتحليلات البيانات الضخمة، ماي 2015

البيانات، وكانت بمثابة النقطة الرئيسية للتعبير عن شروط وضوابط واقعية ينبغي اتخاذ القرارات من خلالها مثل اختيار متطلب تحليل البيانات.

4-دراسة شركة هيتاشي لأنظمة المعلومات[1]

بتقديم منتجات وخدمات مثل **Pentaho لتحليل الأعمال و Pentahoلتكامل البيانات**. وقامت الشركة بإنشاء أداة تدعى (**EDO**)لتحسين التعامل مع البيانات المخزنة والموجودة وذلك لزيادة ولاء واستهداف العملاء للمنتجات.

قدمت الشركة (EDO) كدراسة خاصة لحل المشاكل القائمة. وتركزت الدراسة على تحليل البيانات المتوفرة لديها والتي كانت احدى مشاكلها والتي تمثلت بتقليل الوقت المستغرق لعملية تحليل البيانات من خلال عملية تشغيل ETL(extract, transform and load) والتي كانت بحاجة الى وقت ما يقارب 30 ساعة. وكان الهدف الاخر من تقديم هذه الاداة هو تحسين كفاءة معالجة البيانات، وايجاد احد الحلول للتعامل مع البيانات كبيرة بشكل يومي، إضافة الى زيادة كفاة الانظمة في عملية استخراج البيانات، وتحويل تلك البيانات وتوفيرها لانظمة الاستعلام.

نتائج الدراسة التي قدمتها شركةPentaho :

- تقليل من الوقت المستغرق لعملية تشغيل ETL لتكامل البيانات بنسبة 70.
- الاستعلام عن البيانات أصبحت متاحة، مما يسمح للمحللين والمدراء اتخاذ قرارات مستنيرة بشأن الإستراتيجية بشكل أسرع.
- القدرة على إنشاء تقارير جديدة ومخصصة بالوقت المناسب, تحسين جودة البيانات وتحسين في اداء التكاليف التي يتم تحمليها خلال عملية الإنتاج.
- تقليل الوقت المستغرق لإصلاح الأخطاء الموجودة في البيانات.

5-دراسة شركة PLEXفي مجال تحليل البيانات الكبرى في التصنيع [2]

توفر الشركة العديد الأنظمة التي تدعم التصنيع في مجال التكنولوجيا من خلال انظمتها تقوم بربط الموردين والمواد والآلات والعملاء تحقيق ميزة الابتكار، وتحسين الإنتاجية لتخرج بنظام للمستخدم النهائي وتحقيق مستويات جديدة من الجودة والكفاءة بعد أن تم ربطهم.

---

[1] ج.دين, س.جيموات. تبسيط معالجة البيانات في مجموعات كبيرة، الاتصالات في ا س م ، المجلد. 51، رقم. 1، ص 107

[2] ف.فوجداني, أدوات في الوقت الحقيقي للتكامل الأعمال والتعاون ، نظم الطاقة، المعاملات على اي تربل يبي ، المجلد. 18، رقم. 2، ص 555-562، 2003

جيري فوستر هو المدير التنفيذي للشركة يقوم بوضع استراتيجية التكنولوجيا في أنظمة ,PLEX وقال بان السوق هو تجربة مستمرة من الأدوات والتقنيات لاحتضان بيانات كبيرة جديدة, وقال ان الشركة نشأت نتيجة وجود 55 من عملائها يستخدمون الأدوات والآلات المربوطة الكترونيا، ويتم ضبط هذه الادوات والبيانات الكبيرة الناتجة آليا لتلبية مواصفات المنتج وملاحظاتهم الغنية في الوقت الحقيقي.

6- دراسة جيمس موين[1]

في كيفية نقل البيانات الضخمة والتحليلات التنبؤية في تصنيع iconductors حيث البيانات الضخمة أصحبت حلول رئيسية لتوسيع نطاق التحكم في العمليات حددت الدراسة أبعاد ومشاكل البيانات الضخمة في iconductorsعلى النحو التالي حجم البيانات، سرعة البيانات بالنمو، تنوع البيانات، صحة وقيمة (التحليل) ولحل هذه التحديات هذه وضعت حلول مختلفة

- تحليل البيانات لحظة تجميعها لزيادة السرعة بعمليات التحليل لدعم عملية اتخاذ القرار في الوقت الحقيقي.
- استخدام Hadoop أثناء التعامل مع البيانات الكبيرة ومعالجتها بشكل موازي (الأنظمة الموزعة)، لتصبح جاهزة ومتوفرة.
- MapReduc هو نموذج لمعالجة البيانات واسعة النطاق.
- مستودع البيانات يوفر القدرة على التخزين وازالة القيود المفروضة على حجم البيانات.

الاستنتاج

من خلال دراسة التجارب السابقة من خلال دراستنا للأنظمة التي تمت في مجال البيانات الضخمة والتصنيع تم استنتاج مكونات يمكن تلخيصها في إطار نظام جديد والذي يتكون من ثلاثة أقسام رئيسية

1- مصدر البيانات من خلاله يمكن تحديد مصادر البيانات التي تكون بدورها البيانات الضخمة مثل بيانات المنشئة الداخلية أو من المنشئات الخارجية وهي تكون بتنسيقات مختلفة (XML,TXT) ويتم تقسيم هذه البيانات إلى:

- سجلات الآلات في المصنع هي سجلات تحتوي على بيانات المنتجات وسير عملية الانتاج في سلسلة الإنتاج والتي يتم استخدامها لتحديد الفشل أو الخطأ من هذه الآلات للمنتج في عملية الإنتاج.
- أنظمة ألآلات والمعدات وهي أنظمه تحتوي على البيانات المتعلقة بالمواد الخام والجودة لكل منتج.
- نظام ERP الداخلي هو نظام يحتوي على بيانات التكاليف العامة وجميع الموردين بما في ذلك تكاليف المواد.

[1] ج.اسكندر، ج.موين، م.ارماكوست، الصيانة التنبؤية في تصنيع أشباه الموصلات ، في مؤتمر تصنيع أشباه الموصلات المتقدمة(ASMC) ، العدد 26 سنة نصف سنوية، ص. 384-389

– ملفات بصيغها المختلفة تحمل مكونات وتكاليف كل منتج والمعلومات حول نوعية المنتج والعملاء مع القليل من المعلومات عن شراء المنتجات كالعرض والطلب.

– إعادة تنسيق وصياغة البيانات باستخدام أدوات تكنولوجيا المعلومات لتنسيق جميع مدخلات البيانات والحد من أخطاء المدخلات لتتناسب مع بنية قاعدة البيانات استنادا إلى معايير محددة مسبقا لهذه الأدوات.

2– استخدام تقنية Hadoop مع قواعد البيانات وخوادم متعددة لتوفير المعلومات لنظام BI لاستخدامها بغرض التحليل لاحقا

MapReducingيتم استلام البيانات بقواعد بيانات خاصه بشكل منسق من المرحلة السابقة, ويتم ترحيلها الى قاعدة بيانات واحدة من خلال تحديد معايير لتحليل كل مصدر وتحسين البيانات باستخدام هذه المعلمات لحفظ جميع نتائج البيانات في خادم إدارة قواعد البيانات والتي تحمل أنواع البيانات.

3– نظام تطبيق ذكاء الأعمال يحتوي التطبيق عمليات متعددة الخطوات باستخدام قاعدة بيانات واحدة والإعداد وهو يتكون من:

- قاعدة بيانات التصنيع التي تم توفيرها من نتائج عملية Hadoop السابقة عند استرجاع البيانات.
- خادم الشركة يتم تعيين نظام وتحليل أدوات استقصاء المعلومات حتى على هذا الخادم. يتم إنشاء اتصال مع النظام لتعريف المتغيرات لتحليل البيانات النظام وعلاقات التصميم لنظام ذكاء الاعمال ليتم عرضها للمستخدم النهائي.

النتائج المتوقعة من تصميم هذا الاطار خلال عملية التعامل والتحليل للبيانات الضخمة في مجال التصنيع النظام هي

- تقليل وقت معالجة البيانات لأن البيانات تصبح منظمة ومنسقة.
- وتقليل الوقت في اكتشاف الأخطاء والتعجيل في اتخاذ صنع القرار وتعزيز جودة المنتجات.
- اختيار أفضل الموردين والمواد لكل منتج بالاضافه إلى توفير حلول لزيادة أو نقصان خطوط الإنتاج.

الخاتمة:

تعتبر مراقبة البيانات عملية صعبة لأنها تضم عدة عناصر هامة بدءا من الأمان والخصوصية، وحتى تلبية معايير الامتثال الاستخدام العقلاني للبيانات، خاصة عندما يتعلق الأمر بالبيانات الضخمة فإن المشكلات والتحديات تزداد حجما حيث تشير الإحصائيات إلى أنه بحلول عام 2020 سيكون هناك نحو 44 تريليون جيجابايت من البيانات.

تكافح الشركات التجارية الكبيرة لاكتشاف طرق لتخزين البيانات وإدارتها واستخدامها لتحقيق أقصى استفادة منها، لذلك يجب عليها فهم تحديات التعامل معها واستغلالها و إيجاد الحلول التي يجب اعتمادها للتغلب على هذه التحديات وتوسيع مجالات الاستفادة والاستخدام.

قائمة المراجع:

1- تي سي ريدمان، **تحسين جودة البيانات الضخمة لميزة تنافسية**، معهد ماساتشوستس للتكنولوجيا، المراجعة الادارية، المجلد 36الرقم 2 ص99

2- ج.اسكندر، ج.موين، م.ارماكوست، الصيانة التنبؤية في تصنيع أشباه الموصلات ، في مؤتمر تصنيع أشباه الموصلات المتقدمة (ASMC)، العدد 26 نصف سنوية، ص. 384-389

3- ج.دين, س.جيموات. **تبسيط معالجة البيانات في مجموعات كبيرة، الاتصالات** في ا س م ، المجلد. 51، رقم. 1، ص 107

4- حايك, هيام،(2013) **الحوسبة السحابية تغزو مؤسسات التعليم العالي**، مدونة نسيج الإلكترونة،2014 على الرابط:http://blog.naseej.com/

5- د .كاريش صليحة ، محاضرات **مقياس تطبيقات نظم المعلومات ، تخصص ادارة اعمال و انظمة المعلومات** كلية العلوم الاقتصادية و العلوم التجارية و علوم التسيير جامعة الجزائر3 ، 2016 .

6- د .مقدم علال ، **محاضرات مقياس اعلام الي و اتخاد القرار ، تخصص ادارة اعمال و انظمة المعلومات** كلية العلوم الاقتصادية و العلوم التجارية و علوم التسيير جامعة الجزائر3 ، 2016 .

7- د .نسيم قائد عبده ، **اطار مقترح لتطبيق نظام ذكاء الاعمال في التسويق** جامعة تعز الجمهورية اليمنية الجمعية العلمية للمعلوماتية.

8- رامي شلاح نبذة عن Big Data موقع على : https://3alam.pro/rami-

9- ستيف. ديرتين، **تعريف البنية التحتية للتحليلات البيانات الضخمة**، ماي 2015

10- سيد،رحاب فايز أحمد.(2013)،**نظم الحوسبة السحابية مفتوحة المصدر:دراسة تحليلية مقارنة**. المجلة العراقية لتكنولوجيا المعلومات ، ص17-41.

11- شاهين,أحمد(2013).**محرك Google Drive واستخداماته التعليمية**.مدونة مصمم تعليمي، على الرابط: http://www.id4arab.com/2013/04/google-drive.html

12- ف.فوجداني, **أدوات في الوقت الحقيقي للتكامل الأعمال والتعاون ، نظم الطاقة، المعاملات** على اي تربل بيي ، المجلد. 18، رقم. 2، ص 555-562، 2003

13- م،نيسر، **ذكاء الأعمال كتقنية دعم عملية صنع القرار في المؤسسة الحديثة**، مؤتمر ادارة المعرفة يومي 26و28 يونيو

14- ل.كولومبوس، **الطرق العشر للبيانات الضخمة** و هو ثروة التصنيع، نوفمبر 2014

15- ك.هامر، راجاجوبيول **كيف يمكن للبيانات الضخمة تحسين الصناعات التحويلية** ، ماكينزيو الشركات، 2014

1- **cloud computing** ksu.word press.com/05/03/2012

2- Chapter 1 **;Machine learning and data mining,** http :goo.glfsgdjeaccessed 2018.-

3- Global Pulse, **Big Data for Development Challenges and Opportunities**, New York, May 2012, p24

4- Naci Akkok, Pal Halvorsen,Arthur M.Keller, Vera Geobel, **Semistructured Data and XML**, union 2006.

5-Saif and S wazir, **Performance analysis of big data and cloud computing technique Asurvery**, science direct May13 ;2020.