قواعد البيانات الافتراضية
وأثرها على نظم استرجاع المعلومات
بالمؤسسة المعلوماتية : دراسة استطلاعية

د. أحمد إبراهيم شاهين

مدرس علم المعلومات

قسم المكتبات والوثائق - كلية الآداب، جامعة بني سويف

استاذ مساعد علم المعلومات

كلية علوم الحاسب والمعلومات، جامعة المجمعة

Librarian_ahmed@yahoo.com

مستخلص:

يناقش البحث قضية قواعد البيانات الافتراضية وتوظيفها في قطاع المكتبات والمعلومات، وذلك من خلال إلقاء الضوء على مدى إمكانية الإفادة منها في نظم استرجاع المعلومات، وتوظيف إمكاناتها في خدمة الاحتياجات المستفيد المعلوماتية، وتوظيفها لخدمة النظم الآلية للمكتبات في التطبيقات المعلوماتية المستخدمة' في المؤسسة المعلوماتية.ويوصي الباحث بمزيد من الدراسات العربية في هذا المجال لما له من فائدة على تخصص المكتبات والمعلومات.

الكلمات المفتاحية: قواعد البيانات الافتراضية

Virtual Database effectiveness on Information Retrieval Systemin information organization: a Pilot study□

Abstract:

The paper discusses the issue of Virtual databases and using them in the library and information sector, through to shed light on the extent of the potential benefit from the information retrieval systems, employing its potential to serve needs of the beneficiary information, and use it to automated systems for libraries in informatics applications used in the informational enterprise. researcher Recommends more Arabic studies in this area because of its benefit to the specialization of library and information Science□

١/١تمهيد:

تعد قواعد البيانات الافتراضية من التطبيقات البرمجية لتكنولوجيا المعلومات التي لم يتم إلقاء الضوء عليها في الانتاج الفكري العربي بالقدر الكافي ومدى إمكانية توظيفها في المؤسسة المعلوماتية، حيث تعد نقلة نوعية في مفهوم قواعد البيانات الذي تعتمد عليه المؤسسة المعلوماتية في حفظ المعلومات وتخزينها واسترجاعها، كما يمكن توظيفها في تطوير نظم استرجاع المعلومات وآلياتها في المؤسسة المعلوماتية.

وفي ظل سعي المؤسسات المعلوماتية لمواكبة التطور في تكنولوجيا المعلومات وبرمجياتها، وحرصها على توظيفها في خدمة المستفيد، تسعى هذه الدراسة لإلقاء الضوء على برمجيات قواعد البيانات الافتراضية وآليات تطبيقها في المؤسسة المعلوماتية وتوظيفها مع التقنيات المتوفرة حالياً في المجال.

٢/١ أهمية البحث

تنبع اهمية البحث من ما يلي:

١. ندرة الدراسات العربية في مجال قواعد البيانات الافتراضية وتطبيقاتها بالمؤسسات المعلوماتية.
٢. التطور الحادث في برمجيات قواعد البيانات مما يستلزم معه مواكبة التطور في التقنيات المستخدمة في المؤسسات المعلوماتية.
٣. اهتمام المؤسسة المعلوماتية بتطوير التقنيات المستخدمة في الخدمات المعلوماتية، بما يعود بالنفع على خدمات المستفيدين.

١/٣ مشكلة البحث

تنبع مشكلة البحث الحالي من التطور الواضح في برمجيات قواعد البيانات وتطور مفهوم قواعد البيانات الافتراضية وتقنياتها، وفي ظل سعي المؤسسة المعلوماتية إلى مواكبة التطور في تقنيات المعلومات وبرمجياتها لتحقيق اهدافها في خدمة المستفيد من خلال انظمة استرجاع اكثر تطوراً. يسعى هذا البحث لتوضيح مفهوم قواعد البيانات الافتراضية ومدى إمكانية توظيفها في المؤسسة المعلوماتية وخدماتها.

لذا ينطلق البحث من فرضية أن " قواعد البيانات الافتراضية يمكن توظيفها في نظم استرجاع المعلومات وآلياتها بما يخدم المستفيدين من خدمات استرجاع المعلومات بالمؤسسة المعلوماتية ".

١/٤ الأهداف

١. إلقاء الضوء على مفهوم قواعد البيانات الافتراضية.
٢. توضيح بنية قواعد البيانات الافتراضية وآليات عملها لإبراز اختلافها النوعي عن أنواع قواعد البيانات المستخدمة.
٣. إلقاء الضوء على توظيف قواعد البيانات الافتراضية في نظم استرجاع المعلومات.
٤. إلقاء الضوء على نماذج تطبيق قواعد البيانات الافتراضية في الأنظمة الآلية للمؤسسات المعلوماتية.

١/٦ حدود الدراسة

الحدود الموضوعية: تهتم الدراسة ببرمجيات قواعد البيانات الافتراضية وإمكانية تطبيقها في المؤسسات المعلوماتية من زاوية معلوماتية.

الحدود النوعية: تركز الدراسة على برمجيات قواعد البيانات الافتراضية دون التعرض لأنواع قواعد البيانات المستخدمة لأغراض البحث والإسترجاع في المؤسسات المعلوماتية.

٧/١ منهج الدراسة وأدواتها

تعتمد هذه الدراسة على المنهج الوصفي التحليلي في رصد مفهوم قواعد البيانات الافتراضية وبداية ظهوره وتطوره. كما سيتم التعرض بالتحليل لتصميم قواعد البيانات الافتراضية؛ وبنيتها؛ وبرمجياتها؛ وآلياتها؛ وتطبيقاتها في المؤسسات المعلوماتية، وذلك من خلال استعراض أدبيات الموضوع للتعرف على آليات تطبيق قواعد البيانات الافتراضية لخدمة أهداف المؤسسة المعلوماتية.

وقد اعتمد الباحث على أداة البحث الوثائقي في القراءات النظرية للعديد من الدراسات التي تناولت هذا الموضوع.

٨/١ أدبيات الإنتاج الفكري

استعراض الباحث أدبيات الموضوع في أدلة الإنتاج الفكري في مجال المكتبات والمعلومات وقواعد البيانات العالمية والدوريات المتخصصة التي تحتوي على النصوص الكاملة، وكذلك محركات البحث بمصطلحات الدراسة، وهي:

- قواعد البيانات الافتراضية Virtual Database

لاحظ الباحث أن أدبيات الإنتاج الفكري العربي والأجنبي لم تتناول العلاقة بين قواعد البيانات الافتراضية ونظم استرجاع المعلومات من زاوية معلوماتية، غير أن الإنتاج الفكري الأجنبي عالج العلاقة بين المفهومين من زاوية تقنية. وكان اهم ما افرزه الإنتاج الفكري مرتبط بموضوع الدراسة:

- Wada, Yuji, others. Virtual Database Technology for Distributed Database in Ubiquitous Computing Environment.- American Journal of Database Theory and Application 2012, 1(2) 13-25 DOI: 10.5923/j.database.20120102.02

الهدف من البحث تطوير تقنية التمثيل الافتراضي لقاعدة البيانات من أجل السماح لمحللي البيانات او المستخدمين الآخرين لتطبيق اساليب استخراج البيانات في وظائفهم من خلال استخدام كافة قواعد البيانات المتوفرة على شبكة الإنترنت كما لو انها قاعدة بيانات واحدة، وبالتالي يساعد على تقليل أعباء العمل الخاصة بهم مثل جمع البيانات من قواعد بيانات الإنترنت.

في هذه الدراسة، تم دراسة مزايا مخطط XML واقتراح اسلوب المحاكاة الافتراضية لقواعد البيانات method of virtualization of ubiquitous databases المتاح الوصول إليها في اي مكان ومنها قواعد البيانات العلائقية relational databases، وقواعد بيانات الكائنات object-oriented databases، وقواعد البيانات XML database والتي يمكن الوصول إليها في اي مكان كما لو انها قاعدة بيانات واحدة. وكانت اهم نتائج الدراسة ان طريقة المحاكاة الافتراضية من قواعد البيانات في كل مكان التي يمكن وصفها بمخطط قاعدة البيانات ubiquitous database schema المتاحة في كل مكان تستخدم في التطبيق مخطط XML.

- Hiroaki Ohshima, Satoshi Oyama, Katsumi Tanaka. Cloud as Virtual Databases: Bridging Private Databases and Web Services.- Lecture Notes in Computer Science Volume 5981, 2010, pp 491-497

تتناول هذه الدراسة مقترح لتصميم منصة متجددة لتواصل قواعد البيانات الخاصة وخدمات الويب. والفكرة الرئيسة للبحث تتمثل في تكوين

جداول افتراضية لنتائج البحث على خدمات الويب البحثية في بيئة قاعدة البيانات العلائقية RDB. بهدف الكشف عن البيانات الخاصة المخزنة في قواعد البيانات الخاصة RDBs التي لا يمكن الولوج إليها بدون تفويض، لذا فإن هذه الجداول الافتراضية تحقق ربط بين قواعد البيانات الخاصة وخدمات الويب.

- Frehner , Marcel , Brändli , Martin, Schenker , Jürg. The Virtual Database – A Tool for Integrated Data Processing in a Distributed Environment.- EnviroInfo 2004 (Geneva).- pp 537-548.

تتناول الدراسة نظم المعلومات الجغرافية الجاهزة وإمكانية الإعتماد على قواعد بيانات افتراضية لتقديم نفس الخدمات في بيئة الويب، وتمكن من تقليل التكاليف المادية لاقتناء الأنظمة الجاهزة مع قدرات عالية على الاسترجاع ومقارنة البيانات المسترجعة وتحليلها مع الاعتماد على البرمجيات مفتوحة المصدر، وذلك من خلال تصميم قاعدة بيانات افتراضية وتجربتها.

وتعرض الدراسة فكرة قاعدة بيانات افتراضية باعتبارها بنية لدمج مستودعات البيانات الموزعة. بحث توفر إمكانية تكامل البيانات في تصور مشترك، وتحليل البيانات الموزعة، وبناء الأساس لنظام المعلومات البيئية والمناظر الطبيعية الشاملة. وتهدف قاعدة البيانات الافتراضية المصممة بالدراسة إلى دمج قواعد البيانات الموزعة المختلفة مثل Swiss Agency for the Environment, Forests and Landscape (SAEFL). وكانت من أهم النتائج التي توصلت إليها الدراسة، أنه يمكن استخدام قواعد البيانات الافتراضية في مجال العلوم البيئية والجغرافية.

- Jacob Berlin, Amihai Motro. Autoplex: Automated Discovery of Content for Virtual Databases.- Lecture Notes in Computer Science Volume 2172, 2001, pp 108-122

ورقة بحثية تنطلق من فرضية أن انظمة قواعد البيانات الافتراضية تعد مناسبة للبيئات التي تضم مجموعة صغيرة من مصادر المعلومات، وتتسم بالاستقرار، وعدم التعديل المستمر والتغيير. وبناء على ذلك، فإن اتساع نطاق استخدام نظم قواعد البيانات الافتراضية الحالية لا يعد مناسب لأغراض كثير من المستخدمين. والسبب الرئيسي هو تعقيد عملية دمج مصادر معلومات جديدة في قاعدة البيانات الافتراضية وارتفاع التكاليف هذه العملية.

وتركز هذه الورقة على وصف مشروع تصميم نظام Autoplex المستخدم في تقنيات تعلم الآلة، لتوفير تطبيق آلي لأكتشاف محتويات جديدة لنظم قواعد البيانات الافتراضية، حيث يفترض نظام Autoplex ان العديد من مصادر المعلومات تم دمجها وتحديدها في نظام قواعد البيانات الافتراضية من قبل خبراء بشر كما تم سابقاً في نظام تم تصميميه يطلق عليه " نظام قواعد البيانات الافتراضية القياسي". حيث يبني نظام Autoplex نمط تعليمي من خلال استقراء ملامح الأمثلة السابقة لتطبيقات الخبراء، ومن ثم بناء نمط معرفي، ثم تطبيق هذا النمط المعرفي على مصادر معلومات جديدة يتم ترشيحها في محاولة لاستنتاج مقترحات تشبة النمط المعرفي المبني على الأمثلة السابقة.

- Amihai Motro, Philipp Anokhin, Jacob Berlin . Intelligent Methods in Virtual Databases.- Advances in Soft ComputingVolume 7, 2001, pp 580-591

توضح الورقة البحثية الحالية التقدم الكبير الذي تحقق في مجال نظم قواعد البيانات الافتراضية، ويبرز المشكلات التي ما تزال قائمة. وقد ناقشت هذه الورقة اتجاهين من المشكلات الحالية التي ابرزتها البحوث والدراسات، أولهما الاختلافات الجوهرية بين قواعد البيانات المشاركة في قاعدة بيانات افتراضية، وثانيهما الإدراج التلقائي لمصادر المعلومات الجديدة استنادا إلى عملية البحث عنها ورصدها.

وقد قام الباحثون بتطبيق ما يشار إليه بتقنيات الحوسبة المرنة soft computing techniques لحل كلا المشكلتين.

- Masunaga, Y., Watanabe, C., "Design and implementation of a multi-modal user interface of the Virtual World Database system (VWDB)," Database Systems for Advanced Applications, 2001. Proceedings. Seventh International Conference on , vol., no., pp.294,301, 21-21 April 2001.-
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=916390&isnumber=19788

يناقش البحث الحالي تقارير عن تصميم وتنفيذ واجهة المستخدم متعددة الوسائط لنظام قاعدة البيانات العالمية الافتراضية (VWDB). حيث عرضت الورقة البحثية ما قامت به جامعة Ochanomizu University لتطوير نظام VWDB بهدف تحقيق نظام قاعدة بيانات الجيل الجديد لتطبيقات الفضاء الإلكتروني cyberspace applications. من خلال إنشاء نظام يعمل دمج نظام الواقع الافتراضي، ونظام قاعدة بيانات الكيانات.

وتوفر VWDB واجهة قاعدة بيانات للمستخدم في العالم الافتراضي، ويدعم النظام لغات متعددة لنماذج قواعد البيانات والتي تمكن المستخدمين من التفاعل مع VWDB عن طريق الإيماءات والصوت. وتم الاعتماد في البحث على منهجية تصميم اللغة متعدد النماذج. وذلك اعتماداً على ان الكيانات تغير حالتها في العالم الافتراضي وفقا لاستعلام المستخدم أو طلب التحديث. وذلك من خلال تحديد مجموعة من الخصائص الآلية لتوصيف التفاعل المتعدد للوسائط ووصف منهجية الشبكة الدلالية لترجمة رسائل المستخدم متعددة الوسائط المدخلة للنظام إلى رسالة معادلة للوصول والتعامل مع كائنات قاعدة البيانات.

٩/١ المصطلحات الإجرائية

لأغراض هذه الدراسة سيتم وضع التعريفات للمصطلحات التالية:

- قاعدة البيانات الافتراضية: قاعدة بيانات تعمل كواجهة موحدة لعدد من قواعد البيانات الحقيقة المختلفة في النوع، بهدف البحث في قواعد البيانات الحقيقية عن مصادر المعلومات، واسترجاعها، ودمجها، واستبعاد المتكررات، وعرض النتائج في واجهة واحدة ".
- نظم استرجاع المعلومات: نظم الاسترجاع المبنية على قواعد بيانات والتي يتم تخزين البيانات بها، وحفظها، وتحديثها، واسترجاعها، ليتم معالجتها وتقدم مخرجاتها بناء على الاستفسارات البحثية المقدمة من المستفيدين.
- المؤسسة المعلوماتية: مؤسسة تهدف لتقديم خدمات معلوماتية للمستفيدين من واقع مصادر المعلومات الرقمية التي تملكها أو تملك حق الوصول إليها من خلال العاملين بها، وتوظف نظم استرجاع المعلومات لخدمة مستفيديها.

١/٢ قواعد البيانات الافتراضية في المؤسسة المعلوماتية اصطلاحاً ومفهوماً

تعد قواعد البيانات الافتراضية Virtual Database من المفاهيم التي لم يلقى عليها الضوء في أدبيات الانتاج الفكري العربي، لذا يسعى الباحث في هذا الجزء من الدراسة إلى إلقاء الضوء على ثلاث مفاهيم أساسية قواعد البيانات بوجه عام Database؛ مفهوم افتراضي Virtual؛ وصولاً إلى مفهوم قواعد البيانات الافتراضية في أدبيات الانتاج الفكري في مجال علم المكتبات والمعلومات، بهدف وضع تعريف إجرائي للدراسة.

١/١/٢ مفهوم قواعد البيانات في ادبيات الانتاج الفكري في مجال المكتبات والمعلومات

هناك العديد من التعريفات لمفهوم قاعدة البيانات في الانتاج الفكري الأجنبي والعربي، ويسعى الباحث إلى عرض مجموعة من

التعريفات على سبيل المثال وليس الحصر لمفهوم قاعدة البيانات في ادبيات المجال والمفاهيم ذات الصلة بها.

حيث يرد تعريف قاعدة البيانات في معجم ODLIS على اوجه متعددة منها " مجموعة شاملة من البيانات ذات الصلة التي تنظم لسهولة الوصول باستخدام الحاسب".[١]، وكذلك "مجموعة منظمة من البيانات التي يمكن الوصول إليها على الفور، والتحكم بها من قبل نظام تجهيز البيانات لغرض محدد". وايضاً بانها " مجموعة من البيانات مرتبة لسهولة وسرعة البحث والاسترجاع". و" مجموعة من البيانات المنظمة ذات الصلة في شكل إلكتروني التي يمكن الوصول إليها والتحكم بها من قبل برامج الحاسوب المتخصصة". و" مجموعة من البيانات مرتبة لسهولة وسرعة البحث والاسترجاع من خلال جهاز الحاسب".[٢]

يعرف معجم FOLDOC (Free on-line dictionary of computing) قاعدة البيانات بأنها مجموعات كبيرة من البيانات المهيكلة التي يتم تحديثها بشكل مستمر، وتدار من خلال برنامج للتحديث والاستعلام عن البيانات. وقاعدة البيانات البسيطة قد تكون عبارة عن ملف واحد يحتوي على العديد من السجلات، كل منها يحتوي على نفس عدد الحقول. وتكون قاعدة البيانات جزء من نظام إدارة قواعد البيانات DBMS.[٣]

كما يرد تعريف أكثر تفصيلاً لقاعدة البيانات في معجم ODLIS بأنها ملف كبير يحدث بانتظام من المعلومات الرقمية (التسجيلات الببليوجرافية، والمستخلصات، وثائق النص الكامل، والصور، والإحصاءات، وغيرها) تتعلق بموضوع او مجال محدد، ويتكون من سجلات في شكل موحد منظمة تيسر سهولة وسرعة البحث والاسترجاع وتدار بواسطة برمجيات نظم إدارة قواعدة البيانات DBMS. ويتم إنشاء المحتوى من قبل منتجي قاعدة البيانات ويتم تأجير المحتوى إلى واحد او اكثر من موردي قواعد البيانات مثل EBSCO، OCLC وغيرهم، الذي يوفر

الوصول إلكترونيا إلى البيانات بعد ان تم تحويله إلى شكل مقروء آليا، وعادة ما تكون قاعدة البيانات متاحة على CD- ROM او عبر الإنترنت من خلال شبكة الإنترنت، وذلك باستخدام برنامج البحث الملكية. ومعظم قواعد البيانات المستخدمة في المكتبات والفهارس وكشافات الدوريات، خدمات الاستخلاص، والمصادر المرجعية ذات النص الكامل المؤجرة سنويا بموجب اتفاقات الترخيص التي تحد من الوصول إلى المستعيرين المسجلين وموظفي المكتبة.[٤]

كما تعرف قاعدة البيانات في معجم Tech Terms بانها بنية لتخزين البيانات بطريقة منظمة. وتحتوي معظم قواعد البيانات على جداول متعددة، والتي قد تضم عدة حقول مختلفة. وتستخدم تقريباً جميع مواقع التجارة الإلكترونية قواعد بيانات لتخزين بيانات المنتجات ومعلومات العملاء. هذه المواقع تستخدم نظام إدارة قاعدة البيانات DBMS، مثل Microsoft Access، او FileMaker Pro ، او MySQL كخلفية للموقع الإلكتروني. وعن طريق تخزين الموقع للبيانات في قاعدة بيانات تكون قابلة للبحث بسهولة إلى جانب تصنيفها وتحديثها. وهذه المرونة امر هام لمواقع التجارة الإلكترونية، وانواع أخرى من المواقع الحيوية.

وقد تطورت قواعد البيانات من الشكل المسطح flat المقتصر على الصفوف والأعمدة البسيطة، مثل جدول بيانات. إلى قواعد البيانات العلائقية التي تسمح للمستخدمين للوصول للبيانات، وتحديثها، والبحث عن المعلومات على أساس العلاقة بين البيانات المخزنة في جداول مختلفة. ويمكن ايضا لقواعد البيانات العلائقية تشغيل الاستعلامات التي تعمل على قواعد بيانات متعددة. في حين كانت قواعد البيانات المبكرة تخزن النصوص او البيانات الرقمية فقط، فإن قواعد البيانات الحديثة تسمح للمستخدمين بتخزين انواع البيانات الأخرى مثل مقاطع الصوت والصور واشرطة الفيديو.[٥]

كما ورد تعريف قاعدة بيانات بمعجم مصطلحات المكتبات والمعلومات والأرشيف بأنها مجموعة من الكاينات Objects ذات العلاقة ببعضها البعض، تشمل على الجدول tables، والنماذج forms، والتقارير reports، الاستفسارات Queries، والنصوص المكتوبة Scripts، والتي نشئت وانتظمت بواسطة نظام لإدارة قواعد البيانات (Database Management System) DBMS. وقد تشتمل قاعدة البيانات على معلومات من أي نوع، مثل قائمة بأسماء المشتركين في مجلة، او بيانات شخصية عن أسماء العاملين في هيئة من الهيئة، او مجموعة من الصور والرسومات الجغرافية، او قصاصات. كما يشير المصطلح إلى أي مجموعة من المعلومات المؤتمتة او غير المؤتمتة دون النظر إلى كيفية الوصول إلى المعلومات او كيفية تنظيمها. وهذه المجموعة المنظمة من التسجيلات المقروءة آلياً تحتوي على واحد او أكثر من توصيفات ببليوجرافية، معلومات في موضوعات محددة، محتويات قاموسية، بيانات (في شكل ارقام أو معلومات في نص)، وهذه البيانات تكون في العادة ذات بنيان ويتم إنتاجها طبقاً لمعايير متسقة.[٦]

ويرى الباحث ان التعريفات السابقة لمفهوم قواعد البيانات ركزت على محاور رئيسة تتمثل في حفظ البيانات، وتنظيمها، وتخزينها بغرض استرجاع المعلومات باستخدام برمجيات وتقنيات الحاسبات، كما تعد قواعد البيانات أساس بناء نظم استرجاع المعلومات للبيانات المهيكلة والتي لا غنى للمؤسسة المعلوماتية عنها في تقديم خدماتها المعلوماتية للمستفيدين، لذا فإن تناول قواعد البيانات من هذا المنطلق يعد الأساس الذي يبني عليه الباحث تناوله لتطبيق مفهوم قواعد البيانات الافتراضية في نظم استرجاع المعلومات.

كما تناولت ادبيات الانتاج الفكري تصنيف لأنواع قواعد البيانات في تفصيل مفهوم قواعد البيانات، ومن الأنواع التي تم إلقاء الضوء عليها من زاوية معلوماتية "قاعدة البيانات الببليوجرافية Bibliographic Database" والتي تعد بمثابة ملف file بالكمبيوتر يتالف من مداخل

إلكترونية electronic entries تسمى تسجيلات records، وكل تسجيلة تشتمل على وصف للوثيقة او المادة البليوجرافية bibliographic item، يمكن استرجاعها عادة عن طريق المؤلف أو العنوان أو رؤوس الموضوعات أو الكلمات الدالة. وبعض قواعد البيانات الببليوجرافية تكون عامة وبعضها يوفر الوصول إلى الانتاج الفكري في موضوع أو موضوعات معينة. ويوجد العديد من قواعد البيانات الببليوجرافية التي توفر النص الكامل او على الأقل جزءا من نصوص المواد المكشفة. ومعظم قواعد البيانات تكون مملوكة لأصطحابها والوصول إلى معلوماتها يحتاج إلى ترخيص واتفاق مع البائع.[٧]

كما ان هناك "قاعدة البيانات المتخصصة Niche Database" والتي تعرف بانها " قاعدة بيانات إلكترونية مصممة لتوفير المعلومات حول موضوع محدد جدا، لجمهور محدود. مثال على ذلك Thomson's SDC Platinum، وتوفير البيانات الحالية والتاريخية على عمليات الدمج والاستحواذ للشركات المحلية والدولية، القضايا الجديدة، وتمويل رأس المال الاستثماري".[٨]

تعرف "قاعدة البيانات المتضمنة Embedded Database" بانها " قاعدة بيانات معلوماتية يمكن الوصول إليها من خلال تطبيقات برامج اخرى، مثل Microsoft Office Access، والذي يتضمن ميزة تسمح للمستخدمين تظليل كلمات أو عبارات في وثيقتهم او استخدام خيار "look up" لفتح نافذة البحث وادخل الكلمة الكلمة البحثية في مربع إدخال البيانات، ثم تحدد المصدر أو الخدمة التي ترغب في البحث. وقد تعاقدت مايكروسوفت مع موفري المحتوى للطرف الثالث third-party content providers لدمج بعض المصادر مثل (موسوعة إنكارتا Encarta Encyclopedia، القواميس dictionaries ، والمكانز thesaurus) مباشرة في تطبيقاتها. ويوفر عدد من مقدمي الخدمات مثل Factiva, LexisNexis, and Gale بعض المعلومات مجانا للاسترجاع عن طريق تسجيل الدخول إلى موقع المورد، هذا إلى جانب المعلومات

المتوفرة مقابل رسوم. في معظم الحالات، يجب على المستخدمين دفع ثمن النص الكامل باستخدام بطاقة الائتمان إذا لم يكن لديه حق الوصول من خلال الاشتراك المؤسسي. والميزة الرئيسية لقواعد البيانات المتضمنة توفير الراحة للمستخدم".[٩]

تعرف "قاعدة البيانات المتعددة Metadatabase" بانها " قاعدة بيانات لقواعد البيانات، ويتم إعدادها من خلال تجميع قاعدتين بيانات صغيرتين أو اكثر ليتاح للمستخدم البحث في محتوياتها ككل متكامل، بدلا من تكرار نفس البحث في كل قاعدة على حدة مثل قاعدة OneFile من شركة Gale". وقد تسارعت وتيرة ظهور قواعد البيانات المتعددة من خلال تجميع قواعد البيانات من خلال سيطرة موردي قواعد البيانات على حق الوصول لقواعد بيانات الدوريات. وقد يكون هذا النوع من قواعد البيانات مفيد في حال البحث عن الموضوعات البحثية في التخصصات البينية للمستخدمين الذين يفتقرون المهارات البحثية في تحديد قواعد البيانات المثلى لموضوعات تخصصهم.[١٠]

ويرى الباحث أن الأنواع السابقة ركزت على التقسيم المعلوماتي لقواعد البيانات، في حين أن هناك العديد من التصنيفات لأنواع قواعد البيانات؛ منها تصنيف يتعلق بالمحتوى مثل قاعدة بيانات الكايانات Object-Oriented Database (OODB) التي تعد نظام يقدم تسهيلات نظم إدارة قواعد البيانات في بيئة مبرمجة للكائنات objects. حيث يتم تخزين البيانات ككائنات ويمكن أن قرائتها فقط باستخدام البرامج المعدة لقراءة هذه الفئات. وتحفظ العلاقات بين الكائنات المماثلة او المشتركة كمرجع بين الكائنات. ويمكن للاستعلامات ان تكون أسرع بسبب ان العلاقات غالباً غير مطلوبة كما في قواعد البيانات العلائقية. حيث ان الكائنات يمكن استرجاعها مباشرة دون بحث من خلال معرف الكائن.

ويمكن أن تستخدم لغة البرمجة نفسها في تعريف البيانات ومعالجتها، وتكمن قوة نظام قاعدة البيانات في نوع لغة البرمجة التي

يمكن ان تستخدم نموذج هيكلة البيانات والعلاقات بين عناصر البيانات المختلفة. ومما يسهل تطبيقات الوسائط المتعددة أن ارتباط البيانات او الكائنات بنوعية برمجيات محددة قادرة على قرائتها وتفسيرها.

وتوفر OODBs الدعم للإصدارات البرمجية المختلفة مما يساعد على ظهور Object في كافة الاصدارات دون مشكلات تقنية، كما يمكن لاصدارت Object معالجة الكائنات المشابهة لنفس النوع. ووتوفر ايضا OODBsالدعم المنهجي لتقييد الوصول وتعديل البيانات التي تعد اساس انشطة قواعد البيانات. فإن معظم، إن لم يكن جميع البرامج والتطبيقات object-oriented التي تتطلب قاعدة بيانات تستفيد من استخدام OODB."

وهناك أدبيات للإنتاج الفكري تناولت تصنيف قواعد البيانات من حيث تنظيم البيانات، حيث يذكر معجم FOLDOC ان هناك ثلاث طرق للتنظيم أكثر شيوعا في نظم إدارة قواعد البيانات، قاعدة البيانات الهرمية Hierarchical Database، وقاعدة البيانات الشبكة Network Database، وقاعدة البيانات العلائقية Relational Database. قد يدعم نظام إدارة قاعدة بيانات طريقة تنظيم واحدة أو اثنتان، أو جميعهم. كما تستخدم القوائم المقلوبة Inverted lists إلى جانب طرق اخرى. وتعتمد البنية والهيكلية الأكثر ملائمة على معدلات الاستخدام والاسترجاع التي سيتم إجراؤها. [١٢]

ويعرف معجم مصطلحات المكتبات والمعلومات والأرشيف "قاعدة البيانات الهرميةHierarchical Database" بانها قاعدة بيانات منظمة تكون فيها لعناصر البيانات علاقة ثابتة بعضها ببعض. وكل عنصر عال من الهرم مرتبط بعدة عناصر دنيا، في حين أن كل عنصر ادنى مرتبط بعنصر واحد عال فقط. والوصول إلى أي عنصر يتم بالوصول إلى العنصر الأعلى رتبة مباشرة في البنية.

كما يعرفها ذات المعجم في تعريف اخر بانها قاعدة بيانات تتجمع فيها السجلات بطريقة تتفرع فيها العلاقات كالبنايات الشجرية، ومثل هذا البناء يناسب قواعد البيانات الموجودة في الكمبيوترات الكبيرة، كما يناسب تنظيم البيانات التي يمكن تفريعها منطقياً.[١٣]

في حين يرد تعريف قاعدة البيانات الهرمية بمعجم FOLDOC بانها شكل من اشكال قواعد المعتمدة في نظم إدارة قواعد البيانات وتعتمد على ربط السجلات معا مثل شجرة العائلة بحيث يكون كل سجل له مرجع أو أصل واحد فقط. واستخدمت البنية الهرمية على نطاق واسع في انظمة إدارة قواعد البيانات للحاسبات الكبيرة الأولى. ولكن بسبب القيود الخاصة بهذه البنية فإنها غالبا لا تستخدم كثيراً لعدم قدرتها على ان تعكس الروابط والعلاقات الموجودة في العالم الحقيقي.[١٤]

اما "قاعدة البيانات الشبكة Network Database" تعرف بمعجم FOLDOC بانها من اشكال قواعد المعتمدة في نظم إدارة قواعد البيانات وتعتمد على ان كل سجل يكون له اكثر من مرجعية وهو ما يتناقض مع قاعدة البيانات الهرمية او قاعدة بيانات العلائقية.[١٥]

في حين تعرف بمعجم مصطلحات المكتبات والمعلومات والأرشيف بانها قاعدة بيانات لكل عنصر اساسي فيها ملفه الخاص، وتكون الملفات موصولة ببعضها بواسطة مؤشرات، ويمكن الوصول إلى عنصر معين بواسطة أي ملف آخر متصل به، وليس بواسطة مسلك محدد وحيد كما في قواعد البيانات الهرمية.[١٦]

اما "قاعدة البيانات العلائقية Relational Database" فيعرفها معجم مصطلحات المكتبات والمعلومات والأرشيف بانها قاعدة بيانات تتالف من جدول، والجدول تتالف من صفوف واعمدة والتي يمكن تحويرها وتكييفها لأغراض متعددة. ومن العمليات الهامة في قواعد البيانات العلاقية، هي وصل جدولين، بمعنى عمل إحالات مزدوجة Cross

Reference للمعلومات بينهما. بدا هذا النظام Edger Coss حوالي سنة ١٩٧٠. وكل جدول يتالف من اعمدة لها اسماء وهي التي تسمي بالخواص attributes. وكل تسجلية منطقية logical record تتالف من صفوف لها علاقة ببعضها مختارة من جدول مختلفة. وقاعدة البيانات التي تتالف من جدول واحد تسمى قاعدة بيانات الملفات الثابتة flat-file database. وكل قاعدة بيانات علاقية لها لغة تساؤلات query language للتعبير عن الأومر لاسترجاع المعلومات.[١٧]

في حين يرد تعريف "قاعدة البيانات العلائقية Relational Database" او ما يعرف بـ نظام إدارة قواعد البيانات العلائقي RDBMS بمعجم FOLDOC بانها قاعدة بيانات تبنى على النوذج العلائقي الذي وضع بواسطة EF Codd. وتتيح قاعدة البيانات العلائقية تعريف بنية البيانات، وعمليات التخزين، والاسترجاع، ووقيود الحفاظ على نزاهة البيانات. ويتم في قاعدة البيانات العلائقية تنظيم البيانات والعلاقات بينهما من خلال الجداول. ويعتبر الجدول مجموعة من الصفوف أو السجلات وكل صف في جدول يحتوي على نفس الحقول، ويمكن تعيين بعض الحقول كمفاتيح أساس، مما يعني ان عمليات البحث عن قيم معينة من هذا المجال سوف تستخدم الفهرسة لتسريع البحث. ويمكن إجراء عملية الربط حين يكون للحقول في جدولين نفس نوع البيانات والقيم وذلك لتحديد السجلات المرتبطة في الجدولين عن طريق مطابقة القيم في تلك الحقول. وفي كثير من الأحيان، ولكن ليس دائما، يكون للحقول نفس الاسم في كلا الجدولين. ويمكن ربط حقول متعددة في عدة جداول لأنه يتم توظيف هذه العلاقات فقط في وقت الاسترجاع، وتصنف قواعد البيانات العلائقية على انها نظام إدارة قواعد البيانات ديناميكية.[١٨]

ويرى الباحث في إطار تناول مفهوم قواعد البيانات في ضوء نظم استرجاع المعلومات انه من الضروري إلقاء الضوء على المفاهيم ذات الصلة به، ولعل من اهمها مفهوم "نظام إدارة قواعد البيانات Database

(DBMS) Management System)" في ادبيات الإنتاج الفكري باعتباره وثيق الصلة بمفهوم قواعد البيانات.

يعرف نظام إدارة قواعد البيانات (DBMS) بانه " تطبيق حاسب مصمم لمراقبة تخزين البيانات واسترجاع وامنها وسلامتها، وإصدار التقارير عن البيانات المخزنة في شكل سجلات موحدة منظمة في ملف بحث كبيرة يطلق عليه قاعدة بيانات. تتنوع برامج نظم إدارة قواعد البيانات المتاحة ما بين انظمة بسيطة مخصصة لأجهزة الحاسب الشخصية؛ وأنظمة معقدة للغاية تهدف إلى تشغيل اجهزة الحاسب الكبيرة".[١٩]

ويعرف نظام إدارة قواعد البيانات DBMS بمعجم FOLDOC بانه "مجموعة من البرامج تدير مجموعات كبيرة من البيانات المهيكلة التي يتم تحديثها بشكل مستمر ، وتساعد على الاستعلام عن المحتوى لعدد كبير من المستخدمين وذلك من خلال توفير آليات للبحث والاسترجاع. وتستخدم على نطاق واسع في تطبيقات الأعمال".[٢٠]

ويرد تعريف "نظام إدارة قواعد البيانات (DBMS)" بمعجم ODLIS (Online Dictionary of Library and Information) بانه برنامج تطبيقي يساعد على خلق قاعدة بيانات والتحكم فيها، حيث يتولى التنظيم الشامل لها، وتخزين البيانات Storage، واسترجاعها Retrieval، وامنها Security، وسلامتها Data Integrity. ومثل هذا النظام به إمكانيات لتنسيق التقارير وتجهيزها للطباعة، واستجلاب البيانات من تطبيقات اخرى تستخدم اشكال ملفات معيارية. كما تحتوي على لغة للتحكم data manipulation language بالبيانات وإعادة تشكيلها لمساندة الاستفارات من قاعدة البيانات.[٢١]

ويمكن لنظام إدارة قاعدة البيانات (DBMS) ان يكون مجموعة معقدة للغاية من البرامج التي تتحكم في تنظيم وتخزين واسترجاع البيانات (الحقول والسجلات والملفات) في قاعدة بيانات. كما يتحكم

نظام إدارة قواعد البيانات في امن وسلامة قاعدة البيانات. ويتلقى نظم إدارة قواعد البيانات طلبات الحصول على البيانات من البرامج التطبيقية ثم يوجه نظام التشغيل لنقل البيانات المطلوبة. وعند استخدام نظم إدارة قواعد البيانات في المؤسسة، يمكن تعديل نظم المعلومات المستخدم بسهولة وفق التغيرات التي تطراء على الاحتياجات المعلوماتية للمؤسسة، كما يمكن إضافة فئات بيانات جديدة دون التأثير على النظام القائم. وتوفر نظم إدارة قواعد البيانات الأمن للبيانات من خلال منع المستخدمين غير المصرح لهم من الدخول على قاعدة البيانات وتحديثها. وذلك من خلال استخدام كلمات سر للمستخدمين المسموح لهم بالوصول لقاعدة البيانات ولمستوياتها حسب مستوى الصلاحية، ويطلق عليها المخططات الفرعية. [٢٢]

ويمكن لنظم إدارة قواعد البيانات أن يحافظ على سلامة قاعدة البيانات من خلال عدم السماح لأكثر من مستخدم بتحديث نفس السجل في ذات الوقت. ويمكن لنظم إدارة قواعد البيانات الابقاء على سجلات مكررة من قاعدة البيانات. وتسمح لغات الاستعلام وإعداد التقارير للمستخدمين بالبحث التفاعلي في قاعدة البيانات وتحليل البيانات المسترجعة. إذا يوفر نظم إدارة قواعد البيانات وسيلة لدخول تبادلي وتحديث قاعدة البيانات، فضلا عن الاسترجاع للبيانات، وتتيح هذه الامكانات إدارة قواعد البيانات الشخصية. وبالرغم من ذلك فإن عملية المراجعة المستمرة للضوابط والإجراءات امر ضروري في حال تعدد المستخدمين في المؤسسة. وتتوفر هذه الضوابط في حال تهيئة برامج التطبيقات بطريقة مخصصة لوظائف إدخال البيانات وتحديثها. [٢٣]

وقد تستخدم المؤسسات نوع واحد من نظم إدارة قواعد البيانات لمعالجة المعاملات اليومية ثم تنقل التفاصيل إلى حاسب آخر يستخدم نظام DBMS آخر اكثر ملائمة للبحث والتحليل العشوائي. يتولى تنفيذ القرارات الشاملة للنظام المصمم مدير البيانات data administrators

ومحللي النظم systems analysts. كما يتولى تنفيذ تفاصيل تصميم قاعدة البيانات مدير قاعدة البيانات database administrators. [٢١]

ويتكون عالم المعلومات من بيانات، ونصوص، وصور، وصوت. وتدير العديد من DBMSs النصوص فضلا عن باقي انواع البيانات، ولكن عدد قليل جدا منها يدير كلاهما بكفاءة متساوية. وفي خلال التسعينات من القرن الماضي ومع استمرار زيادة القدرات الاستيعابية لوسائط التخزين بدات DBMSs في التعامل مع جميع اشكال المعلومات معاً.[٢٥]

ويعرف نظام إدارة قواعد البيانات المسئول عن إجراء التحقق من صحة البيانات وتجهيزها ومعالجتها من قبل برامج التطبيقات "بقاعدة البيانات الذكية Intelligent Database"، حيث توفر معظم نظم إدارة قواعد البيانات قواعد التحقق من صحة البيانات مثل رفض إدخال التواريخ أو الحروف الأبجدية في حقول العملة، ولكن معظم المعالجات تتم من قبل برامج التطبيقات. ولايوجد حد معين لكمية البيانات التي يمكن معالجتها وفق الوظائف القياسية بقاعدة البيانات الذكية. ومن امثلة التقنيات التي يتم تنفيذها بقواعد البيانات الذكية تقييد الوصول وفق صلاحيات الدخول لقاعدة البيانات فيما يعرف بـ constraints، وقيام النظام باتخاذ إجراءات تمنع تعديل البيانات من قبل غير المصرح لهم؛ او التغييرات غير الصحيحة للبيانات؛ او التضارب في البيانات للحفاظ على سلامة قاعدة البيانات فيما يعرف بـ triggers، إجراءات التخزين stored procedures، حيث يتجه تجهيز قواعد البيانات نحو الحفاظ على نزاهة البيانات حيث انها تضمن التناسق في كافة البيانات المستخدمة. واصبحت قواعد البيانات الحاسبات المركزية اكثر ذكاء وتسعى لمتابعتها نظم قواعد بيانات الحاسبات الشخصية.[٢٦]

ويرى الباحث ان المفهوم السابق لقواعد البيانات ونظام إدارة قواعد البيانات DBMS يرتبط بشكل وثيق بمفهوم نظم استرجاع المعلومات القائم على حفظ البيانات وتخزينها بغرض إسترجاعها من خلال إدارتها

للإجابة عن الاستفسارات المعلوماتية للمستفيدين مع الحفاظ على الثوابت المتعلقة بسرية المعلومات وامنها وسلامتها باستخدام البرمجيات الملائمة لذلك.

٢/١/٢ مفهوم افتراضي في ادبيات الانتاج الفكري في مجال المكتبات والمعلومات

يسعى الباحث إلى عرض مجموعة من التعريفات التي ورد بها مصطلح "افتراضي Virtual" في ادبيات الانتاج الفكري الأجنبية والعربية في مجال المكتبات والمعلومات ومجال الحاسبات باعتبارهما المهتمين بهذا المفهوم، وصولاً إلى مصطلح قواعد البيانات الافتراضية بهدف تاصيل المفهوم في مجال المكتبات والمعلومات.

يرد مصطلح "افتراضي Virtual" من زاوية تقنية بمعجم FOLDOC باعتباره مصطلح تقني يعني بديل شائع للمنطقي، ويستخدم غالباً للإشارة إلى الكائنات الاصطناعية مثل الذاكرة الافتراضية المعنونة addressable virtual memory باعتبارها أكبر من الذاكرة المادية physical memory التي تم إنشاؤها بواسطة نظام حاسوبي لمساعدة نظام التحكم في الوصول إلى الموارد المشتركة. كما يرتبط بمفهوم المحاكاة Simulated التي تعني أداء مهام معينة غير موجود بالفعل.[٢٧]

أما المصطلحات التي ركزت على الجانب المرتبط بالبيئة الرقمية وبيئة الإنترنت فقد وردت بمعجم ODLIS مثل مصطلح "الواقع الافتراضي virtual reality" والذي يعرف بأنه " بيئة إلكترونية انشئت خصيصا لمستخدمي الحاسب من خلال استخدام البرمجيات التي تحاكي المظهر المرئي للواقع ثلاثي الأبعاد ولكن تفتقر إلى الجوهر المادي، وتستخدم أساسا لأغراض التدريب والترفيه".[٢٨] ويعد مجال الواقع الافتراضي Virtual reality (VR) واحد من مجالات تقنيات الحاسب الأكثر جذباً لتطوير تطبيقات متقدمة ومتنوعة في مجال الفضاء الالكتروني. ويتم في الواقع الافتراضي VR صناعة العالم الافتراضي

virtual world أو البيئة الافتراضية virtual environment من خلال مجموعة من أجهزة الكمبيوتر عالية السرعة، وأجهزة عرض متقدمة، وأجهزة طرفية HMD (أجهزة عرض للراس head-mounted display)، وقفازات البيانات، وفارة ثلاثية الأبعاد (3-D). يمكن للمستخدمين العمل في العالم الافتراضي كما لو كانوا في العالم الحقيقي. ويبحث المستخدمين عن اماكن مادية مختلفة للعمل معاً بيئة افتراضية اصطناعية مشتركة. وهناك تقارير عن تصميم وتنفيذ واجهة المستخدم متعددة الوسائط لنظام قاعدة البيانات العالمية الافتراضية (VWDB). حيث تم تطوير VWDB في جامعة Ochanomizu University بهدف تحقيق نظام قاعدة بيانات الجيل الجديد لتطبيقات الفضاء الإلكتروني cyberspace applications. من خلال إنشاء نظام لدمج نظام الواقع الافتراضي، ونظام قاعدة بيانات الكيانات.[٢٩]

وأيضاً مصطلح "الجولة الافتراضية virtual tour" الذي ورد في ذات المعجم بأنه " جولة عبر الإنترنت عن خدمات المكتبة وتسهيلاتها المقدمة للمستفيدين، وعادة ما تكون متاحة عبر شبكة الإنترنت. وتختلف اشكالها ولكنها غالباً تتضمن عدد من الكلمات ذات الارتباط التشعبي clickable المرتبطة بصور مع النص المصاحب واصفاً المجموعات والخدمات المتاحة في كل موقع".[٣٠]

ومن جانب آخر فقد ورد مصطلح Vitrual من زاوية معلوماتية تتعلق مثل مصطلح "المكتبة الافتراضية virtual library" الذي ورد بمعجم ODLIS على بأنها "مكتبة بلا جدران" لا تحتوي على مصادر معلومات ورقية، وبطاقات microform، أو أي شكل ملموس لمصادر المعلومات داخل مبنى المكتبة ولكن يمكن الوصول إليها إلكترونيا في شكل رقمي عبر شبكات الحاسب. وتوجد مثل هذه المكتبات على نطاق محود جداً، ولكن في معظم المكتبات القائمة الأن في الولايات المتحدة الأمريكية تعتمد على المصادر التقليدية المطبوعة، والفهارس وكشافات الدوريات المتاحة على الانترنت، والنص الكامل لعدد من الدوريات

والمصادر المرجعية المتوفرة بشكل رقمي. وتطلق بعض المكتبات والنظم المكتبية على انفسها مصطلح " مكتبة افتراضية" نظراً لتقديمها خدمات عبر الانترنت. ويعد مصطلح "المكتبة الرقمية Digital library" اكثر ملاءمة من "المكتبة الافتراضية" حيث ان مصطلح "افتراضي" تم اقتباسه من "الواقع الافتراضي virtual reality " حيث ان التجربة اثبتت ان المكتبات استبدلت واقع عرض مصدر المعلومات بالشكل المطبوع إلى عرضه على شاشة الحاسب، وبرغم اختلاف جودة العرض ولكن المضمون المعلوماتي واحد. [٣١]

وايضاً مصطلح "الفهرس الموحد الافتراضي virtual union catalog" الذي ورد في ذات المعجم بأنه " نظام آلي للبحث عن مصدر معلومات او اكثر في عدد من فهارس المكتبات المنفصلة في نفس الوقت، وذلك باستخدام معيار Z39.50 و/او آليات اخرى للبحث والاسترجاع، وذلك عكس الفهرس الموحد المركزي centralized union catalog والذي يعتمد على حفظ التسجيلات الببليوجرافية في قاعدة بيانات واحدة". [٣٢]

اما مصطلح "قواعد البيانات الافتراضية Virtual Database" ورد في معجم TheFreeDictionary مع عدد من المصادر المرجعية في علوم الحاسب على بأنها " استعراض واحد لعدد من قواعد البيانات كما لو كانت كلها قاعدة بيانات واحدة". [٣٣]

ويرى الباحث ان ادبيات الانتاج الفكري المرجعية في مجالي المكتبات والمعلومات، والحاسبات تناول مصطلح افتراضي من اوجه مختلفة، منها ما يتعلق بالجانب التقني في بيئتي الحاسب والإنترنت من جانب، واوجه اخرى ترتبط بالجانب المعلوماتي المتعلق بمصادر المعلومات الرقمية في البيئة الرقمية المتمثل في المكتبة الافتراضية؛ وبأدوات استرجاع المعلومات المتمثلة في الفهرس الموحد الافتراضي، في حين ان مصطلح "افتراضي Virtual" لم يقترن بمصطلح "قواعد البيانات

Database" سوى في ادبيات الانتاج الفكري المرجعية في مجال الحاسبات والمعلومات.

ويرى الباحث أن مفهوم قواعد البيانات الافتراضية يرتبط بالجانبين التقني والمعلوماتي؛ حيث أن قواعد البيانات في الأساس برمجيات ولكنها أحد أدوات نظم استرجاع المعلومات؛ كما أن الهدف من استعراض أكثر من قاعدة بيانات من واجهة واحدة توحيد مصطلحات الاسترجاع، وتوفير وقت المستفيد وجهده في حذف النتائج المسترجعة المتكررة، لذا فإن مصطلح قواعد البيانات الافتراضية يعد من المصطلحات البينية التي تقع في المنطقة المشتركة بين قطاعي المعلومات والحاسبات.

لذا يقترح الباحث تعريف إجرائي لمفهوم قاعدة البيانات الافتراضية بأنها " قاعدة بيانات تعمل كواجهة موحدة لعدد من قواعد البيانات الحقيقة المختلفة في النوع، بهدف البحث في قواعد البيانات الحقيقية عن مصادر المعلومات، واسترجاعها، ودمجها، واستبعاد المتكررات، وعرض النتائج في واجهة واحدة".

٢/٢ قواعد البيانات الافتراضية Virtual Database

١/٢/٢ بنية قواعد البيانات الافتراضية Architecture of VDB

يهدف هذا الجزء من الدراسة إلى إلقاء الضوء على بنية قواعد البيانات الافتراضية وتصميمها من خلال استعراض ادبيات الانتاج الفكري في مجال الحاسبات والمعلومات التي تناولت بنية قواعد البيانات الافتراضية.

في دراسة مبكرة حول تطوير بنية نظام VWDB (The Virtual Web DataBase System Architecture) من خلال دمج نظام VR (Virtual Reality) ونظام قاعدة البيانات الكيانات، باعتبارها أنسب وسيلة لتحقيق أهداف نظام VWDB، توصل الباحثان Masunaga and

Watanabe إلى ان هناك ثلاث طرق ممكنة لتحقيق النظام، على النحو التالي:

- تعزيز نظام VR لتشمل وظيفة قاعدة البيانات.
- تعزيز نظام قاعدة بيانات لجعله نظام واقع افتراضي.
- دمج نظام VR ونظام قاعدة البيانات.

وقد توصل الباحثان Masunaga and Watanabe إلى ان الاختيار الأول ليس حلا مناسباً، حيث يتطلب تنفذ جميع وظائف قاعدة البيانات من نقطة الصفر، والتي قد تشمل الكثير من إعادة اختراع عمل تم بالفعل في حقل قواعد البيانات. وهناك سلبيات في الاختيار الثاني لنفس الاسباب السابقة والمتعلقة ببناء الواقع الافتراضي من الصفر في قاعدة البيانات. وبالتالي فإن الاختيار الثالث يعد البديل والأسلوب الأكثر منطقية لتطوير نظام VWDB.

وكانت المشكلة التي واجهت الباحثان تتعلق بتحديد أي نوع من أنواع نظم قواعد البيانات مناسبة لنظام VWDB. دون الاهتمام بسرعة المعالجة، أي كان نوع نظام قاعدة البيانات سواء نظام قاعدة بيانات الشبكية او نظام قواعد البيانات العلائقية. وفي حال الرغبة في معالجة الاستعلامات والتحديثات بكفاءة عالية فإن نظام قاعدة البيانات الكيانات هو الأنسب، لأن الكائنات في العالم الافتراضي وعادة ما تكون من الأجسام المركبة. ونظم قواعد البيانات الكيانات أكثر كفاءة لتخزين ومعالجة الكائنات مركبة من انظمة قواعد البيانات العلائقية.[٢٤]

ويرى الباحث ان المحاولات المبكرة لبناء قواعد البيانات الافتراضية ركزت على المزج بين تقنيات الواقع الافتراضي VR وبرمجيات قواعد البيانات DB، إلا ان هذا الاتجاه تغير في الدراسات الحديثة حيث ركزت على مفهوم مختلف يتعلق بالافتراضية virtualization يتمثل في بناء قاعدة بيانات تعمل افتراضياً على قواعد

بيانات فردية حقيقية، وبالتالي كان التركيز على حل مشكلة الاختلاف بين قواعد البيانات من حيث تمثيل البيانات، وهيكليتها، ولغة البرمجة المستخدمة في تصميمها والاستعلامات، وهو ما سيركز عليه الباحث في هذا الجزء من الدراسة.

وتتضمن قاعدة البيانات الافتراضية المتاحة للمستخدم على:

١. مجموعة من قواعد البيانات الحقيقية في الواجهة الخلفية
٢. مدير التوثيق authentication manager الذي يحقق التوافق بين قاعدة البيانات الافتراضية وقواعد البيانات الحقيقية في الخلفية من خلال التحكم في تسجيل الدخول / كلمة السر
٣. مدير الطلب request manager الذي يحدد سلوك وحدة تحكم قاعدة البيانات الافتراضية.[٣٥]

ويوجد أنواع مختلفة من قواعد البيانات المتوفرة في الاسواق من حيث نموذج البيانات ومن حيث الموردين، حيث تختلف في نماذج تمثيل البيانات ومعالجة الارتباطات بها احياناً بشكل فريد من نوعه.[٣٦]

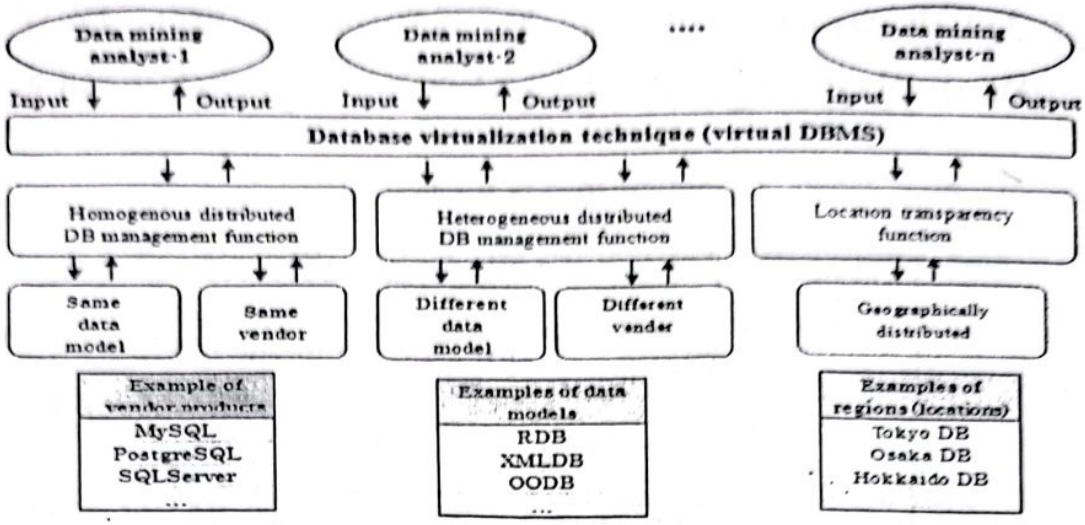
ويمكن أن تحتوي قواعد البيانات الافتراضية VDBs على واحد او أكثر من نماذج تمثيل البيانات والتي تكون في حالة متكاملة ويمكن استعراضها باستخدام التطبيقات البرمجية المتاحة في الأسواق. ويجب أن تكون هذه النماذج في حالة صالحة حتى يمكن استخدامها بقاعدة البيانات الافتراضية VDB من اجل الوصول إلى البيانات. وللتحقق من نموذج واحد يجب أن يكون هذا النموذج في حالة متسقة ذاتيا self-consistent وكاملة، وهذا يعني انه لا توجد اجزاء مفقودة من البيانات، ولا توجد إحالات إلى كيانات غير موجودة. ويتم تنفيذ عملية التحقق من صحة النماذج المتعددة على كافة النماذج وعناصرها الفرعية الموجودة بقواعد البيانات الخارجية. لذا يجب ان تكون قاعدة البيانات الافتراضية VDB دائما

في حالة كاملة، بمعنى ان توجد جميع المعلومات داخل قاعدة البيانات الافتراضية VDB نفسها ولا توجد في قواعد بيانات خارجية.[٣٧]

ومن امثلة نماذج تمثيل البيانات المستخدمة استخدام الجدول كتمثيل في قواعد البيانات العلائقية (RDB)، واستخدام XML كتمثيل في قواعد بيانات XML (XMLDB)، وقواعد بيانات الكيانات (OODB). واحياناً تكون هناك ميزات مختلفة بين الموردين في قواعد البيانات من نفس النوع. وعلى سبيل المثال، نجد فيما يتعلق بقواعد البيانات العلائقية RDB بعض الاختلافات في لغة SQL و/ او نوع تمثيل البيانات. والمثال النموذجي على الاختلاف توافر لغات برمجة مثل MySQL, PostgreSQL, SQLServer لدى موردين مختلفين.[٣٨]

ويرى الباحث ان الاختلاف في نماذج تمثيل البيانات بقواعد البيانات الهدف منه معالجة الأنواع المختلفة من البيانات بما يحقق الهدف من قاعدة البيانات واحتياجات مستخدميها، ولكن تكمن المشكلة الحقيقية في التعامل مع قواعد بيانات عدة في آن واحد تختلف من حيث تمثيل البيانات ونماذجها؛ وهنا تظهر اهمية قواعد البيانات الافتراضية في تقديم حل باعتبارها واجهة واحدة قادرة على التعامل مع قواعد البيانات المختلفة من حيث تمثيل البيانات ونماذجها.

حيث يعد الهدف الرئيس من قواعد البيانات الافتراضية تحقيق التجانس بين قواعد البيانات الموزعة Visualization of Homogeneously Distributed Databases، لذا تعد الخطوة أولى في بناء قاعدة البيانات الافتراضية تقديم وصف لطريقة بناء نظام إدارة قاعدة البيانات الافتراضية لمجموعة من قواعد البيانات العلائقية RDBs المقدمة من موردين مختلفين.[٣٩]

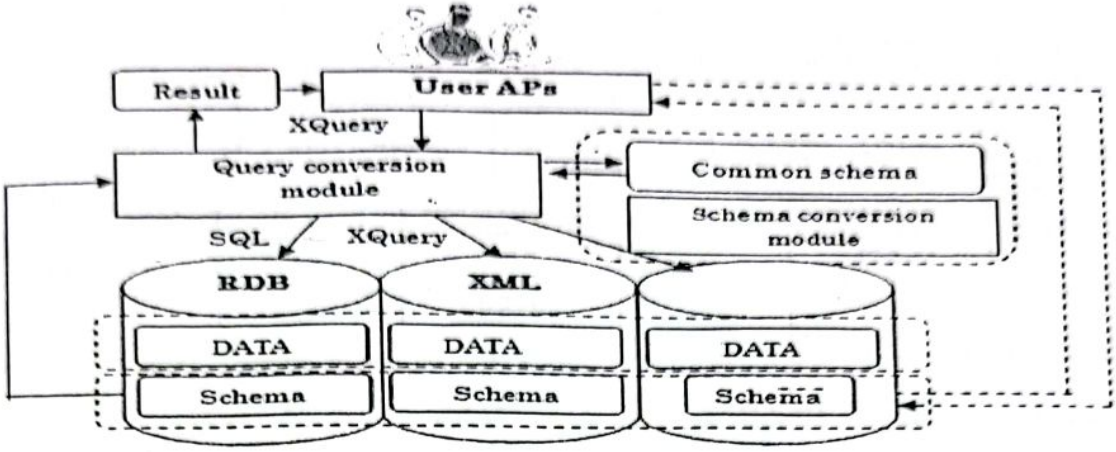


شكل رقم (١) يوضح طريقة بناء نظام إدارة قواعد البيانات الافتراضية لقواعد بيانات علائقية RDB.[٠]

ويوضح الشكل رقم (١) طريقة بناء نظام إدارة قواعد بيانات افتراضية لعدد من قواعد البيانات العلائقية مصممة بلغات برمجة مختلفة مثل MySQL، وPostgreSQL، وSQLServer، كما تختلف في نظام تمثيل البيانات ما بين قواعد بيانات تستخدم التمثيل العلائقي (RDB)، وقواعد بيانات تستخدم XML كتمثيل في قواعد بيانات XML (XMLDB)، وقواعد بيانات الكيانات (OODB)، كما تختلف ايضاً في موقع تواجد قاعدة البيانات الحقيقية ما بين مدن Tokyo، وOsaka، وHokkaido. حيث يكون الهدف الرئيس لنظام إدارة قاعدة البيانات الافتراضية القيام بدور الوسيط لتحقيق التجانس بين قواعد البيانات التي تتفق في نموذج تمثيل البيانات والمورد من جانب؛ وقواعد البيانات التي تختلف في نموذج تمثيل البيانات والمورد من جانب أخر؛ وقواعد البيانات الموزعة جغرافياً في اماكن مختلفة.

ويرى الباحث أن التجانس الذي يحققه نظام إدارة قواعد البيانات الافتراضية يساعد في تطبيق مفهوم التنقيب عن البيانات Data Mining من خلال تحليل النتائج المسترجعة ودمجها واستبعاد المتكررات وعرضها

في واجهة واحدة للمستفيد مما يوفر عليه الوقت والجهد المبذول في عملية البحث، وبالتالي تحقيق هدف عملية استرجاع المعلومات للمستفيد.



شكل رقم (٢) يوضح تحقيق التجانس بين قاعدتي بيانات مختلفتين في النوع RDB وXMLDB باستخدام المخطط المشترك.[١]

ويوضح الشكل رقم (٢) تحقيق التجانس من خلال تطبيق مفهوم الافتراضية على قواعد بيانات مختلفة من حيث النوع، حيث يتم وصف معلومات مخطط نموذج تمثيل البيانات لكل قاعدة بيانات حقيقية؛ وذلك للوصول إلى وضع مخطط مشترك common schema لقاعدة البيانات الافتراضية، ويتم استخدام المخطط المشترك لتحقيق التجانس في الاستعلام والإسترجاع بين قواعد البيانات الحقيقية في قاعدة البيانات الافتراضية، وفي الشكل السابق يجب الحصول على مْعلومات مخطط قاعدة البيانات العلائقية RDB للعمل عليها برمجياً لتحويلها إلى مخطط XML من خلال عملية برمجية محددة وذلك لإنجاز عملية تحويل المخطط schema conversion لنماذج مختلفة من قواعد البيانات الحقيقية.[٢]

ويرى الباحث من الشكل رقم (٢) يوضح أن الاستعلام الموجه من المستفيد لقاعدة البيانات الافتراضية مصاغ بلغة XQuery، لذا يتم إرساله إلى نموذج تحويل الاستعلام Query conversion module ليتم تحويل لغة الاستعلام الموجهة من المستفيد إلى لغتي الاستعلام XQuery ، وSQL

المستخدَمتا في قواعد البيانات الحقيقية والتي تظهر في الشكل؛ وذلك بإرسال لاستعلام إلى المخطط المشترك Common schema الذي يتضمن نموذج تحويل المخطط Schema conversion module الذي يساعد على تنفيذ المخطط المشترك لمهامه برمجياً، ثم يتم توجيه الاستعلام بلغتي XQuery، وSQL لكل قاعدة بيانات حقيقية وفق نمُوذج تمثيل بياناتها، ويتم بعد ذلك إرسال نتائج الاستعلام إلى نموذج تحويل الاستعلام Query conversion module لإرسال النتائج للمستفيد.

ويخلص الباحث في هذا الجزء من الدراسة إلى ان بنية نظام إدارة قواعد البيانات الافتراضية يساعد على تحقيق التجانس بين قواعد البيانات الفردية المختلفة من حيث البنية، وهيكلية البيانات وتمثيلها، ولغة الاستعلام بها، مما ينعكس على اتساع نطاق البحث عن مصادر المعلومات من جانب؛ واسترجاع المعلومات في واجهة واحدة بعد استبعاد المتكررات من جانب اخر، بما يوفر وقت المستفيد وجهده، وجميعها تمثل الوظائف الأساسية لنظام استرجاع المعلومات.

٢/٢/٢ تصميم قاعدة البيانات الافتراضية Design of VDB

يرى الباحث أنه لفهم أعمق لقواعد البيانات الافتراضية وتوظيفها في نظم استرجاع المعلومات، فإنه من المهم إلقاء الضوء على متطلبات تصميم قواعد البيانات الافتراضية.

يتبع تصميم قاعدة البيانات الافتراضية وتنفيذها اتجاه تبادل البيانات على الانترنت، والاستفادة من المعايير المفتوحة open standards، والواجهات المفتوحة open interfaces، وتطوير البرمجيات مفتوحة المصدر open source software. وتعد متطلبات تصميم قاعدة البيانات الافتراضية ما يلي:

١. دمج مستودعات البيانات الموزعة التي يتم تخزينها باستخدام نظم إدارة قواعد بيانات مختلف. بحيث لا تتأثر استقلالية المكونات الفردية لقواعد البيانات بعملية اتحاد البيانات.

٢. تقتصر وظائف قاعدة البيانات الافتراضية على الاستعلامات الموزعة. ويتم التعامل مع إدخال البيانات وتحديثها عن طريق التطبيقات البرمجية لقواعد البيانات الفردية.

٣. تصميم واجهات موحدة للوصول إلى البيانات.

٤. وينبغي أن يكون استرجاع المعلومات، والاستعلام، والتحليل وعرض البيانات من نظم قواعد البيانات الموزعة مفتوح لعدد كبير من المستفيدين، ويتم ذلك باستخدام مستعرض الويب Web Browser.[٣]

يرى الباحث أن متطلبات تصميم قواعد البيانات الافتراضية السابق عرضها تركز على مجموعة من المفاهيم، منها دمج البيانات مع المحافظة على استقلالية نظم إدارة قواعد البيانات، مما يساعد على تحقيق التجانس بين قواعد البيانات المختلفة من حيث التصميم، والبنية، وتمثيل البيانات، كما ان نظم إدارة قواعد البيانات الافتراضية لا تتعامل مع إدخال البيانات وتحديثها مما يحافظ على استقلالية قواعد البيانات الفردية، ويضاف لما سبق أن نظام إدارة قواعد البيانات يركز في متطلبات بناءه على استخدام واجهه موحدة للبحث، والاستعلام من خلال برنامج مستعرض الويب مما يساعد على استخدام قاعدة البيانات الافتراضية على الخط المباشر وبسهولة من قبل المستفيد العادي. وجميع المفاهيم التي تناولتها المتطلبات السابقة تصب في تحقيق نظم استرجاع المعلومات أهدافها من خلال توسيع دائرة البحث في أكثر من قاعدة بيانات فردية في آن واحد ومن خلال واجهة واحدة.

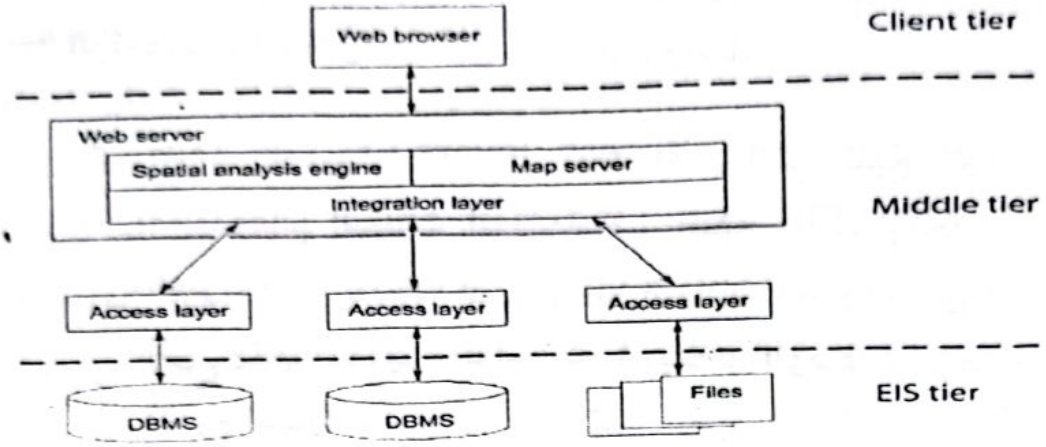


Figure 1: Architecture of the Virtual Database

شكل رقم (٣) يوضح مستويات بنية قاعدة البيانات الافتراضية.[٤]

ويتبع تصميم قاعدة البيانات الافتراضية وتنفيذها مبدأ الاقتران الحر لمكونات قواعد البيانات الفردية، بحيث يكون هذا الاقتران بطريقة منظمة في طبقات منفصلة بوضوح ولكنها مترابطة أيضاً. ويعرض الشكل رقم (٣) مكونات قواعد البيانات وعناصر وحدات البرمجيات اللازمة لقاعدة البيانات الافتراضية في ثلاثة طبقات منفصلة. هذه الطبقات هي كما يلي :

١. طبقة نظام المعلومات المؤسسة (طبقة EIS) Enterprise Information System Tier (EIS Tier): تتكون طبقة EIS من مستودعات البيانات الموزعة التي يجب أن تعمل في تكامل. ويتم تخزين البيانات في أنظمة إدارة قواعد البيانات أوكملفات. ٤٥ حيث تحتوي طبقة EIS في قاعدة البيانات الافتراضية من مستودعات البيانات غير المتجانسة من حيث المكونات الخاصة بكل قاعدة بيانات والمتمثلة في هيكلية البيانات، والتخزين، وانظمة إدارة قواعد البيانات (DBMS)، وايضاً موجودة في مواقع مختلفة.[٦]

٢. الطبقة المتوسطة Middle Tier: ويحتوي المستوى المتوسط على الواجهات من خلال طبقتين فرعيتين، وهما:

- طبقة الوصول access layers، والتي تتيح الوصول إلى مستودعات البيانات المتاحة في طبقة EIS. وتحدد الواجهات الطريقة التي يجب أن تمر بها البيانات من جهة إلى اخرى. هذ إلى جانب توفير الواجهات للبيانات الوصفية لإجراء عملية تقييم البيانات.

- طبقة التكامل integration layer، والتي تتحكم في الوصول لمستودعات البيانات الموزعة؛ وفي عملية دمج البيانات المسترجعة من طبقات الوصول access layers وذلك لتقديم رؤية واضحة للبيانات. وينفذ محرك التحليل المكاني spatial analysis engine عمليات التحليل المطلوبة . ويكون برنامج Map server software مسؤولاً عن تقديم خرائط للبيانات المتكاملة وتحليلها.

٣. طبقة العميل Client Tier: وتختص هذه الطبقة باستعرض المستفيد من بيئة الويب للبيانات المسترجعة من map server بطريقة تسهل عليه العرض والتفاعل مع البيانات.[٤٧]

يرى الباحث أن الطبقات الثلاث لقاعدة البيانات الافتراضية تتوافق مع متطلباتها، حيث تحقق طبقة EIS المتطلب الأول المتعلق بدمج مستودعات البيانات المختلفة دون التدخل في استقلاليتها، في حين تحقق طبقة Middle المتطلب الثاني المتعلق بالتعامل مع الاستعلامات وتوجيهها، واخيراً تحقق طبقة Client المتطلبين الثالث والرابع المتعلق بالواجهة الموحدة للمستفيد للاستعلام واسترجاع المعلومات. هذا على المستوى البرمجي، اما على المستوى المعلوماتي فهناك رؤية اخرى لهذه الطبقات الثلاث، حيث تمثل طبقة EIS المحتوى المعلوماتي الذي يرغب

المستفيد في الوصول إليه والبحث خلاله والذي توافر لأي نظام استرجاع معلومات، أما طبقة Middle فتمثل عملية المعالجة للاستفسار المعلوماتي بنظام استرجاع المعلومات حيث يتم فيها البحث عن المعلومات التي تتوافق والاستفسار البحثي، ثم عملية الدمج واستبعاد المتكررات، ثم تنظيمها لعرضها على المستفيد، وبالتالي هي المرحلة الوسيطة بين المستفيد والمحتوى المعلوماتي، واخيراً طبقة Client وتمثل واجهة نظام استرجاع المعلومات التي يتم من خلالها تقديم الاستفسار وتحديد محدداته لتضييق نطاق البحث المعلوماتي للوصول للنتائج التي تلبي احتياجات المستفيد، وهي ذات الواجهات التي يتم من خلالها عرض النتائج وتعديل البحث.

لذا فإن قواعد البيانات الافتراضية بتصميمها الحالي يحقق التجانس بين قواعد البيانات المختلفة، ويساعد على توسيع دائرة البحث المعلوماتي لتلبية احتياجات المستفيد المعلوماتية، وتوفير الوقت والجهد المبذول في تكرار عملية البحث في قواعد بيانات فردية، ثم المقارنة بين النتائج لاستبعاد المتكررات.

٣/٢/٢ مخطط قاعدة البيانات الافتراضية VDB Schema

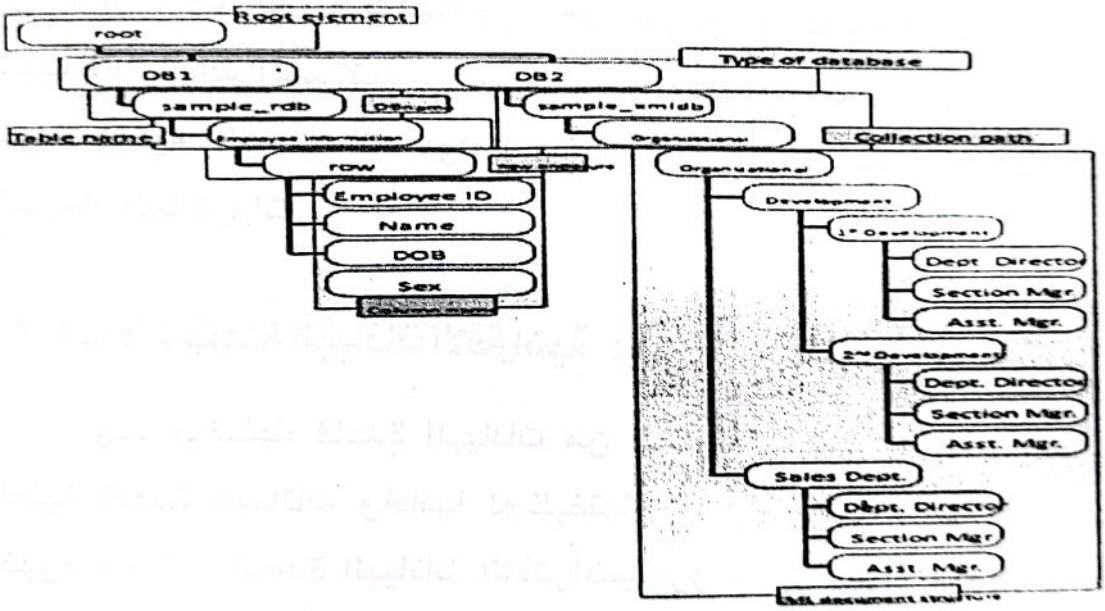
يعد مخطط قاعدة البيانات من المكونات الأساسية التي تؤثر على فاعلية قاعدة البيانات وأدائها لوظيفتها، لذا سيركز الباحث على عرض مفهوم مخطط قاعدة البيانات الافتراضية ووظائفه لإبراز دوره في عملية استرجاع المعلومات.

يعد الهدف من المخطط المشترك Common Schema Structure تقديم معلومات للمستخدم حول بنية قاعدة البيانات الافتراضية، كما يستخدم المخطط في فحص عبارات الاستعلامات وضوابطها وقيودها. حيث يرى المستخدمون مخطط واحد مشترك لجميع قواعد البيانات الموزعة.[٤٨] ومثال على ذلك قيام مجموعة من الباحثين

بتصميم مخطط عام لقاعدة البيانات الافتراضية يقدم وصف للبنية الكاملة لها في شكل XML، لذا تظهر للمستخدم على أنها بنية متكاملة لقاعدة بيانات افتراضية، ويتم تطبيق استعلامات المستفيدين بها باستخدام لغة الاستعلام XQuery."

ويتم بقاعدة البيانات الافتراضية تعريف مخططات قواعد البيانات الثابتة static والمتغيرةdynamic لتحديد قواعد الجمع، والبناء، والتحقق من الصحة بقاعدة البيانات ووضعها في مخطط الذاكرة المشترك المستخدم لتحقيق التوازن في التحميل والحفظ المؤقت. [٥٠]

ويوضح الشكل التالي هيكل مخطط مشترك لقاعدتي بيانات مختلفتين في النوع RDB وXMLDB.



شكل رقم (٤) يوضح يوضح هيكل مخطط مشترك لقاعدتي بيانات مختلفتين في النوع RDB وXMLDB .[٥١]

يوضح الشكل رقم (٤) بنية هيكل المخطط المشترك لقاعدتي بيانات مختلفتين في النوع؛ فالأولى DB1 تعتمد نموذج تمثيل البيانات

العلائقي RDB، بينما الثانية DB2 تعتمد نوذج تمثيل البياناتXMLDB. حيث نلاحظ ان العنصر الأساسي root element* يقع في اعلى المخطط ويمثل المستوى الأول، ثم يندرج في المستوى الادنى منه العناصر التي تشير إلى الأنواع المختلفة من قواعد البيانات مثل DB1، وDB2 ويمثل المستوى الثاني.

ونلاحظ في قاعدة البيانات DB1 التي تعتمد نظام تمثيل البيانات العلائقي RDB ان المستوى الثالث يحتوى على اسم قاعدة البيانات، ثم المستوى الرابع والذي يوضح آلية التخزين المتمثلة في الجداول يحتوي على بنية الجدول؛ والذي يتضمن تفصيلاً مستوى يشير إلى اسم الصف؛ ثم مستوى يشير إلى اسم العمود.

أما في قاعدة البيانات DB2 التي تعتمد نظام تمثيل البيانات XMLDB نلاحظ ان المستوى الثالث يحتوي على اسم قاعدة البيانات، ثم المستوى الرابع والذي يوضح مسار المجموعة، يليه المستوى الخامس الذي يشير إلى هيكل الوثيقة.[٥٢]

كما أن قاعدة البيانات الافتراضية تحتوي على مخطط الاستعلام المشتركCommon Schema Query، والذي يكون له وظيفة ذات بعد معرفي بقاعدة البيانات الافتراضية، ويشار لمخطط الاستعلام المشترك داخل كتابة المخطط المشترك Common Schema حتى يتمكن المستخدم من الحصول على البيانات وفق شكل تخزينها سواء من قاعدة البيانات RDB او قاعدة البيانات XMLDB.

علاوة على ذلك، يساعد مخطط الاستعلام المشترك على تحديد شروط استرجاع القيمة المخزنة التي ينطبق عليها شروط الاستعلام عن طريق تحديد المسار Path واستخدام شرط "Where"، كما يتم مقارنة

* صيغة تكتب بلغة البرمجة لتحديد بداية ونهاية أوامر البرمجة بملف قاعدة البيانات، كما انه في صيغة XML يحتوي على كل عناصر الوثيقة.

النتائج المسترجعة من قاعدة بيانات RDB، وقاعدة بيانات RDB اخرى؛ وكذا النتائج المسترجعة من قاعدة بيانات RDB، وقاعدة بيانات XMLDB؛ او النتائج المسترجعة قاعدة بيانات XMLDB، وقاعدة بيانات XMLDBاخرى، وذلك في حال انطباق شروط الاستعلام على المحتوى المسترجع.

وبالتالي فإن ظهور نتائج الاستعلام للمستخدم بمثابة مؤشر لتنفيذ المخطط المشترك لعملية البحث في قواعد البيانات الحقيقية الموجودة في الخلفية لقاعدة البيانات الافتراضية، وبالتالي فإن النتائج المسترجعة من قواعد البيانات الفعلية يمكن مقارنتها مع بعضها البعض دون ان يعلم المستخدم الاختلافات بينها.[٥٣]

ويرى الباحث أن العرض السابق يوضح الدور الذي يقوم به المخطط المشترك Common Schema في تحقيق هدف قاعدة البيانات الافتراضية من خلال إيجاد تجانس بين قواعد البيانات الحقيقية المختلفة في نماذج تمثيل البيانات لتحقيق مفهوم تكامل البيانات بهدف تسهيل عملية استرجاع المعلومات وتوحيدها للمستفيد وجعلها أكثر فاعلية.

٤/٢/٢ الاستعلام في قواعد البيانات الافتراضية واسترجاع المعلومات VDB Query & IF

يرى الباحث أن الاستعلام بقواعد البيانات الافتراضية يعد الآلية الأهم لتحقيق عملية استرجاع المعلومات، حيث يعد بمثابة الاستفسار المرجعي الذي يصيغه المستفيد للاسترجاع المعلومات التي تلبي إحتياجاته المعلوماتية، لذا سيركز الباحث في هذا الجزء من الدراسة على إلقاء الضوء على آلية تنفيذ الاستعلام بقواعد البيانات الافتراضية.

في حال رغبة المستخدم استرجاع بيانات من أنواع متعددة من قواعد البيانات، فإنه يصدر استعلام لنظام إدارة قواعد البيانات، ثم يقوم

بمراجعة نتائج الاسترجاع للحصول على المعلومات المطلوبة. بينما يُمكن التطوير الحديث لقاعدة البيانات الافتراضية المستفيد من استرجاع المعلومات ومقارنتها كما لو انه يتعامل في استرجاع المعلومات مع قاعدة بيانات واحدة دون ان يعلم المستفيد الاختلافات بين قواعد البيانات"

تعد قاعدة البيانات الافتراضية بمثابة وعاء لعدد من المكوّنات المركبة، والهدف الرئيس منها استرجاع بيانات متكاملة من مصادر بيانات متعددة، لذا يمكن توظيفها للوصول للبيانات باستخدام اسلوب متكامل من خلال واجهة استرجاع واحدة API (Application programming interface)*، حيث تحتوي قواعد البيانات الافتراضية VDB على نماذج تحدد الخصائص الهيكلية لمصادر البيانات وطرق العرض وخدمات الويب."

حيث يقدم المستفيد استعلام واحد في قاعدة البيانات الافتراضية بدلاً من تقديم عدد من الاستعلامات الفردية في كل قاعدة بيانات على حدى، ويتم تنفيذ استعلامات المستفيدين في قواعد البيانات الفردية من خلال لغات البرمجة المختلفة مثل SQL او XQuery وكذلك الأمر في الاستعلامات الأصلية بقاعدة البيانات الافتراضية."

اما آلية تنفيذ الاستعلام في قواعد البيانات في الواجهة الخلفية Database Backend، فيجب ان تعطى كل قاعدة بيانات في الواجهة الخلفية اسم مميز ومنطقي. ثم يتم جمع مخطط كل قاعدة بيانات من قواعد البيانات في الواجهة الخلفية تلقائيا عند إضافتها إلى قاعدة البيانات الافتراضية. ومع ذلك، يمكن تحديد مخطط قاعدة البيانات الثابتة static database schema لاستخدامها بدلا من ذلك. وأخيرا، يحدد مدير الاتصال المخصص specific connection manager استراتيجية تجمع

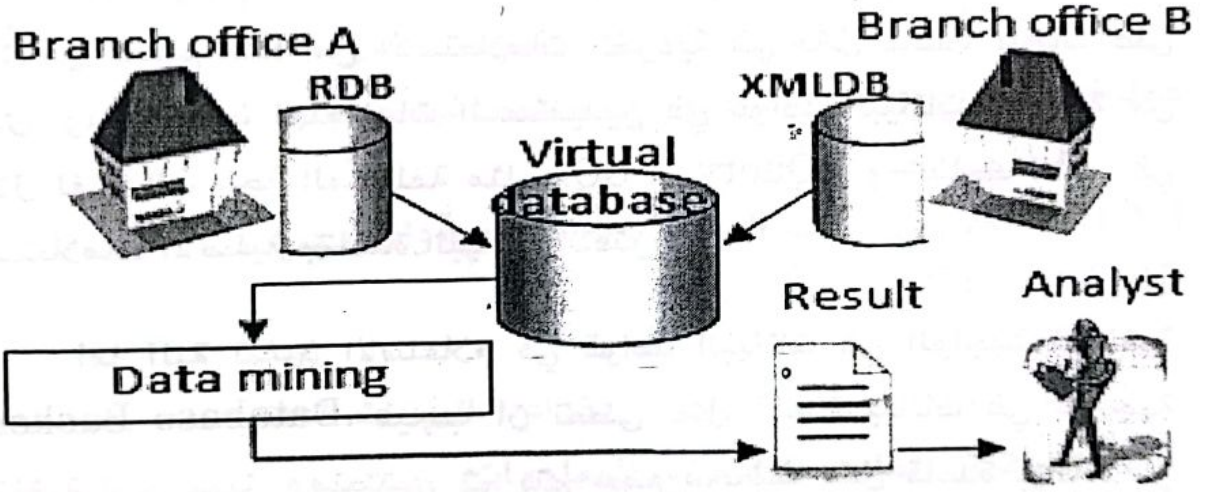
* لمزيد من المعلومات يمكن مراجعة:
http://en.wikipedia.org/wiki/Application_programming_interface

الاتصال لكل تسجيل دخول افتراضي على كل قاعدة بيانات في الواجهة الخلفية.[٥٧]

وقد اقترح Yuji Wada واخرون ثلاث انماط للاستعلام في قواعد البيانات الافتراضية حسب لغة البرمجة، وهذه الانماط كما يلي:

- استعلام XQuery

من خلال استخدام لغة الاستعلام XQuery بالإضافة إلى وظائف موسعة أكثر لم تكن تدعمها لغة الاستعلام XQuery من قبل، ووفق الشكل رقم (٥) التالي فانه بالرغم من ان لغة XQuery تدعم استعلامات مرنة إلا أن مشكلتها الأساسية ان بيان لغة الاستعلام سيكون طويل في حالة الاستعلامات المركبة والمعقدة، والتصور المحتمل يشمل التحكم في الملحقات والتحديث للوظائف والقيود المرتبطة بقاعدة البيانات العلائقية

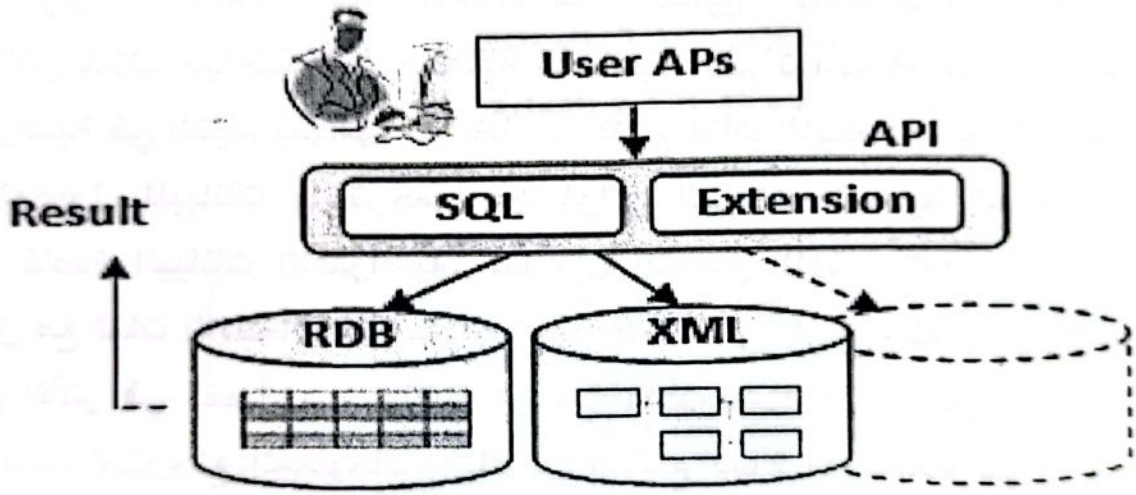


.RDB[٥٨]

شكل رقم (٥) يوضح آلية مقترحة للاستعلام بلغة XQuery في قاعدة البيانات الافتراضية.[٥٩]

- استعلام SQL

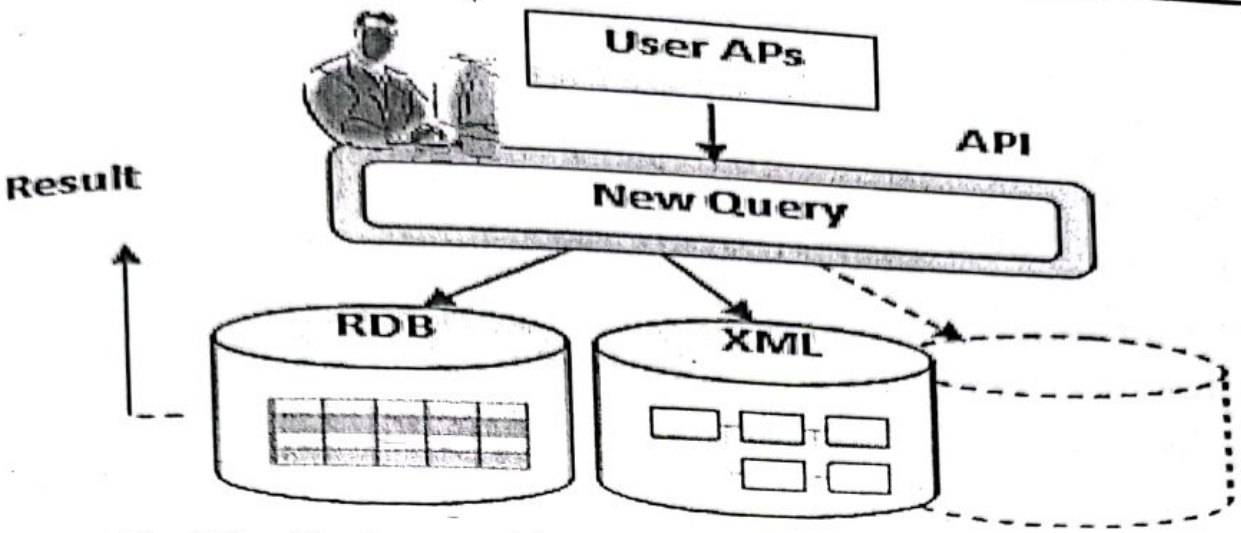
يدعم هذا الاقتراح استخدام لغة الاستعلام SQL بالإضافة إلى وظائف موسعة أكثر لم تكن تدعمها لغة الاستعلام SQL من قبل، وكما يتضح من الشكل رقم (٦) التالي يمكن كتابة الاستعلام بلغةSQL بعبارات بسيطة وقصيرة ولكنها تفتقر للمرونة، لذا يمكن اعتبار العبارات الإضافية كامتداد للاستعلام، والذي يعتبر وظيفة لإصدار الاستعلامات لمثل هذه العناصر في قاعدة بيانات XMLDB، ودالة للتمييز بين العناصر التي لها نفس باعتبارها نقطة جوهرية اساسية.[١٠]



شكل رقم (٦) يوضح آلية مقترحة للاستعلام بلغة SQL في قاعدة البيانات الافتراضية.[١١]

- لغة استعلام جديدة

ويرى Yuji Wada وزملائه الباحثين ان لكل من المقترحين السابقين للاستعلامات إيجابيات وسلبيات، لذا يوضح الشكل رقم (٧) التالي اقتراحهم لتطوير لغة استعلام جديدة تماماً كبديل عن اللغات المستخدمة، والمبنية على أساس انه طالما هناك انماط مختلفة من قواعد البيانات؛ فإنه من الممكن تصميم استعلامات قابلة للتطوير.[١٢]



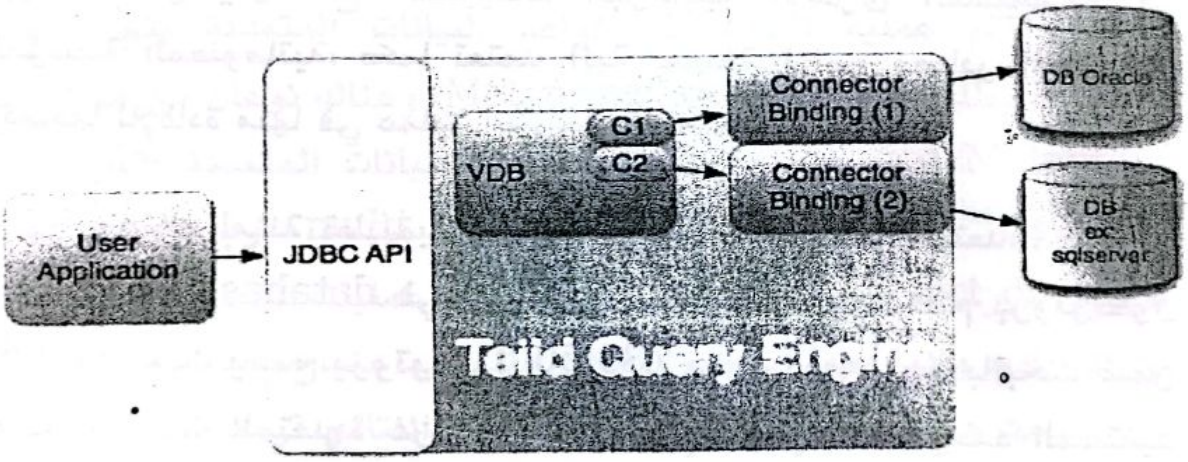
شكل رقم (٧) يوضح آلية تطوير لغة استعلام جديدة بقاعدة البيانات الافتراضية.[٦٣]

ويرى الباحث من الاستعراض السابق لمقترحات Yuji Wadaوزملائه الباحثين، أن المشكلة الأساسية التي تواجه قواعد البيانات الافتراضية في القيام بمهامها تتركز في تنوع لغات الاستعلام المستخدمة في الوصول للبيانات واسترجعها من قواعد البيانات الفردية الموجودة داخل قاعدة البيانات الافتراضية. لذا فإن التوصل للغة استعلام جديدة تتوافق مع لغات الاستعلام المستخدمة بقواعد البيانات المعروفة سيكون له عظيم الأثر في تفعيل دور قاعدة البيانات الافتراضية في قيامها بدورها في عملية استرجاع المعلومات، وبالتالي ارتفاع نسبة الاستدعاء في نظام استرجاع المعلومات.

وما يؤكد هذا الراي انه في بعض الأحيان قد يستلزم تنفيذ الاستعلام إعادة كتابة الاستفسارات على قواعد البيانات في الواجهة الخلفية Rewriting requests on backends، فإذا كانت قواعد البيانات المجمعة والتي يطلق عليها cluster تضم محركات قواعد البيانات من موردين متنوعين، فقد لا يتم فهم استفسار المستفيد من قبل كل قواعد البيانات في الواجهة الخلفية. فعلى سبيل المثال إذا كان التطبيق المستخدم معد بلغة البرمجة PostgreSQL وهناك رغبة في إضافة قاعدة بيانات للواجهة الخلفية بلغة MySQL فإن بعض

الاستفسارات تتطلب تعديل لتنفذ بشكل صحيح على MySQL. ويمكن تحديد قواعد لإعادة كتابة الاستفسارات على قواعد بيانات محددة بالواجهة الخلفية. ويحدد عنصر قواعد إعادة الكتابة Rewriting Rule element كيف يتم إعادة كتابة الاستعلامات المطابقة لنمط معين."

ومن أمثلة تطبيقات قواعد البيانات الافتراضية التفاعلية ما قامت به شركة *Teiid، حيث قامت بتصميم اداة يطلق عليه Eclipse قائمة على استخدام واجهة رسومية للمستخدم GUI لإنشاء قاعدة بيانات افتراضية. وتتيح هذه الأداة للمستخدم تحديد نماذج المصادر وإستيراد البيانات الوصفية والاحصائيات، كما تسمح للمستخدم بتحديد العلاقات وطريقة العرض لملف XML في وصف هذه المصادر، مما يمكن المستخدم من استخلاص بنية المعلومات المسترجعة واستخدامها في التطبيقات الخاصة به وفق البنية الفعلية لهيكلية البيانات. "



شكل رقم (٨) يوضح محرك بحث الاستعلامات لشركة Teiid. "

ويوضح الشكل رقم (٨) استخدم قاعدة البيانات الافتراضية الديناميكية Dynamic VDB ملف XML بسيط، يتولى جمع مصادر المعلومات ودمجها ثم إتاحة الوصول إليها من خلال JDBC

* شركة برمجيات من ضمن اهتمامتها تصميم قواعد البيانات الافتراضية.

(Java Database Connectivity)* حيث يمكن للمستخدم كتابة الاستعلام، وتتعامل قاعدة البيانات الافتراضية مع مصادر المعلومات المختلفة كأنها مصدر معلومات واحد، حيث لا تعرض قواعد البيانات الافتراضية الطبقات الدنيا من قاعدة البيانات.[٦٧]

ويرى Yuji Wada وزملائه الباحثين ان شركة Teiid لديها محرك للاستعلام فريد من نوعه. وعلاوة على ذلك، يتم إنجاز تكامل البيانات في الوقت الحقيقي من خلال ربط برامج تطبيقات الأعمال من خلال طبقة الوصول JDBC / SOAP مع مصادر البيانات التي يتم الوصول إليها من خلال بنية التوصيل.[٦٨]

ويرى الباحث ان اداة Eclipse التي صممتها شركة Teiid نموذجاً للأداة التفاعلية لتصميم قاعدة بيانات افتراضية تستطيع من خلالها المؤسسة المعلوماتية تحديد نماذج المصادر، والعلاقات، وطريقة عرض البيانات لتوظيفها مع التطبيقات البرمجية الأخرى المستخدمة في المؤسسة المعلوماتية، كما تعتمد آلية بسيطة لجمع مصادر المعلومات ودمجها للإفادة منها في عملية استرجاع المعلومات.

وهناك أمثلة تطبيقية للبحث في قواعد البيانات المتعددة Multi-database Searching في المؤسسة المعلوماتية باستخدام بروتوكول Z39.50، حيث يسمح بروتوكول Z39.50 بدون دعم مسبق بالبحث ضمن قواعد البيانات المتعددة عن استفسار بحثي فردي. فقد يحدد المستفيد إرسال الاستفسار البحث إلى قاعدة بيانات محددة. وبالتالي يتلقى الخادم طلب البحث ويوجهه تجاه قاعدة البيانات المحددة.

وقد يحدد الخادم الذي يدعم طلبات البحث في قواعد البيانات المتعددة أي قواعد البيانات يتم البحث فيها بشكل فردي وايها يتم البحث

* تطبيق جافا قياسي لربط مجموعة واسعة من قواعد البيانات المستقلة. انظر:
http://www.tutorialspoint.com/jdbc/jdbc-introduction.htm

فيه بشكل مجمع، فعلى سبيل المثال إذا كان الخادم يوفر 100 قاعدة بيانات فردية، مسماه بالتتابع من " 1 إلى 100" فقد يتم تهيئة الخادم انه يتم البحث في كل قواعد البيانات بشكل فردي ماعدا القاعدتين 1،2 يتم البحث فيهما معاً بشكل مجمع.

وقد يحدد الخادم قاعدة البيانات الافتراضية ليتم من خلالها دعم عملية التجميع للنتائج بدلاً من دعم طلبات البحث في قواعد البيانات المتعددة. فعلى سبيل المثال، إذا كان الخادم يوفر الوصول إلى ثلاث قواعد بيانات فعلية (A,B and C) يتم من خلالها تجميع البيانات المطلوبة؛ فإن الخادم قد يعرض سبع قواعد بيانات افتراضية (A, B, C, AB, AC, BC, and ABC)، وبالتالي يسمح يتوفير جميع التجميعات الممكنة ولكن يسمح بتحديد قاعدة بيانات افتراضية واحدة فقط لتنفيذ الطلب البحثي فيها.[٦٩]

ولتتم عملية البحث في قواعد البيانات المتعددة يتم توظف برمجيات يطلق عليها مرشحات Metaproxy*، وهناك نوعان من مرشحات Metaproxy تدير عمليات البحث بقواعد البيانات المتعددة multiple-database. وهما مرشح virt_db، ومرشح multi. ويمكن ان يعمل مرشح virt_db بمفرده للسيطرة على توجيه عمليات البحث لخادم محدد من الخوادم، بينما يعمل مرشح multi جنبا إلى جنب مع مرشح virt_db لأداء البحث في قاعدة البيانات المتعددة، ودمج النتائج في مجموعة موحدة من النتائج فيما يسمى metasearch in a box.

ويعد التفاعل بين المرشحين أمر معقد بالضرورة: بشكل يعكس الواقع والتعقيد غير القابل للاختزال للبحث في قواعد البيانات المتعددة باستخدام بروتوكول مثل بروتوكول Z39.50 الذي يفصل بين التهيئة والبحث، وفي اي قاعدة بيانات قابلة للبحث غير معرفة في وقت التهيئة.[٧٠]

* برنامج مخفي في قاعدة البيانات

ويعمل مرشح virt_db بمفرده، حيث أن الغرض منه توجيه طلبات البحث إلى واحدة من مجموعة مختارة من قواعد البيانات في الخلفية backend databases. وبهذه الطريقة، يمكن للبروتوكول Z39.50 endpoint من خلال تشغيل برنامج Metaproxy ان يوفر إمكانية الوصول إلى العديد من الخدمات الأساسية المختلفة، بما في ذلك تلك التي لا يمكن الوصول إليها بسبب برامج الجدران النارية firewalls. وتعد من التهيئات المفيدة لقواعد البيانات الافتراضية أن يتم تركيب Metaproxy في قواعد البيانات في الخلفية، وبالتالي يمكن استخدام خوادم Z39.50 في قواعد البيانات في الخلفية.[٧١]

على سبيل المثال، يمكن تعيين عامل الترشيح virt_db وبالتالي يتم توجيه البحث في قواعد البيانات الافتراضية عن موضوع "قانون العمل" إلى خادم الفهرس الببليوجرافي لمكتبة الكونجرس، ثم البحث في "MARC" قاعدة البيانات الافتراضية ثم إرسالها لقاعدة بيانات تسجيلات MARC التي تستضيف بيانات التكشيف لإجراء الاختبار عليها. ويضاف لما سبق أن المرشح يوفر أساس البحث بقاعدة البيانات المتعددة.[٧٢]

عندما يتلقى Metaproxy طلب Z39.50 من عميل، فإنه لا يحيل هذا الطلب فورا إلى خادم خلفية back-end server. وذلك لأن Metaproxy لا يحدد اي خادم من الخوادم الخلفية للبحث عن استفسار العميل حتى يحدد العميل أي قاعدة بيانات يرغب في البحث فيها عن استفساره، وبمجرد أن يحدد العميل قاعدة بيانات افتراضية محددة يتم البحث فيها؛ وخادم الاستضافة المناسب يتم توجيه طلب العميل لقاعدة البيانات، وفي حال تقديم نفس الاستفسار السابق على قاعدة بيانات افتراضية اخرى، وعلى ذات خادم الاستضافة يتم الإجابة على الاستفسار من خلال النسخة المخبأة في الخادم.[٧٣]

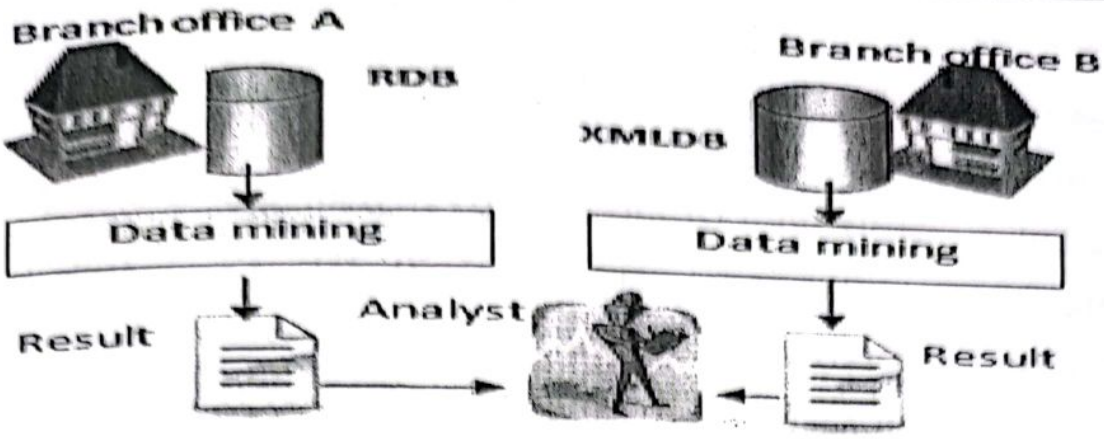
ويرى الباحث من العرض السابق، أن الإشكالية الأساسية للاستعلام في قواعد البيانات المتعددة تتمثل في، لغة الاستعلام المستخدمة في قاعدة

البيانات، وهيكلية البيانات، ودمج النتائج المسترجعة، وجميع هذه الاشكاليات تؤثر على نتائج استرجاع المعلومات. لذا فإن قواعد البيانات الافتراضية تعد الحل الامثل والفعال لهذه الاشكاليات، حيث انها تعالج إشكالية اختلاف لغة الاستعلام سواء بإعادة صياغة لغة الاستعلام وفق برمجيات محددة؛ او تنتج لغة استعلام جديدة تتوافق مع قواعد البيانات الفردية. كما تعالج إشكالية اختلاف هيكلية قاعدة البيانات، ودمج النتائج المسترجعة من خلال إجراء الاستعلام على مرحلتين، اولهما البحث في القواعد الفردية، وثانيهما دمج النتائج واستبعاد المتكررات مما يوفر الوقت والجهد على المستفيد، ويحقق اعلى ناتج استدعاء بنظام استرجاع البيانات.

٣/٢ التنقيب عن البيانات وقواعد البيانات الافتراضية Data Mining & VDB

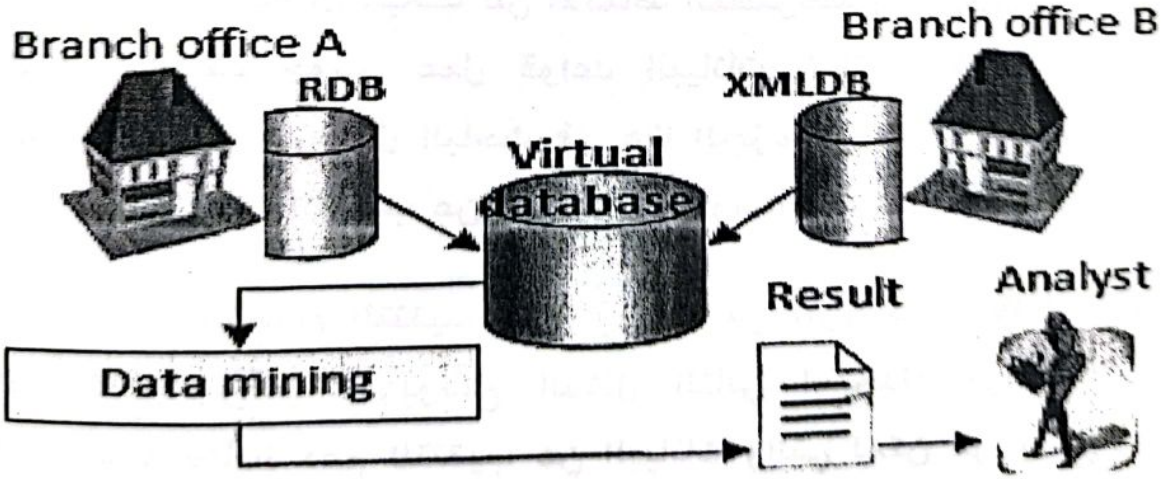
يرى الباحث أن مصطلح مصطلح "التنقيب عن البيانات Data Mining" يرتبط وظيفياً بمصطلح قواعد البيانات الافتراضية، والذي يعرف بمعجم ODLIS بأنه " عملية استخدام تطبيقات قواعد البيانات لتحديد انماط غير مكتشفة سابقا والعلاقات داخل مجموعة موجودة من البيانات".[٧٤] وحيث أن البحث عن الأنماط المشتركة والعلاقات بين مصادر المعلومات تعد جوهر عمل قواعد البيانات الافتراضية في استرجاع المعلومات. لذا سيحاول الباحث في هذا الجزء من الدراسة إلقاء الضوء على تطبيق آليات التنقيب عن البيانات في قواعد البيانات الافتراضية.

يعد استخدام التنقيب عن البيانات من أبرز نقاط فاعلية قواعد البيانات الافتراضية. ويوضح المثال التالي استخدام قواعد البيانات الافتراضية كأداة دعم للتنقيب من البيانات والتي تمكن من وضع تصور مقارن -بين البيانات المسترجعة من التنقيب عن البيانات، بل وتساعد محللي البيانات على فهم الوضع الحقيقي للبيانات.



شكل رقم (٩) يوضح تحليل البيانات من مصادر متعددة بالطريقة التقليدية .[٧٥]

يوضح الشكل رقم (٩) تجربة الطريقة التقليدية التي يتبعها محلل البيانات في تحليل البيانات الواردة من مصادرة متعددة، حيث يقوم محلل البيانات لأغراض التحقق باستخدام بيانات المعاملات الخاصة بالمستفيدين من مؤسسة ما لتحليلها. ووفقاً للشكل السابق فإن البيانات مخزنة في قاعدة بيانات علائقية RDB في الفرع A، وقاعدة بيانات XMLDB في الفرع B. وبالتالي فإن تنفيذ عملية تحليل البيانات تبدأ باسترجاعها من كلا القاعدتين ثم مقارنتها باستخدام ادواة التنقيب عن البيانات.



شكل رقم (١٠) يوضح تحليل البيانات من مصادر متعددة باستخدام قواعد البيانات الافتراضية .[٧٦]

ويوضح الشكل رقم (١٠) طريقة مختلفة لتحليل البيانات؛ حيث انه مع تطبيق معطيات الحالة السابقة ولكن مع إضافة متغير يتعلق بتنفيذ عملية التنقيب عن البيانات باستخدام قاعدة بيانات افتراضية، سيكون هناك اختلاف في المراحل التي تمر بها العملية، حيث يتم استرجاع البيانات من قاعدة البيانات الافتراضية التي تعد بمثابة واجهة لقواعد بيانات متعددة، وبالتالي يتم استرجاع البيانات من قاعدة بيانات واحدة بدلاً من نتائج متعددة من قواعد بيانات مختلفة، وبالتالي سيتم استرجاع نتيجة واحدة من عملية التنقيب عن البيانات، حيث لن يكون حاجة لإجراء اكثر من عملية تنقيب عن البيانات وبالتالي لن يكون هناك حاجة لمقارنة البيانات المسترجعة مما يسهل عملية تحليل البياناتعلى المحللين ووضع تصور فعلي لها.[٣٧]

ويرى الباحث من العرض السابق أن نظام إدارة قواعد البيانات الافتراضية يقوم بدور الوسيط لتحقيق التجانس بين قواعد البيانات الحقيقية باستخدام ادوات برمجية متمثلة في المخطط المشترك Common schema، و نموذج تحويل الاستعلام Query conversion module بهدف التعامل مع السمات البرمجية المختلفة لقواعد البيانات الحقيقية من حيث نماذج تمثيل البيانات، ولغات البرمجة، وذلك لتحقيق هدف محدد يتلخص في توفير وقت وجهد المستفيد في عملية إسترجاع المعلومات. وبالتالي فإن تطبيق مفهوم التنقيب عن البيانات من خلال توظيف قواعد البيانات الافتراضية يسهم بشكل فعال في تحقيق اهداف نظام استرجاع المعلومات

٤/٢ قواعد البيانات الافتراضية تكامل البيانات Data Integration & VDB

إذا كانت الفكرة الأساسية لقواعد البيانات الافتراضية قائمة على دمج محتوى قواعد البيانات الفردية بشكل افتراضي، فمن المهم إلقاء الضوء على مفهوم تكامل البيانات وعلاقته بقواعد البيانات الافتراضية. لذا

سيحاول الباحث في هذا الجزء من الدراسة تناول تقنية تكامل البيانات وتوظيفها في قواعد البيانات الافتراضية.

قد وصفت بعض التقارير المبكرة لدراسة تقنية المحاكاة الافتراضية لقاعدة البيانات database virtualization technology اقتراح لتطوير نظام يستخدم شبكة معلومات واسعة wide-area network لنشر المعلومات الخاصة بنشاط المستخدمين في بيئة الحوسبة المتنقلة mobile computing دون اي خلل في تبادل البيانات، وذلك من خلال اتصاله بمجموعة من انواع قواعد البيانات المختلفة التي تضم مصادر المعلومات التي تهم المستخدمين. ويتم ذلك من خلال تطبيق آلية نسخ صورة للبيانات المتوفرة في مجموعة قواعد البيانات المحلية local database group إلى قاعدة بيانات وصفية meta-database من خلال عمليات البحث الأساسية، وبناء العمليات البحثية، وعلى سبيل المثال، فإنه يمكن أن يجمع البيانات المتوفرة في قواعد البيانات المحلية ويدمجها.[٧٨]

وهناك ميزة خاصة لقاعدة البيانات الافتراضية كونها تعمل بمثابة منصة لدمج البيانات من مستودعات البيانات الموزعة التي تحتوي على البيانات البيئية. وقد اظهرت التجارب الأولى لقواعد البيانات الافتراضية مرونة عالية بشأن إدماج مستودعات البيانات غير المتجانسة. ويتم تحقيق التدرجية العالية للنظام من خلال آلية التخزين المؤقت على أساس تكرار البيانات.[٧٩]

وقد صفت شركة Teiid تقنية تكامل البيانات data integration technique بانها التمكين الافتراضي لأنواع مختلفة من قواعد البيانات، من خلال قواعد البيانات الافتراضية، حيث تمكن المستخدم من الوصول إلى مصادر البيانات في قواعد البيانات العلائقية، وقواعد البيانات على شبكة الإنترنت، وبرامج التطبيقات مثل ERP، وCRM في الوقت الحقيقي ككيان متكامل واحد للاستخدام.[٨٠]

وتعد أحد المنهجيات المستخدمة لتحقيق تكامل البيانات Data integration ما يطلق عليه الاتحاد Federated، حيث يعد حل لتكامل البيانات من خلال الإبقاء على البيانات في مصدرها الفردي والذي تمثله قواعد البيانات الفردية، وبالتالي يتم الحفاظ على البيانات بشكلها وتحديثها، ويتم ببساطة الرجوع إليها عند الحاجة. وفي هذه الحالة ستظهر المصادر المتعددة للبيانات لتكاملها في قاعدة بيانات افتراضية مع إخفاء الأنواع المختلفة لقواعد البيانات خلف واجهة موحدة.[٨١]

يمكن ان يكون تكامل البيانات الاتحادي Federated data integration معقد للغاية، في حال بيئة قواعد البيانات الموزعة غير المتجانسة والمتصلة عن بعد، حيث يتم الاتصال باستخدام بروتوكول two-phase commit؛ مرحلة البحث ثم مرحلة التكامل، واحد الحلول المطروحة لذلك أن يتم الوصول للبيانات وحفظها في مكان محدد مثل المخزن المركزي للبيانات. حيث يتم الوصول إلى البيانات "بشفافية" من خلال تفاعل المستخدم أو التطبيق مع قاعدة بياناتاتحادية أو قاعدة بيانات افتراضية يتم إدارتها من قبل برنامج نظم إدارة قواعد البيانات العلائقية (RDBMS) Relational Database Management Systems RDBMS الأساسي مثل Oracle. ثم يعمل برنامج تكامل البيانات مع برنامج نظم إدارة قواعد البيانات العلائقية الأساسي RDBMS في الخلفية لتحويل مخططات وقواميس البيانات وترجمتها للتوافق مع لغة SQL، وضمان توافق المرور خلال قواعد البيانات الخارجية باستخدام بروتوكول two-phase commit، ثم تصنيف المصادر إلى مصادر متجانسة وغير المتجانس، ثم عرض البيانات الموزعة من خلال شكل موحد على قاعدة بيانات واحدة.[٨٢]

وبالتالي يحتاج برنامج التكامل integration software حتى يستبعد المهام المعقد إلى ان يتكامل مع برنامج RDBMS الأساسي ليحقق الفائدة من الوظائف الداخلية للبرنامج وتنفيذ الاستعلامات بفاعلية. لذا

يجب على برنامج RDBMS توفير الوظائف الهامة لبرامج نظم إدارة قواعد البيانات لتحقيق الاستعلامات والاسترجاع بفاعلية.[٨٣]

وقد قام مجموعة من الباحثين بدراسة حول قاعدة البيانات الافتراضية باعتبارها بنية لدمج مستودعات البيانات الموزعة. بهدف تحقيق التصور المشترك لتكامل البيانات الموزعة، وتحليلها، وبناء نظام للمعلومات البيئية والمناظر الطبيعية الشاملة. وذلك من خلال إنشاء قاعدة بيانات افتراضية تدمج قواعد البيانات الموزعة المختلفة مثل Swiss Agency for the Environment, Forests and Landscape (SAEFL). بحيث يتم توفير البيانات من خلال بيئة تمكينية تمكن من الوصول إليها بسهولة واستكشافها وتحليل شامل لها. ويرى Brändli انه في البداية تم تصميم قاعدة بيانات افتراضية لغرض دمج البيانات والاستعلامات البسيطة.[٨٤]

ويرى الباحث أن مفهوم تكامل البيانات يسهم بشكل فعال في فاعلية عملية استرجاع المعلومات، من خلال البحث في مصادر المعلومات المتعددة بشكل تكاملي، وقد تحقق مفهوم تكامل البيانات بتطبيق تكنولوجيا قواعد البيانات الافتراضية التي ساهمت في الحفاظ على هوية قواعد البيانات الفردية واستقلاليتها من جانب، وفي من جانب اخر تكامل مصادرها بما يحقق اهداف نظم استرجاع البيانات المتمثلة في الوصول لكافة المعلومات التي تلبي احتياجات المستفيد بأسرع وقت ممكن.

٥/٢ قواعد البيانات الافتراضية وأدوات البحث على الإنترنت Internet Resaech Tools & VDB

يرى الباحث ان ادوات البحث على الإنترنت المتمثلة في محركات البحث، والبوابات الموضوعية ترتبط ارتباطاً وظيفياً بقواعد البيانات بوجه عام، كما انه من الممكن توظيف قواعد البيانات الافتراضية بمفهومها لتحقيق أعظم فائدة من أدوات البحث البحث على الإنترنت، وذلك من خلال توسيع دائرة البحث والإسترجاع بما يخدم المستفيد. لذا سيحاول

الباحث في هذا الجزء من الدراسة إلقاء الضوء على تطبيق آليات قواعد البيانات الافتراضية في أدوات البحث على الإنترنت كنظام لاسترجاع المعلومات.

تساعد تكنولوجيا قاعدة البيانات الافتراضية (VDB) على تحويل مصادر الإنترنت، ومصادر المعلومات الخارجية الأخرى إلى امتداد لقواعد البيانات العلائقية لنظام إدارة قواعد البيانات للمؤسسة (RDBMS). ووفقا لبعض التقديرات، فإن ما يقرب من ٩٠٪ من البيانات في العالم هو خارج نظم قواعد البيانات العلائقية. وتنتشر البيانات الحيوية عبر مواقع على شبكة الإنترنت، وملفات النظم، ونظم قواعد البيانات، والتطبيقات القديمة. وتختلف طريقة تنظيم هذه المصادر، ومفرداتها، وآليات الوصول إلى البيانات الخاصة بها. وكثير منها لا تدعم عمليات الاستعلام الأصلية. مما يعكس عدم التجانس لهذه المصادر.[٨٥]

حيث تحتوي شبكة الويب على العديد من مصادر المعلومات ذات البنية المنظمة والتي تحتوي على كميات هائلة من المعرفة، مثل كتب الحقائق التي تضم معلومات سياسية، واقتصادية، وجغرافية، عن كل بلدان العالم؛ مواقع التراجم التي تضم معلومات عن أكثر من خمس وعشرون الف شخصية شهيرة؛ وقواعد بيانات الأفلام التي تضم مئات الآلاف من الأفلام بما فيها من معلومات حول الممثلين، والمخرجين. والعديد من هذه المصادر جزء من "الويب العميق Deep Web " أو "الويب غير مرئية Invisible Web "، والتي لا يمكن الوصول إليها من خلال التصفح العادي؛ حيث يتم تخزين هذه المعلومات في قواعد بيانات لا يمكن الوصول إليها إلا من خلال واجهات بحث معينة. وليتم تحقيق الاستخدام الفعال لهذه المصادر في الرد على الاستفسارات يجب أن تكون متكاملة تحت واجهة مشتركة او لغة استعلام موحدة، وهذا ما يمكن تحقيقه من خلال قاعدة البيانات الافتراضية.[٨٦]

وتقوم تكنولوجيا قواعد البيانات الافتراضية VDB بتحويل كافة مصادر البيانات في قاعدة بيانات ملائقية افتراضية واحدة. باستخدام تطبيق مبني على تكنولوجيا قواعد البيانات الافتراضية VDB. [٨٧] حيث تعد اهم ميزات قاعدة البيانات الافتراضية انها تتيح لأي مستفيد باستخدام متصفح ويب واتصال بشبكة الإنترنت تنفيذ عمليات بحث شامل عن البيانات مع إمكانيات تحليل متقدمة، وقد اتضح ان اسلوب الواجهات الموحدة الذي تم اختياره للوصول إلى البيانات الموزعة يتمتع بمرونة عالية حيث لا يتطلب تغييرات او تعديلات في مستودعات البيانات المشاركة في قاعدة البيانات الافتراضية. [٨٨]

حيث ان الهدف من تطوير تقنية التمثيل الافتراضي لقاعدة البيانات من أجل السماح لمحللي البيانات أو المستخدمين الآخرين بتطبيق أساليب استخراج البيانات في وظائفهم من خلال استخدام كافة قواعد البيانات المتوفرة على شبكة الإنترنت كما لو انها قاعدة بيانات واحدة ، وبالتالي مما يساعد على تقليل أعباء العمل الخاصة بهم مثل جمع البيانات من قواعد بيانات الإنترنت.. [٨٩]

ويرى Katz and Others انه على الرغم من أن شبكة الويب تحتوي على كمية هائلة من المعلومات، إلا أن الحصول على المعلومات الصحيحة امر بالغ الصعوبة نتيجة الافتقار للأساليب الفعالة للوصول للمعلومات وفصل المعلومات المفيدة للمستفيد عن غيرها؛ هذا إلى جانب عدم وجود بنية موحدة للمعلومات على الويب. لذا فقد اتجه Katz and Others إلى تقديم حل لهذه المشكلة من خلال تحويل الويب إلى قاعدة بيانات افتراضية يتم البحث فيها والوصول للبيانات باللغة الطبيعية. وذلك من خلال تطوير إطار مبسط للعلاقات stylized relational framework يطلق عليه object-property-value model الذي يقوم بتكشيف الكلمات التي ترد في الأستفسارات الموجهه لمحرك البحث باللغة الطبيعية، والكلمات الواردة في مصادر الويب. [٩٠]

حيث ظهرت انظمة Question answering في وقت مبكر من القرن الحادي والعشرين تقوم بالإجابة على الاستفسارات وتوفير المعلومات بطريقة اكثر يسراً على المستفيد وبطريقة تختلف عن النموذج التقليدي لنظم استرجاع المعلومات التي تعتمد على صياغة الاستفسارات وعرض الوثائق المسترجعة، حيث تتلقى "نظم الرد على الاسئلة" الاستفسارات باللغة الطبيعية للمستخدم وتكون الإجابات مختصرة. [٩١] ويعتمد نموذج البيانات object-property-value model على تنظيم البيانات وتجانسها، والمصادر المهيكلة وشبه المهيكلة، حيث يقوم النموذج بتكشيف المحتوى الدلالي للإستفسارات باللغة الطبيعية الموجهه من المستفيدين. [٩٢]

ويرى Katz and Others ان التحديات الثلاث الرئيسية للرد على الاستفسارات باللغة الطبيعية تتمثل في فهم السؤال باللغة الطبيعية؛ وتحديد مكان العثور على المعلومات؛ واسترجاعها، لذا قام Katz and Others بتطوير نظام يمر بمرحلتين لمعالجة هذه التحديات اولهما START natural language question answering system لفهم اسئلة المستفيدين، وترجمتها من خلال structured bjectproperty-value queries، ثم Omnibase التي تعد قاعدة بيانات افتراضية تمثل واجهة موحدة لمصادر المعرفة على الويب وقادرة على تنفيذ الاستعلامات المبنية بواسطة START. [٩٣]

ويرى الباحث ان العرض السابق ناقش اهمية توظيف تكنولوجيا قواعد البيانات الافتراضية في الوصول لمصادر المعلومات على شبكة الإنترنت، كما تناول قضية البحث باللغة الطبيعية في قواعد البيانات الافتراضية المبنية على قواعد بيانات الويب، وغيرها من المصادر المتاحة على الويب، مما يعكس الدول الوظيفي الفعال الذي من الممكن ان تلعبه قواعد البيانات الافتراضية في استرجاع المعلومات من شبكة الانترنت.

اما على مستوى محركات البحث، فنجد ان محركات البحث المتعددة Meta Search Engine تواجه قضية التجميع لمصادر المعلومات، وتحديات تقنية متمثلة في تصنيف وترتيب مصادر المعلومات واستبعاد الرسائل المزعجة spam والشوشرة noise. وقد تناول الباحثون المتخصصون في استرجاع المعلومات واسترجاع البيانات هذه القضايا بالدراسة المتعلقة بالجمع بين النتائج من قواعد بيانات متعددة ومجموعات الوثائق او من مؤشرات متعددة عن نفس البيانات. ويمكن تطبيق هذه المعرفة في البحث والتجميع.[١١]

ويرى الباحث انه من المهم لتناول قضية البحث بمحركات البحث المتعددة ضرورة إلقاء الضوء على مفهوم البحث الموحد Federated Search، باعتباره مصطلح يستخدم لوصف الأعمال البحثية ذات الصلة بمحركات البحث المتعددة في مجال علم المكتبات والمعلومات والتي توظف عادة بروتوكول Z39.50 في استرجاع المعلومات. ويعد معيار SRU (Search/Retrieve via URL) البحث والاسترجاع باستخدام العنوان الموحد لموقع الويب أحدث برتوكول معياري لتمثيل الاستفسارات بمكتبة الكونجرس، وقد قارن LeVan بين SRU وOpenSearch.[١٠]

ويمكن تلخيص المزايا المحتملة من البحث المتعدد على النحو التالي:

- تقديم واجهة واحدة لمصادر متعددة
- تبحث في العديد من المصادر بقواعد بيانات محركات البحث التي قد لا يعلم عنها الباحث
- تقدم كل المعلومات المطلوبة في الاستفسار البحثي، حيث لا توجد قاعدة بيانات تضم مجموعة مصادر معلومات بمحرك بحث واحد تحوي كل المعلومات المطلوبة

- الحد من الوقت الذي يقضيه الباحث في عملية البحث
- يمكنها إضافة مستخلصات او مقارنة المصادر المسترجعة
- التقليل من الإعلانات الموجودة بمحركات البحث العادية.[٩٦]

ويرى الباحث أن المزايا السابق ذكرها تتفق مع آلية عمل قواعد البيانات الافتراضية، كما تساهم في تحقيق فاعلية نظم استرجاع المعلومات بالشكل الذي يلبي رغبات المستفيد واحتياجاته.

ولإبراز العلاقة بين محركات البحث المتعددة وقواعد البيانات الافتراضية، نجد أن محركات البحث المتعددة Meta-search engines تنشئ ما يطلق عليه قاعدة بيانات افتراضية لمعالجة الاستعلام السريعة من خلال دمج استعلامات متعددة وإرسالها لمصادر متعددة. وقد واجه الباحثون Madhavan واخرون في Google تحدي البحث المتعدد meta-search والبحث الموحد federated search بدلاً من استخدام كشاف كبير واحد في مجموعة اجهزة حاسب. وقدم Madhavan واخرون حجة مضمونها أنه مع عدد الاستفسارات في محرك البحث الرئيس، فمن الأهمية أن يتم إرسال الاستفسارات ذات الصلة إلى مواقع الويب العميقة deep web sites، وعلى وجه آخر فإن نسبة المرور العالية قد تؤدي إلى تعطل الموقع crash the sites. وعند التعامل مع قضية مرحلة معالجة الاستعلام يمكن اختيار محركات البحث المناسبة. وقد اكد الباحثون أن المنهجية الافتراضية تجعل محرك البحث يعتمد في آداءه لعملية البحث على مصادر الويب العميقة. ومن المهم أن يكون محرك البحث متعاون ومحرك البحث المتعدد يتجاهل المصادر غير المستجيبة.[٩٧]

اما على مستوى البوابات الموضوعية، فقد وضع Chandrinos و Vasarhelyi تصور لتطبيق البوابات الموضوعية subject gateway واستخدامها في التعامل مع مصادر المعلومات عن بعد باستخدام قاعدة

البيانات الافتراضية الاتحادية Federated Virtual Database عبر الويب.

وقد اعتمد الباحثان على منهجية أن موقع الويب الأساسي Primary Web Site (PWS) يعد بمثابة "بوابة ذكية شفافة intelligent transparent gateway " إلى المعلومات ذات الصلة. وذلك بافتراض أن مواقع المصادر الثانوية Secondary Resource Sites لديها حرية نشر المصادر بالشكل الذي يرغبون فيه، لذ اقترح الباحثان بناء طبقة وسيطة في مواقع الويب الأساسية PWS للإجابة على الاستعلامات من خلال ما يسمى برامج التضمين wrappers والذي يتولى الاتصال بمواقع المصادر الثانوية SRS (Secondary Resource Sites) ويكون برنامج التضمين wrapper قادر على التواصل مع SRS محددة، واستخراج المعلومات المطلوبة وتمريرها مرة أخرى إلى الوسيط لتحقيق التكامل. ولكن تظهر مجموعة من المشكلات يمكن تحديدها فيما يلي:

- تتجه برامج التضمين wrappers إلى أن تكون ثابتة، بحيث يمكن وضعها وتنفيذها فقط مع مصادر معروفة وموجودة بالفعل، من خلال إعطاء "لقطة مجمدة frozen snapshot " للنظام. ومن الواضح، انه لا يمكن إنشاء مجمع لأنواع من المصادر في المستقبل

- تتطلب برامج التضمين wrappers مبرمجين على مستوى عالي، إلا إذا كانت المجمعات شبه آلية semi-automatic، او في حال مقدمي المحتوى البعيد remote content providers قبل الاتفاق على مخطط صيغة تخزين المعلومات. في كثير من الأحيان هذا الحل غير قابل للتطبيق.[٩٨]

وبوجه عام، تساعد تكنولوجيا قواعد البيانات VDB برامج التطبيقات في التعامل مع:

- الأعداد الكبيرة من مصادر البيانات

- مصادر البيانات المستقلة والتي لا يتم التحكم فيها مركزياً

- مصادر البيانات المتنوعة من حيث البنية ما بين البيانات المهيكلة وغير المهيكلة

وجميع الخصائص السابقة تظهر في شبكة الإنترنت، وفي معظم الشبكات الداخلية. ومع ظهور لغة XML والمعايير ذات الصلة مثل RDF ستساعد على سرعة انتشار قواعد البيانات الافتراضية VDB لما لديها من قدرة على خفض تكلفة دمج مواقع الإنترنت وصفحاتها على قواعد البيانات الافتراضية.[١١]

ويرى الباحث أن العرض السابق يعكس دور قواعد البيانات الافتراضية في الافادة من ادوات البحث على الانترنت كنظم استرجاع معلومات، حيث يمكن من خلال توظيف تكنولوجيا قواعد البيانات الافتراضية الافادة من مصادر الويب التي لا يتمكن المستفيد والمؤسسة المعلوماتية من الوصول إليها مما ينعكس سلباً على عملية استرجاع المعلومات.

٦/٢ قواعد البيانات الافتراضية والأنظمة الآلية Automation Library system & VDB -

تعد النظم الالية في المكتبات من أهم ادوات استرجاع المعلومات التي يستخدمها المستفيد في الوصول لمصادر المعلومات التي توفرها المؤسسة المعلوماتية، لذا سيحاول الباحث في هذا الجزء من الدراسة إلقاء الضوء على علاقة النظم الآلية بالمؤسسة المعلوماتية بتكنولوجيا قواعد البيانات الافتراضية.

هناك العديد من المشروعات التي تبنتها مؤسسات معلوماتية لربط مصادر المعلومات بها من خلال توظيف تكنولوجيا قواعد البيانات الافتراضية، وسيحاول الباحث عرض نماذج منها سواء على مستوى

المؤسسة المعلوماتية، او على مستوى الشركات العاملة في مجال برمجيات المؤسسات المعلوماتية، او على مستوى الباحثين في مجال المكتبات والمعلومات.

على المستوى المؤسسي قامت مختبرات Los Alamos National Laboratory (LANL) الأمريكية والمعروفة باهتمامها بعلوم الحاسب بوضع تصور لمشروع بحثي يهدف لتحويل مكتبتها البحثية الرقمية إلى "مكتبة بلا جدران LWW "Library Without Walls بحيث تسمح للباحثين من كافة ارجاء العالم بالوصول لمصادر معلوماتها المحلية من خلال استخدام تكنولوجيا الحاسب عبر الاتصال بشبكة الإنترنت من أي مكان وفي أي وقت.[١٠٠]

في اواخر عام 1997 قامت مختبرات (LANL) بتطوير الجيل الثاني من مشروع مكتبة بلا جدران LWW، وشمل هذا التطوير بناء "قاعدة بيانات ضخمة mega-database"، الهدف منها محاولة جمع قواعد البيانات العلمية التي تقتنيها المكتبة في قاعدة بيانات افتراضية واحدة. ويعد السبب في هذا التطوير التغذية الراجعة للباحثين الذين لايملكون القدرة على التمييز الدقيق بين اختلافات قواعد البيانات العلمية من حيث الموضوع أو حدود التغطية، إلى جانب الوقت الذي يقضيه الباحث للعثور على احتياجاته المعلوماتية من خلال البحث في العديد من قواعد البيانات.[١٠١]

وكذا اتخذ مجلس مكتبة كاليفورنيا للقانون عام 2001 قرار بتنفيذ مشروع لتمويل تطوير روابط Z39.50 الغرض منه تحديد البنية التحتية اللازمة لإنشاء قاعدة بيانات افتراضية على مستوى الولاية التي تمثل مقتنيات المكتبات بكاليفورنيا. وتعد قاعدة البيانات الافتراضية عنصر هام في مجموعة خدمات الوصول المطبقة في مكتبة ولاية كاليفورنيا للقانون.[١٠٢]

كما ركزت مكتبة ولاية كاليفورنيا على تطوير ودعم شبكة الاتصالات السلكية واللاسلكية لتمكين جميع المكتبات من تقاسم مصادرها مع بعضها البعض. وبما يسمح بروتوكول Z39.50 بالبحث في فهارس الخط المباشر لمكتبة واحدة او عدة مكتبات من خلال شبكة الإنترنت، بغض النظر عن البرمجيات والأجهزة المتوفرة في المكتبات ضمن الشبكة. والهدف من قرار مجلس مكتبة ولاية كاليفورنيا بمواصلة وتوسيع برامج الدعم، وبرامج الوصول للمضيف تسهيل الوصول لتطوير شبكة فهارس المكتبات الافتراضية التي يمكن لمستفيدي المكتبة استخدامها للوصول لمصادر المعلومات في كل المكتبات المشتركة في شبكة المكتبات.[١٠٣]

ويلاحظ الباحث أهتمام المؤسسات المعلوماتية بمفهوم فهارس المكتبات من خلال شبكة الإنترنت بهدف ربط مصادر المعلومات للمؤسسات المعلوماتية لتحقيق أعلى فائدة منها في عملية استرجاع المعلومات بما يخدم المستفيد، وهو الأمر الذي يمكن تحقيقه من خلال توظيف تكنولوجيا قواعد البيانات الافتراضية بالمؤسسة المعلوماتية.

أما على مستوى الشركات العاملة في مجال برمجيات النظم الآلية بالمكتبات مثل شركة SHAREit، وشركة Auto-Graphics، فقد قامت هذه الشركات بإتاحة وكيل تقاسم المصادر AGent Resource Sharing يوفر منصة مفتوحة تساعد على تحقيق التقدم الهائل في تطوير نظم تقاسم المصادر للجيل القادم من المكتبات. وذلك من خلال تيسير عملية الترابط بين الأنظمة المختلفة، والقدرة على توفير واجهات تفاعلية للستعارة بالمكتبات، ونظم تقاسم المصادر. حيث يمكن للمكتبات مع اختلاف انظمتها الآلية المحلية تقاسم مصادرها مع بعضها البعض، بغض النظر عن حجم مجموعاتها او انظمتها الآلية. حيث يستوعب وكيل تقاسم المصادر الأنظمة الآلية المختلفة من خلال توفير فرص متكافئة لجميع أعضاء اتحاد المكتبات، ويمكن تنفيذه في واحدة من ثلاث بيئات، اعتماداً على احتياجات اتحاد المكتبات، وهي كما يلي:

- قاعدة البيانات الموحدة Physical union database

من خلال دمج فهارس المكتبات الاعضاء في الاتحاد على الانترنت في فهرس موحد.

- نظام قاعدة البيانات الافتراضية Virtual database system

من خلال تمكن البحث الاتحادي federated searching ، وبروتوكول Z39.50 و البحث المبني على HTTP من خلال قواعد بيانات افتراضية ثم دمج النتائج واستبعاد التكرارات وحفظ النتائج.

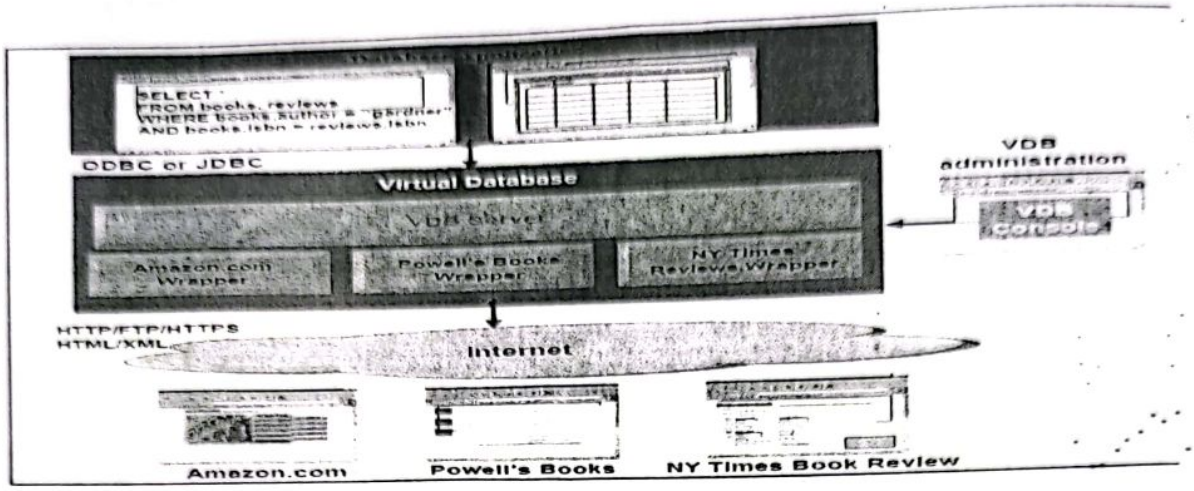
- نظام قاعدة البيانات المختلطة Hybrid database system

حيث يسمح وكيل تقاسم الموارد للمكتبات الأعضاء في اتحاد المكتبات بالمساهمة بتسجيلاتهم في قاعدة البيانات الموحدة، و يسمح للمكتبات الاعضاء بالعمل ضمن نظام افتراضي في بيئة واحدة.[١٠٤]

ويلاحظ الباحث أن الشركات العاملة بمجال برمجيات النظم الالية بالمؤسسات المعلوماتية قامت بإتاحة حلول متنوعة لتقاسم مصادر المعلومات، ومن ضمن هذه الحلول توظيف قواعد البيانات الافتراضية، مما يؤكد دور تكنولوجيا قواعد البيانات الافتراضية في تحقيق مبداء تقاسم المورد بين المؤسسات المعلوماتية، مما ينعكس على استرجاع المعلومات وخدمة المستفيد.

اما على مستوى الباحثين في مجال المكتبات والمعلومات، فقد قام الباحثان Prasad، وRajaraman بتصميم نموذج قاعدة بيانات افتراضية بسيطة VDB لاسترجاع الكتب المخزنة في VDB كما في الشكل رقم (١١). حيث تقوم قاعدة البيانات الافتراضية في هذا النموذج VDB بدمج محتويات مكتبات Amazon.com، وPowell's Books، و New York Times Book Reviews وعرض مخطط موحد، مع اثنين من الجداول، وعروض الكتب. حيث يقوم تطبيق قاعدة البيانات بعمل برنامج المخطط

الموحد، وإصدار استعلامات SQL من خلال JDBC او ODBCAPI، ويمكن بناء تطبيق itelf باستخدام ادوات RAD القياسية مثل Delphi، Visual Basic ، PowerBuilder، او Java toolkits.[١٠٥]



شكل رقم (١١) يوضح نموذج قاعدة بيانات الكتب الافتراضية.[١٠٦]

ويوضح الشكل رقم (١١) انه يتم الوصول إلى قاعدة البيانات الافتراضية VDB من خلال VDB Server، ويتم إدارتها من خلال مستعرض browser مبني على وحدة تحكم VDB Console. كما تحتوي قاعدة البيانات الافتراضية VDB على برنامج تضمين wrapper لتوجيه مصادر البيانات الخارجية إلى VDB Server. حيث يتولى برنامج التضمين wrapper توجيه مصادر البيانات الخارجية مثل مواقع الإنترنت ليتم التعامل معها من خلال قاعدة البيانات الافتراضية كما يتم التعامل مع نظم قواعد البيانات العلائقية RDBMS، في حين يدمج VDB Server قواعد البيانات العلائقية المنفصلة في قاعدة بيانات افتراضية موحدة (VDB).[١٠٧]

ويتعامل برنامج التضمين Wrapper مع مواقع الإنترنت المصممة بلغات HTTP، وHTML، اوXML، حيث يتعامل مع القضايا المتعلقة بالبروتوكولات مثل forms، وcookies، وauthentication. ويمكن الوصول لبرنامج التضمين من خلال تطبيق الواجهة الموحدة JDBC API،

والذي يمكن للمستفيد من خلالها تقديم الاستعلام بلغة SQL. حيث ان الاستعلام المصاغ بلغة SQL في برنامج التضمين Warpper يتم تعبئته في نموذج بلغة HTML على موقع Amazon.com، ثم يتم الابحار وتحليل صفحات HTML التي تظهر في نتائج الاستعلام وتحويلها إلى الجدول العلائقي. حيث يستخدم برنامج التضمين عدد من القواعد، منها قواعد لغوية متطورة لاستخراج سمات صفحات الويب، كما يستخدم قواعد لتحويل البيانات وتنسيقها بما يتناسب ومخطط البيانات الموحد، وايضاً قواعد التحقق من صحة البيانات لضمان سلامة البيانات.[١٠٨]

تستخدم التطبيقات المتطورة التي تتعامل مع عدد قليل من مصادر البيانات الوظائف الكاملة لنظام إدارة قاعدة البيانات الافتراضية VDBMS، حيث يبني نظام إدارة قاعدة البيانات الافتراضية جداول افتراضية لمصادر البيانات المتعددة في قاعدة بيانات افتراضية فردية VDB، كما تدعم الوظائف الكاملة لقاعدة البيانات العلائقية RDBMS في الجداول الافتراضية والتي تشمل عرض التعاريف، معالجة الاستعلامات للمصادر.[١٠٩]

وكما في الشكل رقم (١١) تقوم قاعدة البيانات الافتراضية VDB بإظهار الكتب لداري نشر amazon، وpowell بشكل موحد في جداول افتراضية. وعندما يتلقى نظام إدارة قاعدة البيانات الافتراضية استعلام، يتم تحليل الاستعلام ثم إرسال مفرداته لمصادر المعلومات الفردية، ثم يتم استرجاعها مجمعة للجمع بين نتائجها. كما يتم الاحتفاظ بنسخة مخباءة من النتائج، بالإضافة إلى أن نظام النشر يقوم بعمل لقطات دورية للجداول الافتراضية ويحتفظ بها في مخزن بيانات علائقي خاص بقاعدة البيانات لزيادة سرعة الوصول للبيانات.[١١٠]

ويلاحظ الباحث ان الجهود على مستوى الباحثين في مجال الحاسبات ركزت على بناء نماذج لقواعد بيانات افتراضية يتم تطبيقها على مؤسسات معلوماتية في استرجاع المعلومات، وهو ما يؤكد على

فاعلية تطبيق تكنولوجيا قواعد البيانات الافتراضية في استرجاع المعلومات بالمؤسسة المعلوماتية.

ويرى الباحث أن العرض السابق يشير إلى وجود جهود على المستوى العالمي من خلال المؤسسات، والشركات العاملة في مجال برمجيات المؤسسات المعلوماتية، والباحثين في توظيف تكنولوجيا قواعد البيانات الافتراضية في خدمة المؤسسة المعلوماتية بوجه عام، ونظم استرجاع المعلومات بوجه خاص.

١/٣ الخلاصة

ناقش البحث قضية قواعد البيانات الافتراضية وتوظيفها في قطاع المكتبات والمعلومات، إنطلاقاً من فرضية أن " قواعد البيانات الافتراضية يمكن توظيفها في نظم استرجاع المعلومات وآلياتها بما يخدم المستفيدين من خدمات استرجاع المعلومات بالمؤسسة المعلوماتية ". وخلص الباحث إلى أن تكنولوجيا قواعد البيانات الافتراضية يمكن توظيفها بالمؤسسات المعلوماتية لخدمة أغراض استرجاع المعلومات. كما يوصي بمزيد من الدراسات العربية في هذا المجال لما له من فائدة على تخصص المكتبات والمعلومات.

الهوامش :

[1]ODLIS: Online Dictionary of Library and Information.- Access in October 2013.- available at: http://dictionary.reference.com/browse/database?s=t&ld=1173.

[2]ODLIS: Online Dictionary of Library and Information.- Access in October 2013.- available at: http://www.thefreedictionary.com/database

[3] FOLDOC. Free on-line dictionary of computing.- Access in October 2013.- available at: http://foldoc.org/database

[4]ODLIS. Loc. Cit.- Access in October 2013.- available at: http://www.thefreedictionary.com/database

[5] Tech Terms.- Access in October 2013.- available at: http://www.techterms.com/definition/database

[٦] احمد الشامي.معجم مصطلحات المكتبات والمعلومات والارشيف.- متاح في : /http://www.elshami.com

[٧] المصدر السابق.

[8]ODLIS. Loc. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_N.aspx#nichedatabase

[9]Op. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_E.aspx#embeddeddatabase

[10]Op. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_M:aspx#metadatabase

[11] FOLDOC. Loc. Cit.- Access in October 2013.- available at: http://foldoc.org/object-oriented+database

[12]Op. Cit.- Access in October 2013.- available at: http://foldoc.org/database+management+system

[١٣] احمد الشامي.مصدر سابق.

[14]FOLDOC. Loc. Cit.- Access in October 2013.- available at: http://foldoc.org/hierarchical+database

[15]Op. Cit.- Access in October 2013.- available at: http://foldoc.org/network+database

[16] احمد الشامي،مصدر سابق.

[17] المصدر السابق.

[18]FOLDOC. Loc. Cit.- Access in October 2013.- available at: http://foldoc.org/relational+database

[19]ODLIS. Loc. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_D.aspx#dbmanagement

[20]FOLDOC. Op. Cit.- Access in October 2013.- available at: http://foldoc.org/database+management+system

[21] احمد الشامي،مصدر سابق.

[22]FOLDOC. Loc. Cit.- Access in October 2013.- available at: http://foldoc.org/database+management+system

[23]Op. Cit.

[24]Op. Cit.

[25]Op. Cit.

[26]Op. Cit.- Access in October 2013.- available at: http://foldoc.org/intelligent+database

[27]Op. Cit.- Access in October 2013.- available at: http://foldoc.org/virtual

[28] ODLIS. Loc. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_V.aspx#virtuallibrary

[29] Masunaga, Y., Watanabe, C., "Design and implementation of a multi-modal user interface of the Virtual World Database system (VWDB)," Database Systems for Advanced Applications, 2001. Proceedings. Seventh International Conference on , vol., no., pp.294,301, 21 April 2001 .- Access in November 2013.- available at: URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=916390&isnumber=19788

[30]ODLIS. Loc. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/searchODLIS.aspx

[31]Op. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_V.aspx#virtuallibrary

[32]Op. Cit.- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_V.aspx#virtualtour

[33]TheFreeDictionary.- Access in October 2013.- available at: http://encyclopedia2.thefreedictionary.com/virtual+database

See also: http://www.yourdictionary.com/virtual-database

http://www.pcmag.com/encyclopedia/term/55825/virtual-database

[34] Masunaga, Y.; Watanabe, C., "Loc. Cit.

[35]Virtual database configuration. Access in November 2013.- available at: http://c-jdbc.ow2.org/current/doc/userGuide/html/ar01s11.html

[36] Wada, Yuji, others. Virtual Database Technology for Distributed Database in Ubiquitous Computing Environment.- American Journal of Database Theory and Application 2012, 1(2).- Access in November 2013.- available at:
http://article.sapub.org/10.5923.j.database.20120102.02.html .P14

[37] Virtual Databases. Access in November 2013.- available at: http://www.jboss.org/teiid/basics/virtualdatabases.html

[38] Wada, Yuji, others. Loc. Cit .P15

[39] Op. Cit.

[40]

http://article.sapub.org/image/10.5923.j.database.20120102.02_001.gif

[41]

http://article.sapub.org/image/10.5923.j.database.20120102.02_005.gif

[42] Wada, Yuji, others. Loc. Cit .P16

[43] Frehner , Marcel , Brändli , Martin, Schenker , Jürg. The Virtual Database – A Tool for Integrated Data Processing in a Distributed Environment.- EnviroInfo 2004 (Geneva).- Access in November 2013.- available at: http://enviroinfo.eu/sites/default/files/pdfs/vol109/0537.pdf.- P 540

[44] Op. Cit.

[45] Op. Cit.- P 541

[46] Op. Cit.- P 542

[47] Op. Cit.- P 541

[48] Wada, Yuji, others. Loc. Cit.-PP15-16

[49] Op. Cit. - P19

[50]Virtual database configuration. Loc. Cit.

[51]
http://article.sapub.org/image/10.5923.j.database.20120102.02_007.gif

[52] Wada, Yuji, others. Loc. Cit.-P18

[53] Op. Cit.- P20.

[54] Op. Cit .P24

[55] Virtual Databases. Loc. Cit.

[56] Wada, Yuji, others. Loc. Cit.-P19.

[57]Virtual database configuration. Loc. Cit.

[58] Wada, Yuji, others. Loc. Cit.-P21.

[59]http://article.sapub.org/image/10.5923.j.database.20120102.02_016.gif

[60] Wada, Yuji, others. Loc. Cit.-P21>

[61]http://article.sapub.org/image/10.5923.j.database.20120102.02_018.gif

[62] Wada, Yuji, others. Loc. Cit.-PP21-22.

[63]
http://article.sapub.org/image/10.5923.j.database.20120102.02_018.gif

[64]Virtual database configuration. Loc. Cit.

[65] Virtual Databases. Loc. Cit.

[66] Op. Cit.

[67] Op. Cit.

[68] Wada, Yuji, others. Loc. Cit.-PP13-14.

[69] National Information Standards Organization. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification.-

Published by NISO Press.- Access in December 2013.- available at: http://www.loc.gov/z3950/agency/Z39-50-2003.pdf

[70] Dickmeiss, Adam, Cromme, Marc, Taylor, Mike. Metaproxy - User's Guide and Reference.- October 4, 2013.- Access in December 2013.- available at:http://www.indexdata.com/metaproxy/doc/metaproxy.pdf.- P23

[71] Op. Cit.

[72] Op. Cit.- PP23-24.

[73] Op. Cit.- P27.

[74]ODLIS. Loc Cit. .- Access in October 2013.- available at: http://www.abc-clio.com/ODLIS/odlis_D.aspx#datamining

[75]

http://article.sapub.org/image/10.5923.j.database.20120102.02_014.gif

[76]

http://article.sapub.org/image/10.5923.j.database.20120102.02_015.gif

[77] Wada, Yuji, others. Loc. Cit.- P21.

[78] Op. Cit.- PP13-14.

[79] Frehner , Marcel , Brändli , Martin, Schenker , Jürg. Loc. Cit.- P 539.

[80] Wada, Yuji, others. Loc. Cit- PP13-14.

[81] Laszewski, Tom, Williamson, Jason, Nauduri, Prakash. Getting Started with Information Integration.- Access in November 2013.- available at: http://www.oracle.com/technetwork/articles/datawarehouse/orcl-info-integration-chap-1-494739.pdf.- P11

[82] Op. Cit.- P13.

[83] Op. Cit.- P14.

[84] Frehner , Marcel , Brändli , Martin, Schenker , Jürg. Loc. Cit.- P 539.

[85] Prasad, STS and Rajaraman, Anand. Virtual Database Technology, XML, and the Evolution of theWeb.- IEEE Computer Society Technical Committee on Data Engineering .- Volume 21, Number 2, June 1998.- Access in March 2014.- available at:

http://www.ics.uci.edu/~wscacchi/DBM-EC/Course/Virtual-DB-Web.pdf .- P1

[86] Katz, Boris, and Others. Viewing the Web as a Virtual Database for Question Answering. In Mark T. Maybury, editor, New Directions in Question Answering. Cambridge, Massachusetts: MIT Press, 2004, Chapter 17, P4.- Access in March 2014.- available at: http://www.umiacs.umd.edu/~jimmylin/publications/Katz_etal_NewDirections2004b.pdf

[87] Prasad, STS and Rajaraman, Anand. Loc. Cit.- P2.

[88] Frehner , Marcel , Brändli , Martin, Schenker , Jürg. Loc. Cit.- PP 546-547.

[89] Wada, Yuji, others. Loc. Cit.-P13.

[90] Katz, Boris, and Others. Loc. Cit.- P2.

[91] Op. Cit.- P3.

[92] Op. Cit.- P4.

[93] Op. Cit.- P5.

[94] O'Riordan, Adrian P. Open Meta-search with OpenSearch: A Case Study.- Access in December 2013.- available at:http://www.cs.ucc.ie/~adrian/Metasearch-OpenSearch-AORiordan.pdf

[95] Op. Cit.

[96] Op. Cit.

[97] Op. Cit.

[98] Chandrinos, Konstantinos, Vasarhelyi, Pal.COURSEWARE FOR TRAINING OF TRAINERS AND USERS ON THE SPECIAL APPLICATIONS OF INTERNET-BASED SERVICES IN THE FIELDS OF CULTURAL EDUCATION: CHAPTER 2.CULTURAL SUBJECT GATEWAYS ON THE WEB ACCESSING DIGITIZED INFORMATION VIA NEW INTERNET-BASED TOOLS.- P7.- UNESCO.- Access in November 2013.- available at:http://www.mch.mii.lt/UNESCOeducation/chapters/chapter2.pdf.

[99] Prasad, STS and Rajaraman, Anand. Loc. Cit.- P5.

[100] Luce, Richard E. Integrating the Digital Library Puzzle:The Library Without Walls at Los Alamos.- Access in December 2013.- available at: http://lib-www.lanl.gov/la-pubs/00412932.pdf

[101] Op. Cit.

[102] ISSUE TO COME BEFORE THE BOARD AT THIS MEETING: Consideration of funding for Z39.50 server software grants.- Access in November 2013.- available at:http://www.library.ca.gov/loc/board/agendas/TelecommReports/TelecomInfraZ39-jun00.pdf

[103] Op. Cit.

[104] AGent Resource Sharing™ .A robust interlibrary loan (ILL) and consortial borrowing solution.- Access in November 2013.- available at: http://www.libraryworks.com/featuredlibrary/FL_EastHamptonLibrary.pdf

See also: http://www4.auto-graphics.com/docs/SHAREit_Brochure.pdf

[105] Prasad, STS and Rajaraman, Anand. Loc. Cit.- PP 1-2.

[106] Op. Cit.- P2.

[107] Op. Cit.- PP 2-3.

[108] Op. Cit.- P3.

[109] Op. Cit.

[110] Op. Cit.- P2.