6

الفصل السادس
نظم قواعد البيانات
ومستودعات البيانات التعليمية

1 - المقدمة

2 - لماذا نظم قواعد البيانات؟

3 - المفاهيم الأساسية لنظام إدارة قاعدة البيانات

4 - نماذج بيانات قواعد البيانات

5 - تصميم قاعدة البيانات

6 - بزوغ نظم مستودعات البيانات

7 - مفهوم نظم مستودعات البيانات

8 - أنواع نظم مستودعات البيانات

9 - معمارية مستودع البيانات

1 - المقدمة

تحتاج المنظمات التعليمية المختلفة لمعلومات دقيقة ومتوازنة وفورية لاتخاذ قرارات رشيدة. ويعنى ذلك أن البيانات يجب أن تجمع وتعالج وتخزن بطريقة ملائمة حتى يمكن توليد وخلق معلومات تساعد رؤساء ومديرى ورؤساء الأقسام والمدرسين... الخ، فى المنظمات التعليمية وفى اتخاذ قراراتهم بفعالية وكفاءة. والمعلومات المنتجة لا يجب أن تكون موثوق منها وفورية فقط؛ بل يجب أيضا أن يكون ممكن الوصول إليها. وبعبارة أخرى يجب أن تكون المعلومات معالجة ومخزنة بطريقة ملائمة تستخدم وسائط ذات سعات عالية تخزن هذه المعلومات فى شكل قواعد بيانات إلكترونية.

وكما سبق تحديده فى الفصل السابق عن "نظم دعم القرار التعليمية" التى لا تعتبر جيدة وسليمة إن لم ترتكز على نظم قواعد بيانات قوية وجيدة أيضا، التى من خلالها يمكن تداول بياناتها والوصول إليها بكفاءة.

ولكى يقدم نظام قاعدة البيانات بيانات صحيحة، يجب وجود هيكل يحل محل عمليات المعالجة القديمة التى تختص بتخزين كميات كبيرة من البيانات وتداولها للوصول إليها لاسترجاع معلومات منها، ويعتبر نظام إدارة قاعدة البيانات الصيغة الهيكلية لذلك التى تغذى نظام دعم القرار بمدخلاته من البيانات. كما أن تصميم وإدارة قاعدة البيانات الذى يعتبر عملية معقدة وطويلة فى كثير من الأحيان غير مدرك فى الواقع الفعلى. فحتى الآن لا يوجد فى معظم المدارس قواعد بيانات معتمدة على برمجيات نظم إدارة قاعدة البيانات المتقدمة DBMSs وتتضمن على

سبيل المثال، بيانات عن الطلاب بكل صف من صفوف أى مدرسة، كما تشتمل على بيانات عن المعلمين المتوافرين بكل مدرسة حسب تخصصاتهم مما يساعد على تحديد نسبة العجز الحقيقى فى عدد المعلمين الذى يؤدى للتوصل إلى تحديد العدد المطلوب أو النموذجى لسد النقص الذى تعانى منه كثير من المدارس بصورة حقيقية وملموسة فى كل تخصص. وفى هذا الإطار يؤكد نظام إدارة قاعدة البيانات الجيد سلامة البيانات ويقلل من تكرارها وإسهابها ويرتبط بتتابع منطقى ومتوافق فى أدائه.

وقواعد البيانات المنطقية تعتبر فى كثير من الحالات البنية التحتية التى تبنى عليها نظم المعلومات الإدارية ونظم دعم القرار بسبب مرونتها المرتبطة بالعلاقية والشيئية التى تسمح بالتقنين والمعايرة للحد من تكرار البيانات وصيانتها، حيث إن قدرة قواعد البيانات فى تعريف العلاقات بين الكيانات تجعل المعلومات تمثل أحد أوجه نجاح النظام المستخدمة فيه. ولا تزال حتى الآن تستخدم بعض نماذج قواعد البيانات الهرمية والشبكية، إلا أنها باهظة التكاليف إلى حد كبير.

وحيث إن قاعدة البيانات تتضمن وصلا حيويا يرتبط بقدرات أى نظام معلومات إدارية أو نظام دعم القرار، فإن هيكلتها وتصميمها يجب أن تقيم وتنفذ بعناية فائقة تراعى التطبيق الذى تسانده.

وفى الحقبة الحديثة من عقد التسعينيات من القرن الماضى ومستهل القرن الحادى والعشرين الحالى المرتبط بالتطور التكنولوجى المعتمد على شبكات المعلومات وخاصة شبكة الإنترنت، وكذلك من زحم نظم قواعد البيانات الموزعة وتضخم حجم البيانات بزغت نظم مستودعات البيانات التى تخزن كميات كبيرة من المعلومات وتعالجها على الخط تفوق ما تتعامل معه نظم قواعد البيانات التقليدية. وارتبط هذا التوجه الحديث باحتضان نظم دعم القرار التى تتسم بسمات التحليل المتعمق الممكن استيعابه فى نظم مستودعات البيانات فى نطاق المعالجة التحليلية على الخط OLAP .

وعادة قد تقدم قاعدة بيانات المنظمة التعليمية كالمدرسة معلومات حديثة عنها لتحديد التصرفات القائمة بها، إلا أنها تفشل فى تقديم معلومات ثرية وتاريخية التى تكون أكثر أهمية لعملية اتخاذ القرارات. لذلك ظهرت نظم دعم القرار لكى تغطى هذه الفجوة من خلال إدخال المعلومات التشغيلية وعرضها فى شكل له معنى ومدلول أعم باستخدام نظم إدارة قاعدة بيانات علاقية ترتبط بعمليات نظم دعم القرار. وعلى ذلك تتواجد نظم قواعد البيانات ونظم دعم القرار معا لتقديم مخرجات تسهم فى اتخاذ القرار وتساند النظم المبنية عليه. كما أدى التطوير المتسارع فى مجالات قواعد البيانات ومستودعات البيانات إلى استخدام نظم الوكيل الذكية Intelligent Agents التى تساعد فى سرعة وسهولة التساؤلات العشوائية التى قد تتواجد فى هذه النظم.

ويستعرض هذا الفصل الموضوعات التالية: مكونات نظام إدارة قاعدة البيانات، نماذج بيانات قواعد البيانات، تصميمك قاعدة البيانات؛ بزوغ نظم مستودعات البيانات، وأنواع نظم مستودعات البيانات ومعمارياتها ... الخ.

مع العلم بأن موضوع نظم قواعد البيانات قد حظى من جانبنا باهتمام كبير يتمثل فى الأعمال التى نشرت من قبلنا عن هذا الموضوع فى الربع قرن الماضية. ففى الثمانينيات من القرن الماضى ومن خلال عدة مقالات اهتم بتصميم قواعد البيانات بصفة عامة وارتباطها بشبكات المعلومات، كما أن كل المطبوعات الخاصة بنظم المعلومات الإدارية التى صدرت فى أواخر الثمانينيات وفى التسعينيات اشتملت على فصول ترتبط بقواعد البيانات سواء الهيكلية، المرتبطة بالنص وإدارة الوثائق والوسائل أو الوسائط المتعددة.

2 - لماذا نظم إدارة قاعدة البيانات؟

1/2 نظم إدارة قاعدة البيانات:

نظام إدارة قاعدة البيانات DBMS هو حزمة برمجيات مستخدمة لأداء عدد من

العمليات على قاعدة بيانات مثل إدخال البيانات واستبعادها وتحديثها واسترجاعها... الخ. والغرض الرئيسى لنظام إدارة قاعدة البيانات هو السماح للمستخدم فى تخزين البيانات وتحديثها واسترجاعها فى شكل مختصر وجعل المعلومات سهلة الحفظ والاسترجاع من قاعدة البيانات. ويعفى نظام إدارة قاعدة البيانات المستخدم من المعرفة المتعمقة عن العروض الطبيعية الخاصة بالبيانات وتحديد الألجوريثمات المفصلة الخاصة بتخزين البيانات وتحديثها واسترجاعها.

وعلى ذلك فإن نظام إدارة قاعدة البيانات هو حزمة برمجيات تقوم بأداء مهام مختلفة كثيرة تتضمن إمداد التسهيلات لمساعدة المستخدم فى الوصول للمعلومات وتعديلها فى قاعدة البيانات. وتمثل قاعدة البيانات وصلة وسيطة بين قاعدة البيانات الطبيعية والكمبيوتر ونظام تشغيله، ومن جهة أخرى مع المستخدمين. ولتقديم تسهيلات عديدة لأنواع مختلفة من المستخدمين يقدم نظام إدارة قاعدة البيانات برمجة متخصصة أو أكثر يطلق عليها فى الغالب لغات قاعدة البيانات.

وتأتى لغات قاعدة البيانات فى أشكال مختلفة، ويحتاج إلى اللغة لوصف قاعدة البيانات لنظام إدارة قاعدة البيانات بالإضافة لتقديم تسهيلات لتغيير قاعدة البيانات ولوصف وتغيير هياكل البيانات الطبيعية، ويحتاج أيضا إلى لغة أخرى لتداول البيانات المخزنة فى نظام إدارة قاعدة البيانات واسترجاعها. ويطلق على هذه اللغات لغات وصف البيانات(Data Description Languages (DDL ولغات تداول البيانات Data Manipulation Languages (DML).

نظم إدارة قاعدة البيانات DBMSs تعتبر جيدة جدا فى تنظيم وإدارة مجموعات بيانات متواصلة كبيرة. وتستخدم للمساعدة فى التغلب على كميات بيانات كبيرة؛ لأنه عندما تصل المشكلات إلى حد كبير، فإنه يصعب حلها. وبذلك فإن استخدام مجموعة أشياء كثيرة غير منظمة، يصبح غير ممكن الاستفادة منها فى الأداء واتخاذ القرارات؛ حيث إن هيكل نظام إدارة قاعدة البيانات يحول البيانات إلى معلومات

مفيدة تعطى مؤشرات تساعد على الأداء. ويؤدى ذلك إلى مثابرة النظام واستمراريته فى تواجد البيانات بصفة دائمة مما يسهم لعدم اختفاء البيانات عند إغلاق الحاسب الآلى.

2/2 وظائف نظم إدارة قاعدة البيانات:

تقوم نظم إدارة قاعدة البيانات بالوظائف التالية:

1. **حماية البيانات** التى تتعامل معها من إمكانية وصول المستخدمين المعتمدين باستخدامها مما قد يؤدى إلى فقد البيانات ذاتها أو شيوع سريتها بسبب:

- فشل الحاسبات الآلية كتوقف التيار الكهربائى أو تعطل الأجهزة والملحقات الخاصة بها.
- فشل البرمجيات كتوقف وتعطل برمجيات التشغيل أو برمجيات التطبيق المختلفة.

2. **الوصول المتزامن** الذى يعنى أن مجموعة بيانات مفردة يمكن الوصول إليها بواسطة أكثر من مستخدم فى نفس الوقت، مما يؤدى إلى التالى:

- بصفة افتراضية تتطلب كل تطبيقات قاعدة البيانات أن يكون لمدخلى البيانات فيها وصولا إليها من قبل كل المستخدمين الحاليين والمتوقعين معا فى نفس الوقت. على سبيل المثال، نظام معلومات شئون الطلاب لا يقتصر الوصول إلى قاعدة بياناته من قبل مستخدم واحد فقط.
- يقدم وصول البيانات المتزامن مشكلات غير مرغوب فيها مسببة بواسطة تداول مستخدمين البيانات بها فى نفس الوقت مما قد يؤدى إلى فشل قاعدة البيانات أو تؤدى لمشكلات فى تشغيل برنامج واجهة التفاعل مع المستخدم لعدم إكمال تساؤله، كما تتشابه هذه المشكلات لتمهيد نقاط تقاطع فى مسار البرنامج.

- تقدم آليات لمنع مشكلات الوصول المتزامن التى يطلق عليها رقابة التزامن.

3. نظام إدارة قاعدة البيانات الموزع يسمح بتقسيم أو تجزيء قاعدة البيانات إلى أقسام على حدة كتلك المتواجدة فى مواقع جغرافية متفرقة.

- إمكانية تقديم تحسينات أداء بواسطة استبعاد إرسال البيانات عبر قناة اتصال عن بُعد قد تكون بطيئة نسبيا ويعتبر من الأسرع تواجد قاعدة بيانات على قمة الكمبيوتر المستخدم بدلا من الوصول إليها عبر معيار شبكة الإيثرنت Ethernet، أو من خلال الموديم Modem.
- إمكانية تقليل اختناقات رقابة التزامن بإعطاء كل مستخدم جزء قاعدة البيانات المحتاج إليه فقط بدلا من إكمال وصول كل المستخدمين إلى قاعدة البيانات كلها فى نفس الوقت.

4.عدم اختصاص نظم إدارة قاعدة البيانات بوظيفة تحليل البيانات بالضرورة؛ حيث يكون ذلك مقصورا أكثر على برمجيات الجداول الإلكترونية أو أدوات تحليل الغرض الآخر، وبذلك فإن نظم إدارة قاعدة البيانات هى التى تكون:

- موجهة لتحقيق غرض معين، لا يرتبط أساسا بما هو مختزن به من قاعدة بيانات. وتقترح معظم مبادئ التصميم على ضرورة فك تزاوج مجال حزم تحلى معينة من نظام إدارة قاعدة البيانات لجعل تقسيم العمل واضح بقدر كاف.
- جيدة جدا فى استرجاع جزء صغير نسبيا من قاعدة البيانات وتمريره للتحليل المتصل أداة مصممة لهذا الغرض.
- تسمح غالبا بسلامة القيود المفروضة على البيانات لتأكيد الصحة والتوافق، وقد تتعارض هذه النظم مع التحليل العشوائى الذى يتداول فيه المستخدم البيانات بدون أى أفكار متصورة نسبيا فى كيفية ترابط البيانات معا.
- لا تشتمل غالبا على تسهيلات ملائمة لأداء الحسابات المعقدة أو لا تشتمل عليها كليا.

3/2 مزايا وعيوب قاعدة البيانات:

(1) المزايا:

تشتمل قاعدة البيانات المبنية على الكمبيوتر على المزايا التالية:

1. **تقليل إسهاب البيانات**: Redundancy بصفة تقليدية، فإن كل إدارة أو قسم فى أى منظمة تعليمية كالمدرسة يشتمل على ملفاته وسجلاته التى قد تتكرر فى إدارة أو قسم آخر فى المنظمة. ومع قاعدة البيانات المتكاملة أو المركزية تخزن البيانات فى مكان واحد ويمكن لكل الإدارات أو الأقسام فى المنظمة الوصول إلى البيانات منه فى أداء المهام بطريقة مقننة غير مكررة لحد ما.
2. **تجنب عدم مطابقة البيانات**: Inconsistency إذا قام كل فرد أو قسم أو إدارة فى المنظمة بتحديث ملفاته وسجلاته، قد توجد بيانات غير متطابقة أو غير متوافقة، حيث إن نفس البيانات قد تحدث فى نطاق إدارة معينة، ولكنها لا تحدث فى إدارة أخرى، كما أن بعض التحديث قد يكون غير صحيح. وعلى ذلك، فإن قاعدة البيانات المركزية سوف تستبعد هذه البيانات غير المطابقة.
3. **السماح بالمشاركة فى البيانات**: Information Sharing تسمح قاعدة البيانات بمشاركة بيانات تطبيقات أو إدارات عديدة بجعل استخداماتها الجديدة متوافرة فوريا.
4. **حفظ استقلالية المعلومات**: Information Independence تسمح قاعدة البيانات باستقلالية البيانات لتطبيقات أو منصات معينة بدلا من الأخرى. ويستبعد ذلك الحاجة لإعادة إدخال البيانات مما يعتبر مهما فى تقديم المرونة المطلوبة.
5. **الرقابة على الوصول**: **Access Control** بحفظ كل البيانات فى قاعدة بيانات كبيرة يمكن رقابة الوصول إليها، أى يمكن إعطاء مستخدمين مختلفين حق الوصول مما يؤكد اعتماد وصولهم إليها.

بالإضافة إلى الخمس مزايا السابق ذكرها، توجد ثلاثة أوجه رئيسية لنظام إدارة قاعدة البيانات التى تجعله جذابا ومقبولا فى الاستخدام من النظم الأخرى. وتتمثل هذه الأوجه فى التالى:

- إدارة بيانات مركزيا.
- استقلالية البيانات.
- تكامل النظم والتطبيقات

فى نظام قاعدة البيانات، تدار البيانات بواسطة هذا النظام، ويتم الوصول للبيانات من خلال تقديم مفتاح لتفعيل معالجة البيانات، ويختلف ذلك مع نظم معالجة البيانات التقليدية، حيث يكون لكل برنامج تطبيق وصولا مباشرا للبيانات التى تقرأها أو تتداولها. وفى برامج معالجة بيانات التطبيق التقليدية، تبنى هذه البرامج فى العادة على معرفة متخصصة بهيكل وهيئة البيانات. وفى هذه البيئة فإن أى تغيير فى هيكل البيانات أو هيئته سوف يتطلب تغييرات ملائمة لبرامج التطبيق، وإذا عملت تغييرات رئيسية للبيانات قد يحتاج برنامج التطبيق إلى إعادة كتابته من جديد.

أما فى نظام قاعدة البيانات فإن نظام إدارته يقدم التفاعل المحتاج إليه بين التطبيق والبيانات. وعندما تعمل تغييرات لعرض البيانات، فإن مواصفة البيانات Metadata المحفوظة فى نظام إدارة قاعدة البيانات تتغير ولكن يستمر النظام فى تقديم بيانات برامج التطبيق بالطريقة التى كانت مستخدمة من قبل. ويتداول نظام إدارة قاعدة البيانات مهمة تحويل البيانات عندما تستدعى الضرورة لذلك. هذا الاستقلال بين البرامج والبيانات يطلق عليه استقلالية البيانات، حيث تكون مهمة فى كل وقت يحتاج إليها لعمل بعض التغيير فى هيكل البيانات، فالبرامج المستخدمة قبل التغيير سوف تستمر فى العمل. وحتى يمكن تقديم درجة عالية من استقلالية البيانات، يجب أن يشتمل نظام إدارة قاعدة البيانات على نظام إدارة مواصفات

البيانات Metadata Management System الذى تتكامل فيه كل الملفات فى نظام واحد يمكن خلال المواصفات المقننة يؤدى إلى الحد من الإسهاب والتكرار، ويجعل إدارة البيانات أكثر كفاءة. إضافة لذلك، يقدم نظام إدارة قاعدة البيانات رقابة مركزية على البيانات التشغيلية.

ومن مزايا استقلالية البيانات وتكاملها والرقابة المركزية عليها ما يلى:

- تقليل الإسهاب وعدم التطابق.
- خدمة أحسن للمستخدمين.
- تحسن مرونة النظام.
- تكلفة تطوير وصيانة النظام أقل.
- إمكانية تطبيق المعايير الموحدة.
- إمكانية تحسين وسائل أمن النظام.
- تحسين سلامة النظام.
- إمكانية تعريف متطلبات المنظمة من البيانات.
- ضرورة تطوير نموذج وهيئة بيانات النظام.

(2) العيوب:

فى العادة يقدم نظام إدارة قاعدة البيانات وصولا على الخط لقاعدة بياناته لمستخدمين كثيرين. وبالمقابل، فإن النظام التقليدى قد يكون مصمما فى الغالب لتلبية حاجة معينة وبالتالى يقدم فى العادة وصولا لبياناته لعدد قليل ومحدود من المستخدمين. ولأن لعدد أكبر من المستخدمين وصولا إلى البيانات عند استخدام قاعدة البيانات، فإن المنظمة كالمدرسة قد تتضمن مخاطر إضافية عند المقارنة بنظام معالجة البيانات التقليدى فى المجالات التالية:

1. **السرية والسلامة والأمن**: عندما تكون المعلومات مركزية وتتوافر للمستخدمين من مواقع بعيدة؛ فإن احتمالات سوء الاستخدام تكون فى الغالب أكثر من

النظام التقليدى. ولتقليل وصول المستخدمين غير المعتمدين للمعلومات الحساسة يصبح من الضرورى اتخاذ إجراءات وقياسات فنية وإدارية وقانونية. وتخزن معظم قواعد البيانات معلومات قيمة يجب حمايتها ضد أى انتهاكات أو تجاوزات أو تخريب متعمد.

2. **جودة البيانات**: حيث يمكن وصول المستخدمين لقاعدة البيانات عن بُعد، لذلك يحتاج إلى أساليب رقابة ملائمة عند تحديث البيانات والتحكم فى جودة البيانات. وبسبب العدد المتزايد للمستخدمين فى الوصول للبيانات مباشرة، قد توجد أمامهم فرص كبيرة لإحداث أضرار بالبيانات. وإن لم تتوافر أساليب رقابة ملائمة ومناسبة فقد يتغاضى عن جودة البيانات.

3. **سلامة البيانات**: حيث إن أعدادا كبيرة من المستخدمين تستخدم حاليا قاعدة البيانات فى نفس الوقت، فإنه يحتاج إلى أساليب وقاية فنية ضرورية لتأكيد أن البيانات تبقى صحيحة خلال وقت التشغيل. ويكمن التهديد الأساسى فى أن المستخدمين المختلفين والمتعددين يحاولون كل على حدة فى تحديث نفس البيانات فى نفس الوقت. وعلى ذلك تحتاج قاعدة البيانات للحماية المستمرة ضد أى تغييرات سواء كانت مقصودة أو غير مقصودة تنبع من المستخدمين.

4. **قابلية تعطل المنشأة المعينة**: مركزية كل بيانات المنشأة فى قاعدة البيانات قد تعنى أن هذه القاعدة تصبح موردًا أساسيًا لا مفر منه. وقد يعتمد بقاء المنشأة أو المنظمة على معلومات موثوق منها وتتوافر من خلال قاعدة بياناتها. وعلى ذلك فإن المنظمة قد تصبح معرضة للخطر والتعطيل عند حدوث أى عطل طبقا لقاعدة البيانات أو لتعديلها غير المعتمد.

5. **تكلفة استخدام نظام إدارة قاعدة البيانات**: تصمم نظم معالجة البيانات التشغيلية لتشغيل عدد من العمليات الميسرة جيدا والمخططة سلفا. وفى الغالب،

تقدم مثل هذه النظم للتشغيل بكفاءة عالية للعمليات التى صممت من أجلها. وعلى الرغم من أن النظم التقليدية تكون عادة مرنة، إلا أن التطبيقات الجديدة تكون صعبة التنفيذ وذات تكلفة تشغيلية عالية جدا فى كثير من الحالات، ومع ذلك فإنها فى الغالب تتسم بالكفاءة للتطبيقات التى صممت من أجلها. من جهة أخرى، يقدم مدخل قاعدة البيانات بديلا غير مكلف نسبيا. ولا يكون المدخل المرن بدون تكاليف وأحد عناصرها يتمثل فى التكلفة الإضافية المتصلة بتشغيل التطبيقات التى صممت من أجل النظام التقليدى. وأن استخدام حزمة برمجيات معيارية يكون دائما أقل كفاءة من البرمجيات المتخصصة.

3 - المفاهيم الأساسية لنظام إدارة قاعدة البيانات:

1/3 البيانات:

البيانات هى حقائق والبعض منها مهم أكثر لمستخدم من البعض الآخر. وبعض الحقائق هى مهمة بالكفاية التى تضمن تتبعها فى طريقة رسمية. والبيانات المهمة تشبه الأشياء النفيسة التى نحتفظ بها وتمثل مجموعة صغيرة لممتلكات الشخص ولكنها مهمة جدا يحتم حمايتها بوضعها فى مكان خاص آمن. والبيانات تمثل مفهوم عرض يشتمل على الصور (أشكال ثنائية)، البرامج والقواعد. فالبيانات تشتمل على البيانات الرقمية التى قد تكون أعدادا صحيحة (كل الأعداد فقط) أو أعدادا عشرية. والبيانات غير الرقمية تمثل الحروف، البيانات المنطقية (صحيحة أو خطأ).وقد تتضمن قواعد البيانات المتقدمة على كيانات بيانات أكثر كالأشكال والنصوص. وبصفة غير رسمية تمثل البيانات الأشكال المراد تخزينها فى قاعدة البيانات.

ويمكن تقسيم البيانات إلى بيانات مكانية Spatial وبيانات غير مكانية. وتشتمل البيانات المكانية على الموقع، الشكل، الحجم والتوجه Orientation . على سبيل

المثال، يمكن اعتبار شكل المربع بأن له مركزًا تتداخل فيه أقطاره ويحدد موقعه، والشكل يمثل المربع، كما أن طول أحد جوانبه يحدد حجمه، وزاوية أقطاره تجعل محوره x-axis يحدد توجهه.

أما البيانات غير المكانية فهى التى يطلق عليها الخواص أو الخصائص تمثل المعلومات التى يجب أن تكون مستقلة فى كل اعتباراتها الهندسية Geometrics . على سبيل المثال، طول الشخص، العمر، الجنس... الخ. كما أنه من المهم ملاحظة أنه بينما تكون بيانات الكتلة غير مكانية فإن الثقل يكون بيانات مكانية؛ حيث إن ثقل الشيء يعتمد على الموقع.

ومن الممكن تجنب أو التغاضى عن التمييز بين البيانات المكانية وغير المكانية. على أى حال، توجد اختلافات جوهرية بين كل من البيانات المكانية وتلك غير المكانية. فالبيانات المكانية هى فى العادة بيانات متعددة الأبعاد التى تترابط معا، أما البيانات غير المكانية فهى عادة بيانات ذات بعد واحد ومستقلة. وتضع هذه الميزة البيانات المكانية وغير الرقمية فى رؤى مختلفة مع تضمينات واسعة الانتشار للأوجه الفكرية والمعالجة والتخزين. على سبيل المثال، من المحتمل أن يكون الفرز وظيفة معالجة البيانات غير المكانية الأكثر أهمية المؤداة.

2/3 قاعدة البيانات:

تشبه قاعدة البيانات محتويات كتاب ما أو المعلومات المخزنة فى صفحاته، فهى مجموعة الحقائق التى يشتمل عليها دليل التليفونات على سبيل المثال أيضا.

3/3 المستودع:

المستودع هو الهيكل الذى يخزن البيانات ويحميها. وتقدم المستودعات الوظيفية التالية: إضافة (إدخال بيانات)، تخزين بيانات، استرجاع (إيجاد واختيار) البيانات من المستودع، استبعاد (حذف) بيانات من المستودع.

وتسمح بعض المستودعات بتغيير البيانات أو تحديثها. وينجز تحديث البيانات بواسطة استرجاع نسخة بيانات من المستودع وتحديثها من خلال استبعاد البيانات القديمة وإدخال البيانات المراد تحديثها. وتتواجد المستودعات أساسا لحماية المحتويات التى قد تخترق عرضيا من خلال التالى:

- الأمن Security تعتبر المستودعات نموذجا لكلمة المرور Password. ويوجد لكثير من النظم لآليات أمن مفصلة جيدا.
- البيانات العرضية كإجراء وقائى Safeguard عبر آلية التصرف الذى يمثل تتابع إجراءات تداول قاعدة البيانات ولها خاصية تتمثل فى أنه عند توقفها قبل الاكتمال سوف يعاد تخزين قاعدة البيانات لحالة متوافقة ذاتيا. وفى العادة قبل بدء التصرف فإنه عند اكتماله سوف تكون قاعدة البيانات فى حالة متوافقة ذاتيا.

4/3 نظام إدارة قاعدة البيانات:

كما سبق ذكره، فإن نظام إدارة قاعدة البيانات هو مستودع بيانات مع واجهة تفاعل المستخدم المقدمة لتداول وإدارة قاعدة البيانات. ويفهم نظام إدارة قاعدة البيانات بأنه نظام برمجيات التى تشغل قاعدة البيانات كمبيوتريا. ومن أمثلة نظم إدارة قاعدة البيانات المتوافرة تجاريا النظم التالية: MS. Acces, MS. SQL Server, Oracle, Sybase, DB2, etc. ويقدم نظام إدارة قاعدة البيانات كثيرا من الأوجه التى يفتقد إليها فى مستودع بيانات عادى كدليل التليفونات، على سبيل المثال.

وتقدم معظم نظم إدارة قواعد البيانات عمليات معمارية مقننة التى توفر قدرات أساسية لتداول البيانات. ومن أمثلة هذه العمليات: إدخال بيانات، تخزين سجلات، استبعاد حقول أو سجلات، استرجاع تقارير أو سجلات، تحديث أو تعديل بيانات... الخ.

وتشتمل نظم إدارة قاعدة البيانات على التالى:

- **لغة وصف البيانات: DDL** هى لغة مستخدمة لوصف محتوى قاعدة البيانات. وتستخدم لوصف مسميات خواص البيانات (مسميات الحقول) وأنواع بيانات الموقع فى قاعدة البيانات... الخ.
- **لغة تداول البيانات والتساؤل:** تساند لغة التساؤل بواسطة نظام إدارة قاعدة البيانات لتشكيل الأوامر الخاصة بالإدخال، التحرير، الإخراج، التهيئة... الخ. وفى هذا النطاق تتحقق درجة معينة من التقنين مع لغة التساؤل الهيكلية Structured Query Language (SQL).
- **أدوات البرمجة:** بجانب الأوامر والتساؤلات، يجب أن يكون لقاعدة البيانات وصولا مباشرا من برامج التطبيق خلال طلبات الوظيفة التى يطلق عليها طلبات الروتين الفرعى فى لغات البرمجة التقليدية.
- **هياكل الملفات:** File Structures لكل نظام إدارة قاعدة البيانات هياكله الداخلية المستخدمة لتنظيم البيانات على الرغم من أن بعض نماذج البيانات المشتركة تستخدم بواسطة معظم نظم إدارة قاعدة البيانات.

5/3 التساؤلات: Queries

تستخدم كثير من نظم إدارة قاعدة البيانات واجهة تفاعل المستخدم التى تشتمل على فرز بعض اللغة الرسمية؛ حيث تستخدم لغة وصف البيانات لتحديد أى البيانات سوف تخزن فى قاعدة البيانات وكيفية ترابطها معا، أما لغة تداول البيانات فتستخدم لإضافة البيانات، استرجاعها، تحديثها، استبعادها... الخ، فى نظام إدارة قاعدة البيانات. ويأخذ التساؤل غالبا عبارة أو مجموعة عبارات فى لغة وصف البيانات أو لغة تداول البيانات، ويرى البعض أن التساؤلات ينظر إليها كعمليات قراءة فقط لا يسمح فيها بتعديلات البيانات. وتعتبر لغة التساؤل Query Language لغة رسمية تنفذ لغة وصف البيانات ولغة تداول البيانات أو كليهما معا، على سبيل المثال، من لغات التساؤل " لغة التساؤل الهيكلية SQL والتساؤل بواسطة المثال Query-by-example" ... الخ.

6/3 نماذج البيانات: Data Models

نموذج البيانات يمثل معادلة رياضية أو صيغة محددة تشتمل على جزأين: فكرة لوصف البيانات، ومجموعة عمليات مستخدمة لتداول البيانات. وبذلك يعتبر نموذج البيانات كطريقة تنظيم مجموعات حقائق ترتبط بنظام تحت الدراسة، كما تقدم طريقة تفكير عن العالم المحيط كطريقة تنظيم ظواهر الاهتمامات المعينة، وبذلك يمكن التفكير فيها كلغة مجردة تمثل مجموعة كلمات مع قواعد النحو الخاصة بها التى من خلالها لا يمكن وصف الموضوع.

والفائدة الرئيسية من إتباع نموذج بيانات تتضح من أساس النموذج النظري؛ حيث إنه من النظرية تظهر قوة التحليل والقدرة على استخلاص الاستدلالات وخلق القياسات التى تبزغ من البيانات الخام.

وتقدم النماذج المختلفة مفاهيم مختلفة عن الواقع، كما أن لها رؤى وأبعادا مختلفة. وبذلك لا يوجد اتفاق كامل عن نموذج البيانات الأحسن. وتشتمل نظم إدارة قاعدة البيانات على ثلاث مستويات من التجريد تتمثل رفى التالى:

- **المستوى الطبيعى**: Physical Level الذى يمثل تنفيذ قاعدة البيانات كمبيوتريا. وتختص بالأشياء كهياكل التخزين، وطريقة الوصول إلى هذه الهياكل.
- **المستوى الفكرى**: Conceptual Level يرتبط بتعبير نموذج العالم الواقعى لمصمم قاعدة البيانات.
- **مستوى المنظر**: View Level الذى فى إمكانه إعطاء وصولا لأجزاء مختلفة من قاعدة البيانات لمجموعات مختلفة من المستخدمين. ويطلق على جزء قاعدة البيانات لمجموعة معينة من المستخدمين "منظر View".

4 -نماذج البيانات الأكثر شيوعا:

يتعرض هذا الجزء من هذا الفصل لأطر أكثر نماذج بيانات نظم إدارة قواعد البيانات الشائعة بالفعل:

1/4 نموذج علاقة الكيانات: Entity-Relationship Model (ERM)

فى إطار نموذج علاقة الكيانات يمكن تصور أن العالم أو أى نظام مكون من كيانات Entities التى ترتبط مع بعضها ببعض بواسطة علاقات Relationships. وتجمع الكيانات فى أنواع معينة يطلق عليها "مجموعات كيانات Entity Sets". ويمكن اكتشاف مجموعات الكيانات والعلاقات من خلال رسمها فى خريطة علاقة الكيانات Entity-Relationship Diagram (ERD).

1. الكيانات:

الكيانات هى أشياء Objects تتواجد فى عالم المنظمة الحقيقى التى يمكن تمييزها مثل الطلاب، المدرسين، المواد الدراسية، جداول الدراسة... الخ. وتعنى إمكانية التمييز أن كل مجموعة من مجموعات الكيانات يمكن تعريفها بطريقة فريدة. وكل مجموعة كيانات تشتمل على خواص مشتركة تفسر وتوصف ما يدل عليه الكيان المعين. ولا يتضمن أى شيء حقيقى بالضرورة غرض الكيان المفرد أو الأحسن؛ حيث إنه لكل شيء معين يمكن اختيار مجموعات خواص أو خصائص مختلفة للشيء التى تهتم بوضعها المعين. هذه النتائج لنفس الشيء تكون منمذجة بطرق مختلفة. كما تجمع الكيانات فى مجموعات يمكن رسمها معا فى خريطة علاقة الكيانات ERD.

2. العلاقات:

العلاقة تمثل قائمة مجموعات كيانات قد يعبر عنها بترقيم أ ، ب ، الخ. والعلاقة يعبر عنها بترقيم ع، مما يشكل معادلة كالتالى: أ ع ب.وتوجد أنواع مختلفة من العلاقات يمكن حصرها فى الأنواع الثلاث التالية:

- علاقة واحد لواحد: One-One إذا كانت مجموعة علاقات تمثل أ ع ب وأن ع هى علاقة واحد لواحد كعلاقة المدرس بفصل دراسى مفرد أو مادة تعليمية معينة عندئذ فإن كل كيان فى ب له علاقة مع كيان أ والعكس صحيح.

- علاقة كثير لواحد: Many-One إذا كانت صيغة علاقة أ ع ب وأن العلاقة ع هى علاقة كثير لواحد، عندئذ كل كيانات مجموعة كيان أ كالمعلمين لها علاقة مع كيان واحد فقط ب كالمدرسة ولكن ليس العكس.على سبيل المثال كل المدرسين أو الطلاب الذين يمثلون كيان أ لهم علاقة مع كيان ب الذى قد يكون المدرسة.
- علاقة كثير لكثير: Many-Many إذا كانت صيغة العلاقة أ ع ب ، وأن العلاقة ع تمثل علاقة كثير لكثير، عندئذ فإن كل كيان من مجموعة أ يمكن أن يكون له علاقة مع عدد من كيانات ب والعكس صحيح. مثال ذلك أن كيانات المدرسين أو الطلاب أ لهم علاقة بكيانات الفصول الدراسية أو قاعات الدراسة أو المقررات الدراسية المختلفة ب... الخ .

وتكتشف العلاقة بواسطة رسم خطوط تربطها بمجموعة كيانات متوافقة فى إطار خريطة علاقة الكيانات ERD التى ترتبط بنموذج علاقة الكيانات ERM الذى يستخدم لتحديد المفاهيم العامة فى إنشاء نماذج طبيعية.

2/4 نموذج الشبكة: Network Model

يبنى نموذج الشبكة على مفهوم هيكل، كما هو متواجد فى لغات برمجة مثل لغة السى C أو لغة البسكال Pascal . ويمكن نمذجة الكيانات كهياكل بخواص الكيان التى تتصل بحقول الهيكل. كما تتميز الكيانات بواسطة الموقع الذى يمثل عنوان الهيكل الطبيعى الذى يحتفظ فيه. وعلى ذلك؛ فإنه يمكن عرض هيكلين من قيمة واحدة لكيانين منفصلين. كما يمكن تنفيذ مجوعات الكيانات كملفات تضاهى سجلاتها الهيكل. وتنشأ العلاقات مع وصلات واضحة (فى مواجهة المؤشرات Pointers) من هيكل لآخر.

ولا يوجد لنموذج الشبكة دلالات رسمية ولغة استفسار عالية المستوى. وتتداول قاعدة البيانات عملياتها من خلال برامج تكتب فى الغالب بلغة البرمجة مثل لغة الكوبول COBOL التى كانت مستخدمة فى الماضى. وقواعد بيانات

الشبكة مشفرة Coded يدويا. وعلى ذلك، يمكن أن تتسم بالكفاءة عند تنفيذها مع مرات تنفيذ التساؤل المعين. وفى هذا الصدد، فإن كل العلاقات تبنى فى قاعدة البيانات الكمبيوترية. وقد يعيب هذا الأداء عدم مرونته المتسم بالجمود وصعوبة الاستخدام من قبل المستخدمين.

3/4 النموذج العلاقى: Relational Model

طور النموذج العلاقى أى المبنى على العلاقات فى بداية عقد السبعينيات من القرن الماضى. ومنذ ذلك الوقت صار النموذج العلاقى منتشرا فى نظم إدارة قاعدة البيانات المبنية على مفهوم العلاقة التى ترتبط بمجموعة عناصر أو حقائق يطلق عليها Set of Tuples ترتبط معا بطريقة ما (وقد يكون ذلك من خلال الحقائق أو العناصر التى توضع معا فى مجموعة معينة). والنموذج العلاقى هو فى الواقع نموذج بيانات كامل رياضى يبنى على الدعامة النظرية.

1. مجموعة الحقائق / العناصر:

وأقل حقيقة فى مجموعة الحقائق تعتبر بيانه تستمد قيمتها من مجال محدد كمجال الأعداد الصحيحة Integers، كما تمثل مجموعة قيم التى يطلق عليها خواص.

2. العلاقات:

تمثل العلاقة مجموعة فرعية لكل مجموعة الحقائق التى تشكل بواسطة مجال محدد كما فى حالة مجال الطلاب. وحيث إن مجموعات العناصر أو الحقائق هى مجموعات قيم وعلاقة، فمثلا يمكن ملاحظة التالى: الملف هو قائمة بكل السجلات فيه، والجدول يمثل قائمة صفوف أعمدة، والعلاقة هى مجموعة عناصر تتواجد فى أعمدة تمثل خواص الكيان.

وتتمثل العلاقات بصفة طبيعية فى الجداول من حيث:

- الجداول فى حد ذاتها ليست علاقات، لأن العلاقات لا يمكن أن تكرر مجموعة

الحقائق أو العناصر. ولا يوجد مثل هذا التركيب على الجداول. ومن المحتمل أن يكون ذلك مريحا للتفكير عن العلاقات كجداول كلما يبقى التميزّ واضحًا. معظم نظم إدارة قاعدة البيانات العلاقية التجارية المتاحة إن لم يكن كلها تقريبا تنتهك هذا لمبدأ لأنها تكرر مجموعات العناصر أو الحقائق.

- استخدام العلاقات كأداة نمذجة البيانات تعتبر واضحة عند توافر هذه العلاقات.

- كل مجموعات عناصر أو حقائق فى علاقة يمكن أن تتميز بواسطة قيم خواصها، أى مجموعة خواص لها قيم بالضرورة تعرف بمجموعة الخواص، ويطلق عليها مفتاح Key الذى يعرف مجموعة عناصر أو حقائق بطريقة فريدة. ويستخدم مفتاح مجموعة عناصر فى الغالب كبديل لمجموعة عناصر كاملة أخرى، أى أن العلاقات تتراكم معا وتعرض بواسطة مفتاح علاقة واحدة مع العلاقات المتواجدة.

ويختار مصممو قاعدة البيانات بعض مجموعة خواص معينة لكى تكون مفتاحًا لعلاقات قاعدة البيانات ويطلق على ذلك "مفتاح أصلى أو أولى Primary Key". وإذا كان هذا المفتاح الأصلى لأحد الجداول يظهر كخاصية علاقة مختلفة فيطلق عليه "مفتاح أجنبى Foreign Key" فى علاقة أخرى. على سبيل المثال علاقة واحد لواحد يكون المفتاح الأصلى لكيان أ متضمنا أيضا فى كيان ب والعكس صحيح أو لكليهما. وأيضا، قد يفترض أن اسم المدرس هو المفتاح الأصلى لسجل الطالب أو الفصل الدراسى، عندئذ يمكن أن تكون للمدرس خاصية وقيمتها هى اسم المدرس فى سجل الفصل أو سجل الطالب. كما أنه فى علاقة كثير لواحد فإنه كما فى حالة علاقة المدرس الواحد بكثير من الطلاب أو الفصول أو المواد الدراسية يتواجد مفتاح المدرس الأصلى فى سجلات الطلاب أو الفصول أو المواد الدراسية كمفتاح أجنبى. أما فى علاقة كثير لكثير بوجود علاقات متشابكة بين المدرسين

والطلاب والفصول والمواد الدراسية، فإن المفتاح الأصلى لكل كيان يظهر كمفتاح أجنبى فى المجالات الأخرى.

3. لغات التساؤل:

منذ عقد السبعينيات فى القرن الماضى وجدت لغتين أوليتين للتعامل مع العلاقات: أحدهما هى اللغة الجبرية Algebraic وبنيت اللغة الأخرى على منطق التنبؤ للأمر الأول ولهاتين اللغتين نفس التعبير القوى:

- **الجبر العلاقى**: حيث يوجد الترقيم الجبرى عندما يعبر عن التساؤلات بتطبيق مشغلات متخصصة للعلاقات.
- **التفاضل والتكامل العلاقى**: Relational Calculus يتواجد ترقيم منطقي، حيث يعبر عن التساؤلات بكتابة صيغ منطقية يجب تلبيتها فى إجابة مجموعات العناصر أو الحقائق.
- **لغة التساؤل الهيكلية**: SQL هى إحدى لغات التساؤل التجارية المنتشرة على نطاق واسع.

4. نظام إدارة قاعدة البيانات العلاقية: RDBMS

يمثل هذا النموذج العلاقى النوع الأكثر شيوعا وانتشارا. وتعتبر قاعدة البيانات العلاقية المرتبطة بهذا النظام مجموعة وحدات البيانات المنظمة فى مجموعة جداول مفسرة رسميا التى من خلالها يمكن الوصول أو إعادة تجميع البيانات بطرق عديدة مختلفة بدون الحاجة لإعادة تنظيم جداول قاعدة البيانات. وتبنى قاعدة البيانات العلاقية على نموذج علاقى مصمم باستخدام شفرات أو أكواد طورتها شركة آى بى إم IBM فى بداية السبعينيات من القرن الماضى.

ونظام إدارة قاعدة البيانات العلاقى هو حزمة برمجيات تسمح بإنشاء وتحديث وإدارة قاعدة البيانات العلاقية. ويتضمن استخدام عبارات يدخلها المستخدم أو

تحفظ فى برنامج التطبيق. وتوجد إصدارات أو تحديثات عديدة لهذه النظم التى من ضمنها: MS. Access, Oracle, Sybase, Informix ... الخ.

وتعتبر هذه البرمجيات سهلة الاستخدام نسبيا فى إنشاء والوصول والامتداد. وبعد إنشاء قاعدة البيانات الأصلية، يمكن إضافة مجموعة البيانات الجديدة بدون الحاجة لطلب تعديل كل التطبيقات المتواجدة بها.

وقاعدة البيانات العلاقية هى مجموعة جداول (يطلق عليها أحيانا علاقات) تتضمن مجموعة بيانات أو أكثر فى أعمدة الجدول. ويشتمل كل صف من صفوف الجدول على حالة بيانات فريدة للمجموعات المفسرة بواسطة الأعمدة.

وتتسم نظم إدارة قاعدة البيانات العلاقية بالمزايا التالية:

- إمكانية الوصول: Accessibility الوصول العشوائى لاستخدام لغات التساؤل العشوائية (مثل لغة التساؤل الهيكلية SQL) توضح أن التساؤلات يمكن تجميعها وتنفيذها وتعظيمها آليا بدون إعادة الفرز للبرمجة.
- الصحة: Correctness تبين دلالات الجبر العلاقى المنطقية التى تكون كاملة وموثوق منها.
- المرونة: Flexibility يفصل النموذج العلاقى النموذج المنطقى من النموذج الطبيعى، ويعنى ذلك فك عملية التزاوج المتواجدة.
- السلامة: Integrity تأكيد سلامة وأمن قاعدة البيانات وأن القيود التى تؤكد التغييرات الهيكلية لا تؤثر على معنى قاعدة البيانات بطريقة معاكسة.
- إمكانية التنبؤ: Predictability تساعد الدلالات والمعانى المتوافقة والمترابطة مجموعات المستخدمين فى توقع نتيجة تساؤلاتهم بسهولة.
- الآراء المتعددة: Multiple Views تساعد قاعدة البيانات العلاقية فى عرض آراء مجموعات المستخدمين المختلفة عن نفس قاعدة البيانات.

- الرقابة التزامنية: Concurrency Control تسهم فى الرقابة المتزامنة على البيانات المتاحة لمجموعات المستخدمين المختلفة.
- ...الخ.

4/4 النموذج الشيئى: Object Model

1. ما هو النموذج الشيئى:

تشبه كلمة "شيء" مفهوم "الكيان" فى علاقة الكيانات على الرغم من أن الشيء يعتبر أكثر عمومية مع التوصية باعتبار الشيء فيما يتعلق بالعالم الطبيعى المرتبط بالمنظمات التعليمية كالمدارس على سبيل المثال. أى أن الأشياء لا تقتصر فقط على الأشياء الطبيعية أو المحسوسة على سبيل المثال، يمكن أن تكون هياكل البيانات أشياء، وكما فى حالة النماذج الأخرى، يفترض لنموذج الشيء جمع الأشياء معا فى مجموعات مفهومة، ويطلق على هذه المجموعات "فصول أو أقسام Classes".

وتجميع الأشياء معا فى فصول تكون مفهومة،لأن أشياء نفس الفصل يجب أن تشترك معا فى مجموعة خواص فريدة بهذا الفصل، كالسلوك والعلاقات المشتركة مع الأشياء. واختلافا عن مجموعة الكيانات والعلاقات، لا تحفظ الفصول فعليا أشياء الفصل المعين، حيث تكون الفصول ذات طبيعة فكرية بصورة مجردة. ولا يوجد أى شيء فى النموذج الشيئى الذى يتساوى مع مجوعة الكيانات أو العلاقات. فمثلا فى نموذج الشبكة تكون العلاقات بين الأشياء هى التى تحدد عبر وصلة أو مؤشر طبيعى بين الأشياء. ويمكن وصف النموذج الشيئى كهيكل الأشياء فى النظام من حيث كياناتها وعلاقاتها للأشياء الأخرى وخواصها وعملياتها.

2. قاعدة البيانات الموجهة شيئيا: Object-Oriented Database (OODB)

قواعد البيانات الموجهة شيئيا هى قواعد بيانات تساند الأشياء والفصول

Classes. وتختلف هذه القواعد عن قواعد البيانات العلاقية التقليدية، لأنها تسمح بهيكلة الأشياء الفرعية وكل شيء له تعريف خاص أى Object-ID فى مواجهة القيمة الموجهة للشيء.

لذلك بزغت قاعدة البيانات العلاقية الشيئية Object-Relational Database (ORDB) التى تجمع خصائص كل من قاعدة البيانات الموجهة شيئيا وقاعدة البيانات العلاقية بناء على الأشياء. ومن مزايا قاعدة البيانات العلاقية الشيئية إمكانيتها فى إدارة بيانات معقدة تعسفية، إلا أن من عيوبها أنها أضعف فى مساندتها للتساؤل العشوائى وفى إنتاج الأوجه المشتركة مع قواعد البيانات الهرمية. ومن أمثلة نظم إدارة قاعدة البيانات العلاقية الشيئية ORDBMS خادم Informix Universal Server الذى يعمل على نظام تشغيل يونكس UNIX ونظام النوافذ والشبكات.

والأوجه الأساسية لنموذج البيانات الموجه شيئيا تتمثل فى التالى:

- هوية الشيء Object Identity يمثل قدرة النظام فى التمييز بين شيئين مختلفين لهما نفس الحالة، وحالة الشيء يمكن المشاركة فيها من خلال أشياء عديدة لكل منها هوية مستقلة.
- التعبئة أو التغليف Encapsulation تحدد نوع من أنواع التجريد الذى يقوى الفصل النظيف بين التفاعل الخارجى (السلوك) لشيء ما وتنفيذه داخليا. وتتطلب التعبئة أن يكون كل الوصول (التفاعل) مع الأشياء يعمل بواسطة تنشيط تنفيذ الخدمة المقدمة عن طريق تفاعلها الخارجى.
- الحالة المعقدة Complex State يكون لها القدرة على تفسير أنواع البيانات التى يشكل تنفيذها هيكل شبكى Nested. وحالة الشيء يمكن أن تبنى من سجلات الأنواع الأولية Primitive ، الأشياء الأخرى أو مجموعات الأشياء.
- نوع قابلية الامتداد أو التوسع Type Extensibility يوضح القدرة لتفسير أنواع

بيانات جديدة مفسرة مسبقا بواسطة تعزيز أو تغيير هيكل أو سلوك الأنواع. ويعتبر نوع الوراثة Inheritance آلية تستخدم لتفسير أنواع جديدة بواسطة تعزيز سلوك متواجدة بالفعل.

- العضوية Genericity تمثل نوع نموذج البيانات الشيئى التى تتعاون فيها لغة تساؤل الشيء التى تتسم بالعضوية فى النظام. ويعتبر ذلك كنوع جديد مضاف للنظام يمكن من عملية التساؤل العضوى.
- لا يوجد نموذج بيانات شيئى متفق عليه دوليا من كل الأشخاص، ولكن بيان نظام قاعدة بيانات التوجه الشيئى يعطى إطارا ينبثق من المعيار المراد تطبيقه.
- تمثل حقبة قاعدة بيانات التوجه الشيئى تجميع نظم لغة برمجة التوجه الشيئى التى تأتى قوتها من معالجة البيانات الشبيهة لما توجد فى قواعد البيانات والبيانات العابرة أى الثابتة Transient، كما توجد فى برامج التنفيذ. ويلاحظ أن التركيز مع قاعدة بيانات التوجه الشيئى تشبه نموذج الشبكة متجه نحو المبرمجين لا مستخدمى النظم النهائيين.

3. ميراث العلاقات: Inheritance Relationships and Typing

نماذج التوجه الشيئى الكثيرة قد تستخدم لغة البرمجة C++ للميراث المتعدد المكرر، ولغة جافا Java للميراث المفرد، ولغة Actor فى حالة عدم وجود ميراث. ويكون الشيء هو مثل أو حالة الفصل الذى تتواجد به، وينظر للميراث من منظورين:

- منظور تعاقبى Incremental ويمثل عملية إضافة الخواص والوظائف لفصل موجود بالفعل (فصل أساس Base Class). ويمكن إضافة خواص أو وظائف جديدة لفصل جديد لم يكن فى فصل الأساس من قبل، ويمثل ذلك أسلوب لإعادة استخدام الشفرة.
- طباعة فرعية Subtyping وهو أسلوب لترتيب تعاريف الفصل فى هرمية ترضى

الشرط أو الوضع فى أن أعضاء الفصل الفرعى هم أيضا أعضاء فى الفصل الأعلى Superclass. وفى هذه الحالة، فإن الخواص أو الوظائف يمكن أن تغير نوعًا كلما كان النوع الجديد (موروث إما مباشرة أو غير مباشرة) من فصل الأساس الأصلى. ولا يمكن استبعاد الخواص أو الوظائف القديمة التى يمكن أن تقدم مع تنفيذات جديدة كلما بقى التفاعل للوظيفة غير متغير (أو متغير عبر التخصص).

4. التعبئة أو التغليف: Encapsulation

يتضمن تعبئة أو تغليف خواص وسلوك الأشياء التالى:

- لا توجد تفاعلية مع الشيء الذى لا يذهب خلال التفاعل المنشور العام.
- يتضمن حالة الأشياء تفسير فصل سلوك الشيء الذى يوضح كوظائف وإجراءات.
- حالة الشيء لا يمكن أن تتداول بواسطة أى شيء خارجيا لها.

على سبيل المثال، فى اللغة الموجهة نحو اللاشيء مثل لغة برمجة C قد يقال أن المبرمج يكتب إجراء لتغيير قيم الهيكل التى تحتفظ بوضع الرسومات الأولية التى تشتمل على المعلومات الشرطية مع الإجراء الداخلى الذى يغير وضعه، ويرسل المبرمج رسالة إلى الشيء طالبا تغيير وضعه.

وميزة التعبئة ترتبط بتنفيذ سلوك يمكن أن يتغير بدون تأثير أى فصل آخر فى النظام، ويساعد ذلك فى تفكيك تزاوج الفصول كما يقلل من تعقيد النظام.

5. مقارنة النموذج الشيئى مع النموذج العلاقى:

النموذج الشيئى يختلف عن النموذج العلاقى فى الطرق التالية:

- يسمح النموذج الشيئى بأشياء معقدة لكى تحدد خصائص المجالات ولا يسمح النموذج العلاقى بذلك.

- النوع المتوافر المعقد فقط فى النموذج العلاقى هو العلاقة، ويقتصر نموذج الشيء على أن كل كيانات النظام تكون أشياءًا أكثر عمومية عن العلاقة (العلاقة قد تكون أشياءًا ولكن الأشياء لا تكون علاقات).
- يسمح النموذج الشيئى بعدم وجود مجموعات خصائص أو حقائق وبالتبعية تعرف الكيانات بواسطة قيم الخاصية. ويفترض النموذج الشيئى وجود هوية أو تعريف ID الشيء بطريقة فريدة غير مرئية للمستخدم.
- الأشياء هى حالات الفصول وتشكل الفصول نظام طبع النموذج، حيث لا يوجد مفهوم مستوى فصل الطبع فى النموذج العلاقى وكل شيء يكون علاقة. ويساند النموذج العلاقى مجالات المستخدم المفسرة جيدا ولكنه يطبق فى مستوى الخاصية؛ حيث إنه مع نموذج الشيء يكون الفصل نوعا، وما لا يتساوى فى العالم العلاقى يكون للعلاقات تشكيل الأنواع أيضا.
- تحفظ العلاقات كل مجموعات العناصر أو الحقائق، ولا توجد مساواة بين الأشياء، كما لا توجد مجموعة أو أى شيء آخر يشتمل على كل أشياء الفصل.
- توجد طلبات أعلى للغات تساؤل غير مبرمجة للنموذج العلاقى، بينما يوجد قليل من الفرص المساوية للنموذج الشيئى.
- يوجه النموذج الشيئى أكثر للمبرمجين لا للمستخدمين النهائيين، والعكس يكون صحيحا للنموذج العلاقى.

5 - تصميم قاعدة البيانات العلاقية :

1/5 منظور نموذج قاعدة البيانات العلاقية :

يشتمل منظور نموذج قاعدة البيانات العلاقية على ثلاث مكونات أساسية، هى:

1. مكون هيكلى يتضمن مجموعة جداول يطلق عليها أيضًا علاقات.

2. مكون متداول يدويا يشتمل على مجموعة عمليات عالية المستوى.
3. مجوعة قواعد لصيانة سلامة قاعدة البيانات.

وكما سبق تحديده فإن المصطلحات المرتبطة بنظرية قاعدة البيانات العلاقية من فرع الرياضيات الذى يطلق عليه نظرية المجموعة Set Theory على الرغم من تواجد مترادفات مستخدمة بتوسع لهذه الألفاظ الرياضية الدقيقة.

ويخزن كل شيء فى قاعدة البيانات العلاقية فى الجداول التى تشتمل على أعمدة وصفوف. ويستمد العامود من مجال محدد الذى يمثل مجموعة القيم التى منها تستمد القيم الفعلية. ويوجد أكثر من عامود فى الجدول تبين قيمة من نفس المجال. ومدخل العامود فى أى صنف يمثل مفردة ذات قيمة SINGLE-VALUED، أى يشتمل على وحدة فقط كالعمر أو الاسم. وتكرار المجموعات، حيث إن الأعمدة التى تشتمل على مجموعات تكون بدلا من قيمة مفردة لا يسمح بها. وكل صف من صفوف الجدول يعتبر مفسر فريد بواسطة مفتاح أصلى Primary Key قد يشتمل على عامود أو أكثر الذى يتضمن فى الجدول الذى قد لا يشتمل على صفوف مكررة. وقاعدة سلامة كيان النموذج Entity Integrity Rule توضح أنه لا يوجد مكون مفتاح أصلى لا يشتمل على لا قيمة NULL. وقاعدة سلامة أفضلية النموذج Referential Integity Rule توضح أنه لكل قيمة مفتاح أجنبى Foreign Key فى الجدول الذى يجب أن يوجد مفتاح أصلى فى جدول آخر فى قاعدة البيانات.

2/5 عمليات قاعدة البيانات العلاقية:

يفسر النموذج العلاقى عددا من المشغلات العلاقية Relational Operators التى يمكن أن تستخدم لتداول البيانات، وتتمثل المشغلات العلاقية فى التالى:

1. الوحدة UNION
2. الفصول المتداخلة INTERSECTION

3. الاختلاف DIFFERENCE

4. التقسيم / التجزيء DIVIDE

5. المنتج PRODUCT

6. الاختيار SELECT

7. المشروع PROJECT

8. الوصل JOIN

وتتطلب مجموعة مشكلات مشغلات "الوحدة" و"الفصول المتداخلة" و"الاختلاف" أن للعلاقات نفس الدرجة (رقم الأعمدة) والخصائص المرتبطة يجب أن تكون لنفس المجال. على سبيل المثال، "وحدة" (أ، ب) تمثل علاقتين مع علاقة ثالثة منتجة تشتمل على كل الصفوف التى إما أن تكون فى العلاقتين الأولى أو الثانية أو العلاقتين معا. ويتساوى ذلك مع العلاقة المنطقية OR. كما أن إدخال مشغل "الفصول المتداخلة" أ فى جدول ب، أى إدخال علاقتين يمثل علاقة تشتمل على كل الصفوف التى تكون علاقات معا. وإدخال مشغل "الفصول المتداخلة" يساوى العلاقة المنطقية AND. وبذلك فإن الاختلاف بين (أ – ب) يمثل اختلاف علاقتين ، أحدهما تشتمل على كل الصفوف التى تكون فى العلاقة الأولى، إلا أنها ليست متواجدة فى العلاقة الثانية. والمشغلات العلاقية الأخرى هى:

- التقسيم / التجزيء (أ / ب) يتطلب أن كل هيكل علاقة واحدة تمثل جزءًا من أجزاء هيكل العلاقة وتحدد أى قيم تظهر فى مع كل الصفوف.
- المنتج (أ x ب) منتج العلاقتين يمثل علاقة تشتمل على كل تجمعات العلاقتين الممكنتين.
- الاختيار يسترجع كل الصفوف التى تلبى وضع معين أى مجموعة فرعية أفقية.

- المشروع يسترجع مجموعة عمودية أى رأسية من الصفوف، وتكون الصفوف المكررة محظورة أو لا تكون محظورة فى العلاقة الجديدة.
- الوصل هو تجمع من صفوف علاقتين بناءًا على خاصتين (خاصية فى كل علاقة) تتواجدين مع مجا مشترك.

3/5 معالم تصميم قاعدة البيانات العلاقية:

1. **المنظور العام**: تقدم برمجيات نظام إدارة قاعدة بيانات مايكروسوفت أكسس MS. ACCESS على سبيل المثال، عددا من الأدوات التى يمكن استخدامها لإنشاء قاعدة بيانات علاقية حتى لو كان الشخص لا يتمتع بخبرة كبيرة مع التصميم العلاقى. على سبيل المثال، يمكن استخدام أداة Wizard لإنشاء عشرة أنواع من قواعد البيانات المعرفة مسبقا من قاعدة بيانات ترتبط بالوقت والتصرف المعين. وعندما تتوافر البيانات فى جدول إلكترونى Spreadsheet أو جداول أخرى من أى نوع ولكن تتكرر البيانات غير المقننة Unnormalized. ويتطلب فصل البيانات فى جدولين أو أكثر لقاعدة بيانات أكسس ACCESS، ويمكن استخدام أداة محلل جدول Table Analyzer Wizard للمساعدة فى تقرير الحقول التى تحتاج للتحرك فى الجداول المنفصلة.
2. **عملية تصميم قاعدة البيانات**: الأساس فى فهم عملية تصميم قاعدة البيانات يقع فى فهم نظام إدارة قاعدة البيانات العلاقية MS. ACCESS الذى يخزن البيانات، ويقدم المعلومات المنتجة بكفاءة وفعالية. كما يحتاج أيضا نظام إدارة قاعدة البيانات العلاقية هذا تضمين الحقائق المرتبطة بالموضوعات المختلفة المخزنة فى جداول منفصلة بعضها عن بعض. على سبيل المثال، قد يوجد جدول واحد يختزن كل الحقائق عن الطلاب فقط، وجدول آخر يختزن بيانات المعلمين فحسب... الخ. وعندما تستخدم البيانات، تجمع الحقائق معا وتعرض فى طرق مختلفة، كطبع التقارير التى تجمع الحقائق عن المعلمين والطلاب فى نفس الوقت.

وعند تصميم قاعدة البيانات المطلوب الاحتفاظ بها كموضوعات منفصلة؛ فإن نظام إدارة قاعدة البيانات العلاقية يخبر ويعرف الموضوعات الأخرى المتضمنة حتى يمكن إمداد أو إتاحة المعلومات الصحيحة معا عند الحاجة إليها.

3. **خطوات تصميم قاعدة البيانات العلاقية**: فيما يلى عرض للخطوات المتضمنة فى عملية تصميم قاعدة بيانات علاقية:

- **الخطوة الأولى – تقرير الغرض من قاعدة البيانات**: حيث إن ذلك سوف يساعد فى تقرير أى الحقائق المحتاج إليها للتخزين.

- **الخطوة الثانية – تقرير الجداول المحتاج إليها**: بمجرد تحديد أغراض واضحة لقاعدة البيانات، يمكن تقسيم المعلومات فى موضوعات أو مجالات منفصلة قائمة بذاتها كالطلاب، المدرسين، المواد الدراسية، الفصول... الخ. وكل مجال من هذه المجالات سوف يخصص له جدول مستقل فى قاعدة البيانات.

- **الخطوة الثالثة – تقرير حقول البيانات المحتاج إليها**: تختص بتقرير المعلومات المطلوب حفظها فى كل جدول. وكل مجموعة معلومات فى الجدول يطلق عليها حقل Field يعرض كعامود من أعمدة الجدول. على سبيل المثال، يمكن تخصيص أحد حقل بيانات جدول الطلاب لإسم الطالب، وحقل آخر لتاريخ ميلاده، وثالث لنوعه "ذكر أو أنثي"... الخ.

- **الخطوة الرابعة – تقرير العلاقات**: يمكن النظر فى كل جدول من الجداول المتاحة فى قاعدة البيانات وتقرير كيف أن بيانات أحد الجداول ترتبط ببيانات جدول آخر، كما يضاف حقول للجداول لإنشاء جداول جديدة لتوضيح العلاقات عند الضرورة.

- **الخطوة الخامسة – تعزيز وتحسين التصميم**: يحلل التصميم المتاح بهدف التعرف على أى أخطاء من أجل استبعادها. وتنشأ الجداول ويضاف إليها سجلات بيانات جديدة، وعند التوصل للنتائج المستهدفة من الجداول ، يصبح التصميم

متقنا وجيدا أو قد يستلزم تكييف وتطوير وتعديل النظام ذاته عندما توجد أخطاء فيه أو تترك أشياءًا من التصميم التمهيدى. وبذلك يجب التفكير فى أن التصميم الأولى يشبه المسودة الأولية المحتاجة للمراجعة بهدف التصحيح. كما يجب تجريب واختبار عينة من البيانات والنماذج التمهيدية المرتبطة بالأشكال والتقارير. وباستخدام نظام إدارة قاعدة البيانات يصبح من السهل تغيير تصميم قاعدة البيانات المطلوب إنشائها. على أى حال، تصبح قاعدة البيانات أكثر صعوبة فى عمل التغييرات الرئيسية فى الجداول بعد ملء بياناتها بعد إنشاء أشكالها وتقاريرها، لذلك يجب التأكد مسبقا من صحة ودقة التصميم قبل بدء العمل.

4. **مشكلات التصميم المشتركة والشائعة**: توجد مشكلات عديدة مشتركة وشائعة التى يمكن التغلب عليها عند تصميم قاعدة البيانات. وقد تؤدى هذه المشكلات لصعوبة استخدام البيانات وحفظها. ويوضح العرض التالى بعض الإشارات المطلوب اعتبارها فى تقييم قاعدة البيانات:

- قد يوجد جدول واحد مع عدد كبير من الحقول التى لا تترابط كلها مع نفس الموضوع أو المجال. على سبيل المثال، قد يشتمل أحد الجداول على حقول ترتبط بالطالب، وحقول أخرى ترتبط بالمعلمين، بالإضافة لحقول معلومات عن المواد الدراسية. على ذلك يجب التأكد من أن كل جدول يشتمل على بيانات عن موضوع واحد.
- توجد حقول تترك خالية فى كثير من السجلات، لأنها غير مرتبطة بها، ويعنى ذلك أن الحقول تتبع جدولا آخر.
- يوجد عدد من الجداول أكبر كثيرا ممن يشتمل على نفس الحقول. على سبيل المثال، توجد جداول منفصلة للمواد الدراسية فى الفصل الدراسى الأول، وأخرى فى الفصل الدراسى الثانى، وحتى الفصل الدراسى الصيفى، وتوجد

أيضا جداول للطلاب المنتظمين وأخرى للطلاب المنتسبين، أو لطلاب التعليم المفتوح... الخ يخزن فيها نوعا من المعلومات المحتاج إليها. لذلك يجب تماسك كل المعلومات المرتبطة بموضوع واحد فى الجداول المعينة. وقد يحتاج أيضا لإضافة حقل آخر لتعريف بيانات الفصل الدراسى أيضا.

5. **تقرير الغرض من قاعدة البيانات**: الخطوة الأولى فى تصميم قاعدة بيانات هى تقرير الغرض من إنشائها وكيفية استخدامها. وبذلك يمكن تقرير الموضوع أو المجال المحتاج لتزين حقائق عنه فى الجداول، وما هى الحقائق المحتاجة للتخزين فى حقول هذه الجداول. وقد يتم ذلك من خلال أساليب جمع البيانات التى ترتبط على سبيل المثال بالمناقشة والعصف الذهنى مع مستخدمى قاعدة البيانات المتوقعين، وتحديد أطر المخرجات والتقارير المطلوب إنتاجها، وجمع الأشكال أو استمارات تقصى الآراء المستخدمة لتسجيل البيانات. وتستخدم كل هذه البيانات فى خطوات عملية التصميم الباقية. مثال ذلك، متابعة الطلاب أثناء الدراسة ، أو تحديد المخزون المتبقى من الكتب المدرسية... الخ.

6. **تقرير الجدول المحتاج إليه**: تقرير الجداول فى قاعدة البيانات قد تكون الخطوة الأساسية فى عملية تصميمها، لأن النتائج المطلوبة من قاعدة البيانات كالتقارير المطلوب طبعها، الأشكال والنماذج والكشوف المطلوب استخدامها، والأسئلة المراد الإجابة عليها... الخ، لا تقدم بالضرورة مفاتيح عن هيكل الجداول التى تنتجها، بل تبين ما يراد معرفته لا فقط فى كيفية جمع المعلومات فى جداول. وتشتمل الجداول على حقائق عن الطلاب مثلا فيما يتصل بعناوينهم، أرقام تليفوناتهم، تواريخ ميلادهم، أنواعهم، وغير ذلك من الحقائق الأخرى. ويقدم الشكل أو النموذج عددا من الحقائق المراد تخزينها فى قاعدة البيانات. وقد يضاف تقديم أخطاء فى تطبيق المعلومات، لأنه قد يفترض أن يسجل الطالب فى ثلاث أو أربع مواد دراسية، يتضمن كل منها حقائق عنه تتكرر فى كل هذه المواد الدراسية. وبافتراض أن أحد الطلاب سجل فى مقرر معين ولكنه ألغاه أو

انسحب منه بعد ذلك، وعند حذف سجل هذا المقرر الدراسى من الجدول الذى يتضمن معلومات عن كل الطلاب تحذف كل البيانات التى تتعلق به. وفى حالة حفظ بيانات عن هذا الموقع فى قاعدة بيانات النظام حتى يمكن إرسالها إلى المدرس، فإنه مرة أخرى يستحسن وضع معلومات الطالب المعين فى جدول منفصل خاص بالطلاب. ومن خلال ذلك، يمكن حذف المقرر الدراسى بدون حذف معلومات الطالب نفسه.

7. **تقرير الحقول المحتاج إليها**: لتقرير الحقول فى الجدول، يجب تحديد ما يحتاج لمعرفته عن الأفراد أو الأحداث المسجلة فى الجدول. ويمكن التفكير فى الحقول لخصائص ترتبط بالجدول المعين. وكل سجل أو صف من صفوف الجدول يشتمل على نفس مجموعة الحقول أو الخصائص. على سبيل المثال، حقل عنوان فى سجل أو ملف الطالب يشتمل على عناوين الطلاب. وكل سجل من سجلات الجدول أى من صفوفه يتضمن بيانات عن الطالب منها حقل بيانات عن عنوانه المتضمن عناوين كل الطلاب.

8. **حقول المفتاح الأصلى**: Primary Key Fields تأتى كفاءة وقدرة نظام إدارة قاعدة البيانات من سرعة التعرف على إيجاد المعلومات المخزنة فى الجداول المختلفة معا وإحضارها من أجل الاستخدام. ولكى يعمل نظام إدارة قاعدة البيانات بكفاءة عظمى، يجب أن يشمل كل جدول فى قاعدة البيانات على حقل فريد أو مجموعة حقول تعرف بطريقة فريدة كل سجل مختزن فى الجدول. وفى الغالب يعتبر ذلك رقم تعريف فريد، مثل رقم الطالب، المدرس، الفصل... الخ. ويطلق على رقم التعريف الفريد هذا المفتاح الأصلى Primary Key بالجدول. وتستخدم نظم إدارة قاعدة البيانات المختلفة حقول المفتاح الأصلى لتسريع ترابط البيانات فى الجداول المتعددة ولإحضار البيانات معا. وعندما يوجد مؤشر أو معرف فريد Identifier للجدول، مثل مجموعة أرقام تخصص دراسى، ويمكن معرفة هذا التخصص من بين مجموعة التخصصات المرتبطة به.

كما يمكن استخدام هذا المؤشر / المعرف كمفتاح أصلى للجدول، إلا أنه يجب التأكد أن قيم هذا الحقل سوف تكون دائما مختلفة لكل سجل. ويلاحظ أن نظام إدارة قاعدة البيانات كنظام أكسسMS. ACCESS لا يسمح بقيم مكررة فى حقل المفتاح الأصلى. على سبيل المثال، لا تستخدم أسماء الطلاب أو المعلمين أو المواقع... الخ، كمفاتيح أصلية، لأن الأسماء ليست فريدة فى حد ذاتها، إذ يمكن بسهولة إيجاد شخصين أو موقعين جغرافيين بنفس الإسم فى نفس الجدول. وعندما لا يحدد المؤشر ، المعرف الفريد للجدول يمكن استخدام حقل يسجل الأرقام بصفة تتابعيه ببساطة. وعند اختيار المفاتيح الأصلية، يجب مراعاة التالى:

- لا تسمح معظم نظم إدارة قاعدة البيانات بتكرار القيم أو عدم وجودها فى حقل المفتاح الرئيسى. لهذا السبب يجب اختيار مفتاح أصلى يشتمل على هذه القيم.
- قد تستخدم القيمة فى حقل المفتاح الأصلى للعثور على السجلات، لذلك يجب ألا يكون هذا الحقل كبيرًا صعب التذكر أو الطبع.
- حجم المفتاح الأصلى يؤثر على سرعة العمليات فى قاعدة البيانات. وعند إنشاء حقول للمفتاح الأصلى، يمكن وضع خاصية تحد من حجم الحقل. وبذلك يستخدم الحجم الأصغر الذى يستوعب كل القيم المحتاج إلى تخزينها فى الحقل.

9. **تقرير العلاقات**: بعد تقسيم المعلومات فى جداول قاعدة البيانات، يحتاج لطريقة تعلم نظام إدارة قاعدة البيانات كيفية استعادتها مرة أخرى بطرق مفهومة. وحيث إن هذه النظم هى نظم علاقية فى الأساس، فيعنى ذلك تخزين البيانات فى جداول منفصلة بعضها عن بعض. عندئذ تفسر العلاقات بين الجداول وتستخدم العلاقات لاستعادة وإيجاد البيانات المترابطة المخزنة فى قاعدة البيانات. على سبيل المثال، عند افتراض الرغبة فى مكالمة هاتفية مع طالب آخر أو مع المعلم للسؤال عن مقرر دراسى معين، فإن أرقام تليفونات الطلاب

والمعلمين تسجل فى جدول الطلاب وجدول المعلمين وجدول المقررات الدراسية. وعند إخطار نظام إدارة قاعدة البيانات عن أى مقرر، فإنها تهتم بذلك، وتحصل على رقم هذا المقرر وتربطه بأرقام تليفونات الطلاب والمعلمين بناء على العلاقات بين الجداول. ويؤدى ذلك بربط رقم تليفون الطالب مع المفتاح الأصلى له الذى يكون أيضا فى جدول المقررات الدراسى، أى أن حقل تعريف الطالب فى جدول المقررات الدراسية الذى يطلق عليه المفتاح الأجنبى Foreign Key، حيث يمثل مفتاحا مستوردا من جدول آخر. وعلى ذلك، فإنه لتحديد علاقة بين جدولين (أ،ب) يضاف مفتاح أصلى واحد لجدول آخر لكى يظهر الجدولان معا. ولكن كيف يتمكن تقرير أى مفتاح أصلى لجدول يجب استخدامه؟. لتوضيح العلاقة بطريقة صحيحة، يجب أولا تقرير طبيعة هذه العلاقة. وكما سبق توضيحه وجد ثلاثة أنواع من العلاقات بين الجداول، هى: العلاقات واحد لواحد، العلاقات كثير لواحد، والعلاقات كثير لكثير. على سبيل المثال، عند إنشاء العلاقة واحد لكثير التى تعتبر نوع العلاقات الأكثر شيوعا فى قاعدة البيانات العلاقية، فإن السجل فى جدول أ يمكن أن يكون له أكثر من سجل مضاهى فى جدول ب. ولكن السجل فى جدول ب له سجل مقارن واحد فى جدول أ. كما أنه عند إنشاء العلاقة كثير لكثير، فإن السجل فى جدول أ يمكن أن يكون له أكثر من سجل مضاهى فى جدول ب، وسجل فى جدول ب يمكن أن يكون له أيضا أكثر من سجل مضاهى فى جدول أ. هذا النوع من العلاقات يتطلب تغييرات فى تصميم قاعدة البيانات قبل تحديد العلاقة بأنها صحيحة فى نظام إدارة قاعدة البيانات.

6 - بزوغ نظم مستودعات البيانات:

1/6 تطور تكنولوجيا قواعد البيانات:

العرض المختصر التالى يوضح التطور التاريخى لتكنولوجيا قواعد البيانات حتى بزوغ نظم مستودعات البيانات.

1. **فى الستينيات** من القرن العشرين الماضى: ظهرت إلى الوجود نظم معالجة التصرفات Transaction Processing Systems (TPS) اعتمدت واتسمت بالتالى:

- تقارير حزم البيانات.
- صعوبة الحصول على المعلومات وتحليلها.
- جمود وعدم مرونة النظم والتكلفة الباهظة لإعادة برمجة كل تصرف جديد.

2. **فى السبعينيات**: ظهرت نظم إدارة قاعدة البيانات DBMS للقيام بجمع البيانات، إنشاء قواعد البيانات، نظم المعلومات الإدارية وشبكة نظام إدارة قاعدة البيانات. وارتبطت نظم إدارة قواعد البيانات بعدة أنواع منها الشبكية والعلاقية والشيئية. كما بزغت نظم دعم القرار DSS ونظم معلومات الرؤساء أو مديرى الإدارة العليا EIS المبنية على النهايات الطرفية وغير المكتملة التى لم تتكامل مع الأدوات التى أتيحت فيما بعد على قمة الكمبيوتر.

3. **فى الثمانينيات**: بزغت أدوات الوصول والتحليل المتعمق على قمة الكمبيوتر التى منها:

- أدوات التساؤل، الجداول الإلكترونية، واجهات التفاعل الرسومية.
- نماذج إدارة قواعد البيانات المتقدمة الممتدة، الشيئية، الاستنتاجية... الخ.

4. **فى التسعينيات**: بزغت نظم مستودعات البيانات التى تتضمن:

- نظم دعم القرار.
- تنقيب البيانات وإعادة التخزين.
- قواعد بيانات النصوص، وقواعد الوسائط المتعددة، وقواعد البيانات الذكية.

2/6 خلفية نظم مستودعات البيانات:

نظم مستودعات البيانات تمثل التحول من النظم التى يطلق عليها "نظم قواعد البيانات الموزعة" التى تتسم بالفورية والوصول بشفافية لكل البيانات فى كل الأوقات. وقد اهتم بتكنولوجيا نظم مستودعات البيانات التى نمت بثبات وديناميكيا، منذ التسعينيات فى أواخر القرن الماضى. وقد بزغت نظم مستودعات البيانات نتيجة للفشل فى تطبيق نظم إدارة قواعد البيانات الموزعة Distributed Database Management Systems (DDBMS).

ومنذ الثمانينيات من القرن الماضى قامت كثير من المؤسسات الرائدة فى تطوير نظم إدارة قاعدة البيانات مثل Digital Equipment, IBM, Oracle, Microsoft, etc. بمحاولات بحثية وتطويرية فى حل مشكلات جزر البيانات المتفرقة من خلال نظم إدارة قواعد البيانات الموزعة. وبزغ من هذه المحاولات منتجات استهدفت تقديم خدمات تتسم بشفافية الموقع وتجزيء البيانات عبر جداول طبيعية متفرقة جغرافيا تقوم بتداول الأساليب غير المتجانسة (المتمثلة فى منتجات ونماذج نظم إدارة قاعدة البيانات المختلفة) مثل خطط دولية محايدة النموذج. ويعنى ذلك أن المنشأة أو المنظمة تقوم بأداء أى شيء تريد عمله على أساس محلى فيما يتصل باختيار المنتج المستهدف والتكنولوجيا، وتفسيرات البيانات، وسياسات الإدارة. وصارت نظم إدارة قواعد البيانات الموزعة العمل السحرى الذى يربط كل جزر البيانات معا ويقدم بنية أساسية من خلالها يمكن لأى فرد فى المنظمة الوصول للبيانات لدعم القرار وحاجات المعلومات الأخرى. وقد كانت أسباب فشل نظم إدارة قواعد البيانات الموزعة عديدة، وأدى ذلك للتحرك نحو نظم مستودعات البيانات فى بداية التسعينيات. ومع مشكلة جزر البيانات المتفرقة التى كانت ما زالت متواجدة وتمثل صعوبة كبيرة، ظهر الاهتمام المتزايد بتكنولوجيا الكمبيوتر الموزعة مثل التطبيقات المبنية على شبكات الكمبيوتر المحلية LANs وشبكات الكمبيوتر

العريضة WANs والشبكات المؤسسية Intranet وشبكة الإنترنت Internet وبصفة خاصة لحلول الإدارات المختلفة بالمنظمة فيما يتعلق بنظم Data Marts.

يجب تذكر أيضا أن بداية عقد التسعينيات كان حقبة التحول من الحاسبات الكبيرة إلى الحاسبات الشخصية، وأصبح من السهل تصور السيناريو الذى كان عليه آخر القرن العشرين وبداية القرن الحادى والعشرين الذى يوضح أن معظم المنشآت والمنظمات، ومنها المنظمات التعليمية كالمدارس والجامعات كانت وما زالت تستخدم تطبيقات عديدة مشغلة على الحاسبات الكبيرة أو معتمدة على النظم المركزية، كما صارت أصول بيانات المنظمة موزعة بين الخادمات الموزعة جغرافيا والمعتمدة على شبكات الكمبيوتر المحلية والعريضة.

وبذلك انبثقت تكنولوجيا مستودعات البيانات من مفهوم بسيط وواضح نسبيا يتمثل فى الإجابة على حل المشكلة المتضمنة فى التساؤل التالى: عند وصل البيانات المخزنة غير المتجانسة الموزعة فى وقت التشغيل بطريقة صعبة ومشكوك فيها، لماذا يصعب استرجاع البيانات الضرورية لموقع واحد؟ وقد أدت الإجابة على ذلك الذى ارتبط بنظم قواعد البيانات الموزعة لبزوغ نظم مستودعات البيانات التى تتضمن كميات ضخمة من البيانات التى صارت ضرورية وناضجة ترتبط بعمليات الاستخلاص والتحليل والنقل على الخط.

وفى نطاق تكنولوجيا نظم مستودعات البيانات صارت للتطبيقات صدى واضح وترجيح للبيانات لم تكن متاحة من قبل فى نظم إدارة قواعد البيانات الموزعة التى تحد وتعرقل تتسم بعدة عوامل منها: مكان حفظ التكنولوجيا حتى تجزيئها، مخططات عمومية، بروتوكولات منفذة فى عدة مراحل. وقدمت التكنولوجيا المتقدمة خدمات برمجيات وسيطة لمستودعات البيانات التى تساعد فى استخلاص البيانات المختارة وتحويلها وتأكيد جودتها وحركتها بين بيئات أصبحت فجأة مكونات أساسية لخطط المنشأة أو المنظمة الإستراتيجية طويلة الأجل. ظهر

فى ظل نظم مستودعات البيانات مفهوم يرتبط بوحدات تنظيمية للمعلومات يسمى Data Mart الذى يرتبط بإعادة توزيع البيانات الجاهزة فى مستودع البيانات طبقا للغرض الوظيفى المعين. وفى نفس الوقت، وجدت النماذج والمعماريات التى اتصلت جذورها بنظم إدارة قاعدة البيانات الموزعة كما فى حالة تجزيء الجداول فى مكانها المناسب الخاص بتكنولوجيا خادم قاعدة البيانات الموازى. وتتمثل التكنولوجيات الجديدة التى صارت تؤثر على مستودعات البيانات الحديثة فى التالى:

- مخزن البيانات التشغيلى وهو بيئة بيانات تعريفية عن الوقت الحقيقى القريب.
- مستودعات البيانات الافتراضية التى تمثل بيئات بيانات تعريفية مع استخلاص وإعادة تخصيص بيانات كبيرة الحجم ملخصة.
- ذكاء أعمال الوسائط المتعددة المتقدمة من حيث التزاوج بين مفهوم التخزين والتكنولوجيا المبنية على الإنترنت لتقديم قدرات أعظم لإمداد المعلومات بغض النظر عن كيف تخزن أنواع المعلومات.

7 - مفهوم نظم مستودعات البيانات

1/7 ما مستودع البيانات؟

مستودع البيانات يمثل مخزن بيانات فردى وكامل ومترابط مكون من تنوع مصادر بيانات مختلفة متوافرة للمستخدم النهائى الذى يمكنه فهما واستخلاصها فى سياق الأعمال المتنوعة. ويمثل قاعدة بيانات كبيرة جدا، ويقوم بعملية تحويل البيانات إلى معلومات وجعلها متوافرة للمستخدمين فى توقيت فورى. وبذلك فمستودع البيانات يمثل وسيلة لتجميع البيانات من مصادر عديدة وإدارتها بغرض إجابة أسئلة عن الأعمال المؤداة أو المطلب أدائها، وبذلك يشتمل على اتخاذ القرارات التى لم تكن ممكنة من قبل، أى أن مستودع البيانات يكون موجها موضوعيا، متكاملا، متنوع الوقت، غير قابل للضرر، تجمع البيانات به للاستخدام فى اتخاذ القرارات.

يعتبر مستودع البيانات قاعدة بيانات تتسم بالخصائص المميزة التالية:

- أداة للمستخدمين، حيث يسمح لهم فحص البيانات التاريخية، وتحليل هذه البيانات فى نوع من الطرق، واتخاذ القرارات المبنية عليها.
- نظام مستقل عن قواعد البيانات التشغيلية يستخدم تنوع تكنولوجيات من البرمجيات كمحركات قاعدة البيانات وأدوات التساؤل، إلى الأجهزة كخادمات المعالجة المتعددة.
- أداة لمساعدة اتخاذ القرارات وتحليل المعلومات، من حيث:
  - عدم إمكانية مستودع البيانات فى اتخاذ قرار للبشر.
  - تقديم معلومات تساعد للتوصل لمؤشرات تمكن من اتخاذ قرارات ذكية ومعلوماتية.
  - تضمين بيانات من عبر المنظمة، حيث تلخص غالبا للمستوى الذى يؤدى لسرعة وكفاءة الاسترجاع.
- موجه نحو الموضوع أى يقدم رؤية بسيطة ومختصرة لها قيمة حقيقية عن موضوع أو مجال أو أكثر مختار لمساندة عملية اتخاذ القرارات.
- منشأ بواسطة تكامل مصادر بيانات غير متجانسة ومتعددة عبر المنظمة.
- مشتمل على بيانات تاريخية تتخلل أفقا زمنيا أطول مما يتاح فى قواعد البيانات التشغيلية.

وبذلك فإن مستودع البيانات يستهدف التالى:

- القدرة للوصول لبيانات المنظمة كلها؛ حيث إنه يعمل على:

* جمع البيانات من إدارات المنظمة المختلفة.

* قد تشتمل إدارات المنظمة على بيانات مخزنة من أشكال متنوعة من قواعد بيانات حاسبات كبيرة ومتوسطة للجداول الإلكترونية وملفات معالجة البيانات.

* يحرك مستودع البيانات كل البيانات المتاحة من مصادر غير متجانسة إلى موقع مركزى.
* يسمح للمنشاة أو المنظمة من الوصول لكل بيانات أعمالها وتطبيقاتها المترابطة.

- القدرة على أداء التحليل المتعمق بسرعة، حيث إنه بمجرد تضمين البيانات فى المستودع فإنها تتناسق داخليا، ويمكن التساؤل فيها لأداء التحليل.
- يرتبط التحليل بإعادة ترتيب البيانات وأداء التنقيب فيها للبحث عن الاتجاهات التى لم تكن معروفة من قبل.

وعند التساؤل عن **لماذا يبنى مستودع البيانات؟** نجد أن ذلك يرتبط بالأسباب التالية:

- تماسك موارد المعلومات فى المنظمة.
- تحسين أداء التساؤل.
- فصل وظائف البحث واتخاذ القرارات من النظم التشغيلية.
- أساس تنقيب البيانات لاكتشاف المعرفة، ورؤية البيانات إلكترونيا، والتحليل المتعمق للتصرفات، وإيجاد علاقات وارتباطات... الخ.

من العرض السابق لمفهوم مستودع البيانات يمكن استخلاص **الأبعاد والخصائص** التالية التى يتسم بها:

1. **التماسك والترابط**: ويعنى **التماسك** جمع البيانات من خلال كل إدارات وأقسام المنظمة للتخزين فى موقع واحد. أما **الترابط** فيعنى أن مستخدمين مختلفين سوف يحصلون على نفس النتائج لنفس التساؤل للعرض فى أوقات مختلفة، كما يعنى أيضا أن البيانات متوافقة فى مستودع البيانات بغض النظر عن كيف تخزن فى مصادرها الأصلية.
2. **الوجهة الموضوعية**: يجب أن تتضمن المعلومات التعليمية فى مستودع البيانات

على معلومات الإدارة التعليمية وما يرتبط بها من معلومات العملية التعليمية ذاتها الخاصة بالمقررات الدراسية، ويجب أن تتحرك هذه البيانات الموجهة موضوعيا نحو مستودع البيانات وبمجرد تواجدها فيه ينظم بناء على الموضوع المعين.

3. **التوجه التاريخى**: تعتبر بيانات مستودع البيانات ذات توجه تاريخى، مما يعنى أنها لا تتغير عبر الزمن إن لم توجد مشكلة مع البيانات فى المصدر. وبذلك تغطى هذه البيانات المتاحة فى المستودع فترة زمنية طويلة، بينما نظم معالجة التصرفات على الخط OLTP، فإنها تعتبر بيانات حديثة فقط. وتسمح البيانات فى فترة زمنية طويلة تحليل الاتجاهات المختلفة عبر الزمن، وتتضمن أيضا الاتجاهات الحديثة بجانب البيانات المتوسطة والطويلة الأجل.

4. **صبغة القراءة فقط**: حيث إن البيانات المتضمنة فى مستودع البيانات ذات توجه تاريخى فإنها تقرأ فقط، إلا أنها قد تتغير فقط فى حالة تواجد أخطاء فى المصدر الأصلى لها. وحيث إن البيانات المتواجدة فى المستودع لا تحدث أو تحذف؛ فإنه لذلك يمكن هيكلته للتعامل مع التساؤلات واستخدام الكشافات بسرعة ومرونة كافية.

2/7 الفرق بين النظم التشغيلية ونظم مستودعات البيانات:

العرض التالى يبين الفرق بين نظم المعلومات التشغيلية التى تساندها نظم قواعد البيانات ونظم مستودعات البيانات الحديثة الذى يتمثل فى الأبعاد التشغيلية الآتية:

- الأداء: نظم المعلومات التشغيلية المساندة بنظم قواعد البيانات مصممة لمعرفة الأبعاد التعليمية المختلفة وتحميلاتها. أما تساؤلات نظم المعالجة التحليلية على الخط OLAP المتعلقة بمستودعات البيانات فقد تقلل الأداء بالنسبة للعمليات التشغيلية، على الرغم من أن طرق تنظيم البيانات والوصول إليها وتنفيذها تحتاج لآراء وتساؤلات متعددة الأبعاد.

- الوظائف: نقص البيانات المتاحة يتطلب دعم القرارات ببيانات تاريخية لا تحتفظ بها نظم قواعد البيانات التشغيلية، كما أن تماسك البيانات فيما يتعلق بدعم القرارات يتطلب تماسك البيانات من حيث تراكمها وتلخيصها التى تنبع من مصادر غير متجانسة فى حالة نظم قواعد البيانات التشغيلية والمصادر الخارجية. وفيما يرتبط بجودة البيانات التى تستخدم من مصادر مختلفة تمثيلات وأكواد وأشكال بيانات غير متوافقة يجب تسويتها، التى تعمل على تحقيقها نظم مستودعات البيانات.
- النظم التشغيلية: يمكن تحديد أبعاد النظم التشغيلية التى تعرف بنظم بمعالجة التصرف على الخط OLTP التى تستخدم نظم إدارة قاعدة لبيانات وخاصة العلاقية منها فى التالى:
  - تشغل المهام الحرجة المرتبطة بالتفرق والاستعادة والسلامة.
  - تستخدم البيانات الحديثة التفصيلية.
  - تتضمن بيانات مكررة هيكليا.
  - تحتاج للعمل مع متطلبات المهام الكتابية الروتينية وتحتوى على قراءة وتحديث سجلات قليلة.
  - تستخدم لتشغيل المهام والأعمال التى تحدث فى المنظمة وخاصة التعليمية.
  - تشغل الأعمال والمهام فى الوقت الحقيقى.
  - مبنية على بيانات ثانوية.
  - معظَّمة لتداول تطبيقات القراءة والكتابة البسيطة.
  - تستخدم بواسطة أفراد يتعاملون مباشرة مع الطلاب مثلا.
  - تستخدم بصفة متزايدة من قبل المعلمين والطلاب والإداريين والفنيين بالمدرسة مثلا.

ويحدد الجدول التالى معالم وأبعاد الفرق بين نظم معالجة التصرف على الخط ونظم مستودعات البيانات:

جدول رقم (1/6) الفرق بين معالجة التصرف على الخط ومستودعات البيانات

معالجة التصرف على الخط	مستودعات البيانات
موجهة نحو التطبيق (التصرف والفعل)	موجهة نحو الموضوع
تتضمن بيانات حالية آنية	تتضمن بيانات قديمة تاريخية
تستخدم لتشغيل الأعمال	تستخدم لتحليل اتجاهات الأعمال
تشتمل على بيانات حديثة على الدوام	يتم تحديث البيانات فى نطاق حزم
البيانات المتوافرة مفصلة ومنفصلة	بيانات مختصرة ومتكاملة
الوصول للبيانات يعتبر تكراري	الوصول العشوائى للبيانات
توجه للمستخدم الكتابى	توجه لمستخدم المعرفة كالمدير
تشتمل على هيكل جداول مقننة (كثير من الجداول، أعمدة قليلة)	تشتمل على هيكل بيانات تهيئة (جداول قليلة، أعمدة كثيرة لكل جدول)
تمثل الأداء المتسم بالحساسية	ترتبط بالأداء الملطف Relaxed
تشتمل على سجلات قليلة يتوصل إليها فى الوقت الحقيقي	تشتمل على أحجام كبيرة من البيانات يتوصل إليها فى وقت معين
تتضمن خاصية القراءة والتحديث	تتضمن القراءة فقط وتحديث الدفعة المتكاملة
لا تشتمل على بيانات مسهبة	يوجد إسهاب فى البيانات المتضمنة
يعتبر حجم قاعدة بيانات التطبيق من 100 ميجابايت إلى 100 جيجابايت	حجم المستودع أكثر من 100 جيجابايت وصولا لعدة تيرابايت
تكون الاستفسارات بسيطة وقليلة	الاستفسارات صعبة ومعقدة
يعتبر إتمام التصرف قياس الأداء	إتمام التساؤل هو قياس الأداء
تتعامل مع آلف من المستخدمين	تتعامل مع مئات من المستخدمين فقط
تتضمن شمولية الإدارة	تتضمن الإدارة بواسطة مجموعات فرعية

مما سبق يتضح أن نظم معالجة التصرفات على الخط هى نظم مستخدمة لتشغيل الأعمال، أما نظم مستودعات البيانات؛ فإنها تساعد فى تعظيم الأعمال وتحليلها موضوعيا.

8 - أنواع مستودعات البيانات:

توجد عدة أنواع لمستودعات البيانات التى قد ترتبط بمخازن البيانات التشغيلية والمعالجة التحليلية على الخط ومعالجة التصرف على الخط وبيانات الوحدة التنظيمية والمستودعات الافتراضية والمرتبطة بذكاء الأعمال. والعرض التالى يحدد معالم كل نوع باختصار:

1/8 مخزن البيانات التشغيلية: Operational Data Store (ODS)

يمكن اعتبار أحد الاختلافات الأكثر دراماتيكية بين توقعات تكنولوجيا نظم إدارة قاعدة البيانات ومستودعات البيانات كما تزاوله وتنفذه معظم المنظمات فى الوقت الحالى ما يتمثل فى تباطؤ الوقت أو فجوة الوقت Time Lag بين أى إدخال بيانات التصرف وتوافرها للأغراض التحليلية، وغير ذلك من الأهداف المعلوماتية الأخرى.

وسواء قام مستودع البيانات المعين بتنفيذ تداول دورة استخلاص البيانات، التحويل، الحركة والتحميل على أساس يومى، أسبوعى، شهرى، ربع سنوى... الخ. فسوف يكون قابلا للجدل، لأن بيانات النشاط التشغيلى الخاص بالتصرف قد لا تتوافر للتحليل الفورى. كما أن نماذج إعادة تخزين البيانات السريعة على أساس يومى أو أثناء اليوم لبيئات أصغر فى مواجهة تلك الأسبوعية أو الشهرية، ما زالت تبين تأخر حقيقى يمكن ملاحظته فى توافر البيانات. ولا يمثل ذلك مشكلة لكثير من حاجات أعمال المنظمات الكبيرة. كما تعتبر مستودعات البيانات المستخدمة بطريقة نموذجية للتحليل التاريخى تتطلب مجموعة بيانات كاملة، على سبيل المثال ملخصات شهرية عن تحصيل الطلاب أو غيابه. إلا أنه قد تتواجد قيمة محددة

للأعمال فى تضمين المعلومات المتوافرة غير الكاملة كأداء التكليفات أو التدريبات... الخ.

على أى حال، قد تتطلب أوجه ذكاء الأعمال والحاجات المعلوماتية فى المنظمة المعينة كالمدرسة وصولا شبه فورى للبيانات من مصادر تطبيقات عديدة عبر المنظمة. وفى هذه الحالات، فإن تباطؤ الوقت الموروث فى مستودعات البيانات الكلاسيكية لا يمكن قبوله تقديم صورة شبه مفصلة فوريا عن الطالب وأطراف العملية التعليمية الآخرين من أى قرارات يمكن اتخاذها استجابة للنشاط المقترح. هذا المتطلب قد لا يكون فى حد ذاته قيدًا كبيرًا أو أى شيء خاص، إلا أن العوامل التالية قد تعقد الوضع فى حالة الطالب بالمنظمة التعليمية وفقًا لما يلى:

- الطلاب وأطراف العملية التعليمية الآخرين هم فى الواقع يشكلون شبكات وهرميات معقدة لكيانات المدرسة على سبيل المثال.
- تواجد تطبيقات غير متجانسة التى قد تنتشر وتتفرق فى كل مكان كما فى حالة تعليم الحضانة، التعليم الابتدائى، والإعدادى، والثانوى العام التى لكل منها وظيفة ومجال يدير وحدات أنشطته المختلفة. ولا يوجد تفكيك للتمييزات بين مجالات التطبيقات المتنوعة ترتبط بالمسئولية، ومعلومات متشابكة ومتداخلة لدرجة معينة فى إطار مستودع البيانات الكلاسيكى.
- أهمية قدرة مستخدمى المعلومات على أداء وظائف ذكاء الأعمال لا فقط تساؤل واسترجاع المعلومات، إلى جانب قدرتهم فى أداء المعالجة التحليلية على الخط OLAP كما فى حالة تنقيب البيانات للوصول لأراء المستخدمين من الطلاب مثلا من خلال سماتهم فى مواجهة كل البيانات المطبقة المتضمنة الأنشطة التى قد يوافق عليه فى أرجاء أخرى.

وفى هذا الصدد، يعتبر النشاط التاريخى مؤديًا بعضًا من الدور الأساسى فى بيئة مخزن البيانات التشغيلية الكمبيوترى للطالب أو المعلم، إلا أنه يعتبر من جهة أخرى

تضخيمًا للبيانات فيما يتعلق بالنظرة السريعة لكل نشاط فى تقرير الربع سنوى السابق مثلا المرتبط بتصنيف عملية اتخاذ القرارات البشرية، وذلك بدلا من بيانات المعالجة التحليلية على الخطط التاريخية الطويلة الأجل كاتجاهات نجاح إنجازات النشاط المقارن مع السنوات الخمس الماضية، أو ما يشبه ذلك من وظائف.

ويستطلع بعض مطورى مستودعات البيانات منتجاتهم فيما يتصل بقدرتها على مساندة بيئات مخازن البيانات التشغيلية بالإضافة إلى مستودعات البيانات التقليدية، ولكن عند النشر يتضح أنها تصف أكثر قليلا من القدرة على حركة البيانات المتغيرة (التى تتضمن إضافات وحذف) فى بيئة مستودعات البيانات بدون طلب الإنعاش أو التجديد بأن يكون الحذف أو الاستبعاد مصحوبا بإعادة تحميل بيانات المستودع. وعلى الرغم من القدرات الموجهة نحو التغيير التى تعتبر بالتأكيد إضافات مرحبًا بها لتتبع الخيارات المتوافرة للتنفيذ؛ فإن متطلبات التحديث فى شبه الوقت الحقيقى لمخازن البيانات التشغيلية الحقيقية ما زالت محدودة جدا.

وعلى الرغم من تلك الأوجه المفقودة فى مخازن البيانات التشغيلية لمستودعات البيانات ، فإن مجال البرمجيات الوسيطة Middleware الخاصة بالمراسلات تجعل مخازن البيانات التشغيلية ممكنة من خلال مساندة تطبيقات الاتصال المتداخلة وتبادل البيانات، كما تقدم برمجيات المراسلات الوسيطة خيارين معماريين مختلفين هما:

- خيار تفاعل نقطة لأخرى، الذى يعمل مع مصدر أو أكثر من مصدر تطبيقات يرسل الرسائل والبيانات الموجهة نحو التغيير الملائم لمخزن البيانات التشغيلية لإضافة، وتعديل، وتحديث أو الاستبعاد المطابق. ويعتبر هذا الاختيار ملائما جدا لتقدم تطبيقات محدودة البيانات (بحيث ألا تزيد على ثلاثة تطبيقات) لمخزن البيانات التشغيلية.
- خيار تجزيء المراسلة أو مكون البرمجيات الوسيطة الذى يرسل منه مصدر

التطبيقات الرسائل المعاد مسارها لمخزن البيانات التشغيلية، أو لمصدر تطبيقات آخر فى نمط النشر والاشتراك. وعلى الرغم من أن هذه المعمارية تعتبر أكثر صعوبة وتعقيدا من مجموعة واجهات نقطة لأخرى البسيطة، إلا أنها أكثر ملائمة عندما تشتمل البيئة المعينة على مصادر مختلفة وتطبيقات تكتيكية كثيرة.

وبصفة نموذجية، يجب على مطور التطبيقات أن يقدر على مساندة مستودع البيانات، أى قدرته على إمداد البيانات الملائمة بواسطة طلب خدمة مراسلة المستودع. وفى غياب هذه القدرة مع صعوبة تعديل التطبيقات المتوافرة بالفعل؛ فإن الوكلاء الخارجيين External Agents المتخصصين أساسا فى تقديم البرامج الروتينية للاستخلاص يمكن أن تكتشف البيانات وتعمل على تغييراتها وتداول تفاعلها مع بيئة المصدر والمراسلات.

وعلى الرغم من الصعاب السابق الإشارة إليها؛ فإن مخازن البيانات التشغيلية تعتبر معقدة فى المزاولة التطبيقية، إلا أنها تقلل من تأخر الوقت بين إنشاء البيانات وتوفيرها للمستخدمين.

2/8 قسم مستودع البيانات لإحدى إدارات المنظمة : Data Mart

من الملاحظ أن مستودع البيانات يغطى بيانات كل منظمة بإداراتها وأقسامها المختلفة العديدة. أما ما يطلق عليه Data Mart فيخصص لإدارة أو قسم أو موضوع معين تهتم به المنظمة ويوجد فيها. وتبنى الوحدة التنظيمية للبيانات المعينة Data Mart للأسباب التالية:

- عدم حاجة وظيفة أو إدارة معينة لكل بيانات مستودع البيانات الشامل فى المنشأة أو المنظمة.
- إمكانية تصنيف وقت للمعالجة من خلال تحديث الوحدة التنظيمية للمعلومات Data Mart مع بيانات جديدة فى العادة.

- تعنى الوحدات التنظيمية للبيانات Data Marts بتساؤلات أسرع لكل وحدة تنظيمية على حدة لوجود بيانات أقل يمكن اختيار إجابة التساؤل منها بسرعة.

والوحدات التنظيمية للبيانات Data Marts تتضمن بيانات مستقلة فى مواجهة البيانات المعتمدة من خلال التالى:

- إمكانية تبنى الوحدات التنظيمية للبيانات Data Marts فى مراحل مختلفة وتطويرها بواسطة فرق مختلفة أيضا.
- غاية فريق تطوير الوحدة التنظيمية للبيانات Data Mart هو التفسير المسبق لخصائص الأبعاد المختلفة المستخدمة خلال وحدات البيانات المتعددة.
- يتوافق إنشاء الوحدات التنظيمية للبيانات على إمكانية السماح بسحب البيانات بسهولة من مستودع بيانات المنظمة.
- يمكن أن تتساوى الوحدات التنظيمية للبيانات مع مجالات متعددة فى بنائها لإرضاء حاجات مستخدمين معينين.
- يمكن لهذه المزاولة منع تكامل مستقبلى وتكلفة تطوير متقدمة للاهتمام فى مشاركة المعلومات عبر إدارات المنظمة.

3/8 نظم المعالجة التحليلية على الخط: On-Line Analytical Processing (OLAP)

تعرف نظم المعالجة التحليلية على الخط بأنها:

- عملية مكتملة بواسطة أفراد وأدوات لتسهيل معلومات الأنشطة الهيكلية فى المنظمة التعليمية.
- تسمح للمستخدمين من الإبحار فى البيانات المتاحة من خلال الولوج لمستوى تفضيل أدنى.
- تقدم مستويات تحليل مختلفة خلال قدرات تساؤل معقدة ومقارنة البيانات وتنقيب البيانات والتقرير.

- مجموعة تكنولوجيات تأخذ البيانات فى مستودع البيانات وتحولها فى هياكل متعددة يطلق عليها Clues تسمح باستجابة أحسن للتساؤلات المعقدة.

مما تقدم يتضح أن المعالجة التحليلية على الخط مصممة فى الأساس لإدخال وتحرير البيانات فى الوقت الحقيقى. وعلى سبيل المثال، يرتبط نظام معالجة التصرف على الخط OLTP بإدخال طلب التحاق الطالب أو تسجيله، حيث ترد استمارة أو طلب الالتحاق / السجيل التى يجب أن تدخل النظام بسرعة، كما يمكن عمل التغييرات كلما أضاف الطلاب أو استبعدوا وحدات بيانات أو إلغاء طلب الالتحاق / التسجيل وحذفه... الخ.

أما فى حالة المعالجة التحليلية على الخط OLAP فإنها تبنى لكى يتساءل فيها، ولكنها غير مختصة بإدخال وتحديث واستبعاد ثابت وتخزن بياناته فى شكل شبه معيارى واكتشاف الجداول لتعظيم سرعة التساؤلات، وبذلك تهدف لتجنب بطء نظم معالجة التصرف على الخط وقواعد البيانات المنفصلة.

4/8 مستودعات البيانات الافتراضية: Virtual Data Warehousing

يوجد تشابه مشترك لوصف مستودعات البيانات من أطراف مثل: المنتجين (تطبيقات المصدر)، المواد الخام (البيانات) التى تورد لمستودعات البيانات. وفى مستودع البيانات يعاد تغليف المواد الخام التى تحول لجداول حقائق ترتبط بالمقرر الدراسى مع معلومات مسحوبة من تطبيقات المصدر المختلفة.

وبدون شك فإن هذا التشابه يصف بدقة مستودع البيانات كما هو منفذ على نطاق واسع من خلال قواعد البيانات فى الوقت الحالي، حيث إنه فيما يتصل بالنشاط المطلوب، فإن هذا التشابه يشتمل على تدفق أساسى من المواد أو المقررات الدراسية الخام التى يجب أن تتحرك من موقع التطبيق، حتى تصل بعدئذ إلى خادم الكمبيوتر بالمدرسة مثلا قبل ما يتاح فعليا للطلاب والمعلمين. ومن الملاحظ أن البيانات تكون فى الشكل الرقمى الإلكترونى، وتتحرك كمية كبيرة منها فى الواقع

من خلال التكرار، الذى لا يكون خاصية مرغوبة، إلا أنه يؤدى إلى الفشل فى الوصول للبيانات فى المكان المعين، وفشل نظام إدارة قاعدة البيانات الموزعة مصحوبا ببطء وفجوة معالجة وسعة نطاق الاتصالات غير المتجانسة.

وفى حقبة تحميل البيانات على الخط فى الوقت الحقيقى وتنفيذ شفرة التطبيق مع لغة تطوير وبرمجة للتطبيقات على الشبكة مثل لغة جافا Java Applets على سبيل المثال، التى تمثل نماذج كمبيوترية موزعة على كمبيوترات الطلاب والمعلمين، مع باقى النماذج المعمارية المتأثرة بتكنولوجيا الإنترانت والإنترنت، ويصبح من المرغوب فيه تحريك البيانات المحتاج إليها للأغراض التحليلية على أساس الحاجة المرتبطة باستخلاص وتحرير وتحميل كميات ضخمة من البيانات تقاس على أقل تقدير بالجيجابايت بدلا من توقع حاجة المعلومات المعينة للتحليل.

وفى العادة يمثل ذلك التوجه نحو مستودعات البيانات الافتراضية التى تتساوى مع نسخة تكنولوجيا نظام إدارة قاعدة البيانات الموزعة القديمة الجاهزة فقط. على سبيل المثال، يمكن أداء عملية افتراضية مشتركة مستخدمة البيانات التشغيلية المتضمنة فى نظام المعلومات الإدارية IMS والمعلومات التشغيلية المرتبطة المخزنة فى قاعدة بيانات أوراكل Oracle على سبيل المثال، مع تحويلات ملائمة تحدث أثناء وقت التشغيل لإعداد التقرير المعين. وبذلك فإن مستودع البيانات الافتراضى يتضمن كثيرًا من المشكلات الكامنة والمورثة من النظم الكمبيوترية القديمة الخاصة بنظم إدارة قاعدة البيانات الموزعة وتأثير النظم التشغيلية وصعوبة تداول اختلافات المعانى والدلالات بين البيانات فى البيئات المتعددة وهكذا. ويرتبط الحل بتوحيد وتقنين وتوازن أصول مستودعات البيانات الافتراضية الأساسية أى تداول البيانات بالقرب من المصدر كلما أمكن ذلك، مع قدرات البرمجيات الوسيطة للتحويل وتأكيد الجودة.

فى إطار المعالجة التحليلية على الخط يبدأ بتلخيص البيانات وبالضرورة استلام تفاصيل بيانات متزايدة لمساندة التحليل واتخاذ القرارات. على سبيل المثال، يبدأ أولا بالنظر فى تحديد إطار تحصيل الطلاب فى كل مقرر كل شهر أو أسبوع، مع إمكانية إضافة تفاصيل بصفة متعاقبة حتى يمكن الوصول لمستوى معين لتحصيل طلاب التخصص أو المقرر فى كل فصل من فصول المدرسة.

وعند التحميل على الإنترنت يجب تعريف ما سوف تؤديه تكنولوجيا الوسائط المتعددة حيال ذلك، حيث إنه فى إطار تشعب الوسائل Hypermedia، فإنها تشتمل على الفيديو والأشكال والنصوص بالإضافة للبيانات التقليدية التى تشتمل على نماذج لتفسير العلاقات المعقدة بين مكونات المعلومات والوصول إليها فى طرق تطبق لحاجات معينة.

وباعتبار السيناريو التالى: عند فحص معلومات تاريخ التحصيل فى مقرر معين وفقا للدرجات المحصلة، قد يلاحظ المحلل انخفاضًا دائمًا فى معدل تحصيل الطلاب فى المقرر المعين فى هذا الشهر عن الشهرين السابقين أو العامين الماضيين. وفى نفس الوقت، فإن المقررات أو المواد الدراسية الأخرى فى المدرسة يكون التحصيل فيها أحسن من المتوقع، ويرتفع التحصيل النهائى أو الإجمالى للعملية التعليمية فى كل المقررات من خلال التنقيب فى مستودع بيانات نظام التحصيل بالمدرسة، وقد لا يجد المحلل سببًا ظاهرًا جاهزًا يدعو للخروج عن القياس أو المعدل المتنبأ به مسبقا، فما هو السبب فى تلك النتيجة؟

فى هذه الحالة قد لا تتوفر الإجابة فى مستودع البيانات بصفة كلية (على الأقل لا كما نفسر تقليديا مستودع البيانات). على سبيل المثال، نشاط مدرسى آخر منافس مثل القراءة والاطلاع الحر فى المكتبة قد يكون هو المسبب فى زيادة التحصيل.

وما يعمله المحلل فيما يتصل بذلك، يرتبط بمد مفهوم ذكاء أعمال التحصيل المدرسى لكى يتضمن أنشطة أخرى مساعدة كالاطلاع الحر، التدريبات المعملية، الرحلات، الندوات التعليمية، استخدام الإنترنت... الخ، الممكن الوصول إليها

كضرورة لما يعبر عليه مصطلح ذكاء الأعمال بالضبط. وبذلك فإن وحدات معلومات صغيرة من مصادر عديدة ترتبط معا للمعالجة التحليلية التى تنتج صورة واضحة ومعبرة عن الوضعية المعينة. وقد يقدم الوصول لأنواع معلومات غير تقليدية عن الحدث المعين المرتبط بشرح نمط التحصيل المطلوب التعرف عليه تحديد التالى:

- امتلاك متخذ القرار أى المدرس معلومات موثوق بها عن أبعاد تحصيل الطلاب على مدى زمنى قد يمتد لستة أشهر سابقة.
- معلومات أساسية عن مدى تحصيل الطلاب من كل مدرس فى نفس المقرر الدراسى يؤدى لشحذ ملكات التنافس العلمى بينهم مما يسهم فى رفع مكانة المدرسة فى بيئتها.

9 - معمارية مستودعات البيانات:

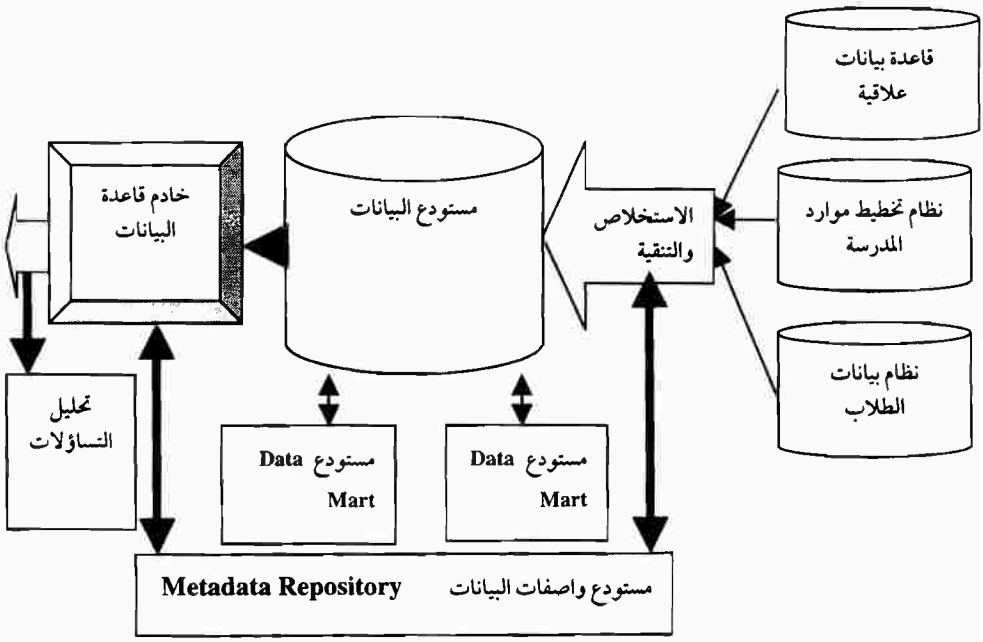
1/9 مكونات مستودع البيانات:

توجد كثير من المكونات التى تسهم فى عمل مستودع البيانات، التى تتضمن التالى:

- مصادر البيانات التى تتمثل فى نظم معالجة التصرف على الخط وهى نقطة البدء التى تنتهى فى مستودع البيانات، وقد تكون قواعد البيانات العلاقية، نظم تخطيط موارد المنشأة ERP ، تطبيق بيانات الطلاب، وتطبيق بيانات المدرسة... الخ، وقد تكون مصادر البيانات تتابعية، علاقية، تاريخية، خارجية... الخ.
- الاستخلاص والتنقية Extraction and Cleaning حيث تلتقط مستخلصات البيانات من المصدر التشغيلى وتعمل على تنقية البيانات المستخلصة، وترتبط بمستودع واصفات البيانات Metadata Repository .
- مستودع البيانات Data Warehouse الذى يتضمن وحدات تنظيمية للبيانات Data Marts وتنبع من خلال تفكيك نظام إدارة قاعدة البيانات.
- أدوات Cubes التى تمثل محاور للتساؤلات فى مستودعات البيانات.

- خادم مستودع البيانات المختص بمحرك المعالجة التحليلية على الخط، الذى يرتبط بنظام إدارة قاعدة البيانات العلاقية لغة التساؤل الهيكلية SQL التى تستمد مفردات تساؤلها من مستودع واصفات البيانات، ويطلق عليها المعالجة التحليلية على الخط ROLAP لتخزين وإدارة بيانات مستودع البيانات، أى التنظيم الموجه نحو الجداول؛ المعالجة التحليلية على الخط متعددة الأبعاد MOLAB المبنى على هياكل البيانات، والمعالجة التحليلية على الخط الهجينة HOLAP.

ويوضح الشكل التالى مكونات معمارية مستودعات البيانات السابق عرضها:



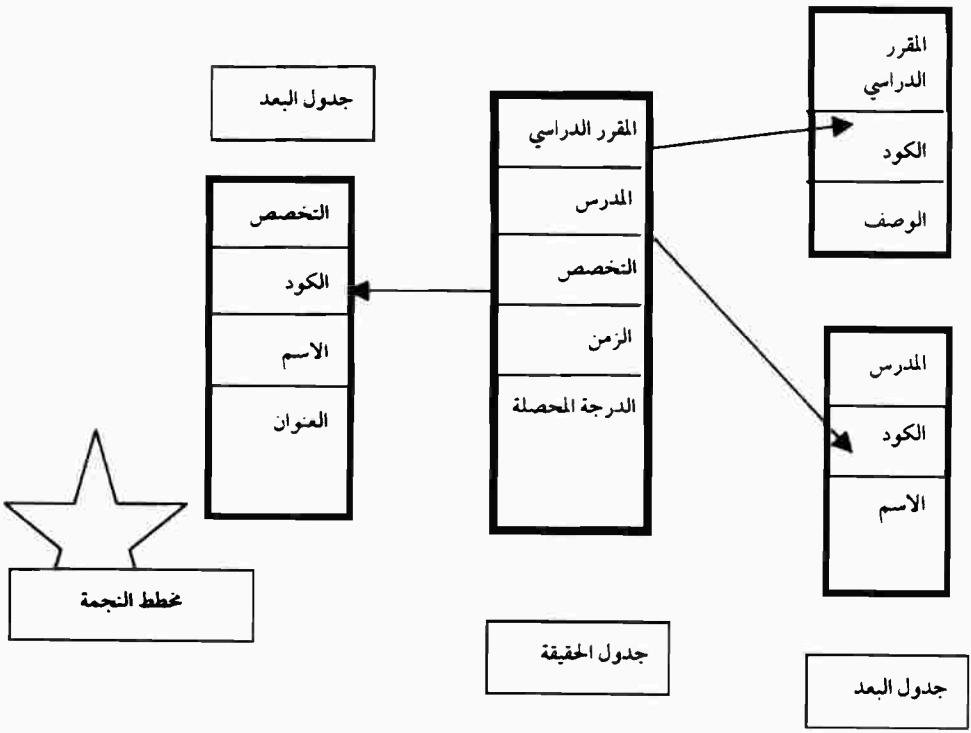
شكل رقم (1/6): مكونات معمارية مستودع البيانات

وفى نطاق مستودع البيانات تنظم المعلومات حول جدول الحقيقة FACT TABLE الذى قد يكون جدولًا واحدًا أو أكثر من جدول، ويجمع كل جدول مجموع أحداث غير المتجانسة (الحقائق) التى تتسم بواسطة أبعاد وخصائص مستقلة. على سبيل المثال، جدول التحصيل الدراسى للمقرر الدراسى:

جدول رقم (2/6): جدول الحقيقة

المقرر	المدرس	التخصص	الفترة الزمنية	نسبة التحصيل، النجاح
مقرر (أ)	مدرس (1)	تخصص 1	فصل دراسي	٪70
مقرر (ب)	مدرس (2)	تخصص 2	عام كامل	٪80

يلاحظ من الجدول السابق أن كل بُعد يمكن أن يشتمل على عدد من الخصائص، وفى هذه الحالة؛ فإن القيمة فى جدول الحقيقة هو مفتاح خارجى كما هو موضح فى الشكل التالى:



جدول رقم (2/6): جدول الحقيقة وجداول الأبعاد والقيم

مما سبق يتضح التالى بخصوص مكونات معمارية مستودع البيانات:

1- تشتمل المكونات على التالى: استخلاص البيانات من قواعد البيانات وتحميلها، وتشكيل مستودع البيانات، والتحليل والتساؤل باستخدام أدوات المعالجة التحليلية على الخط، وواصفات البيانات، وأدوات تنقيب البيانات، الخ.

2- تحميل مستودع البيانات بعد تنقية البيانات قبل تحميلها، ومن أساليب التحميل لغة التساؤل الهيكلية لإدخال بيانات جديدة، وتسجيل واجهة التفاعل، ومنفعة تحميل الدفعة.

3- مصادر البيانات سواء كانت بيانات داخلية أو خارجية، عند التفكير فى إنشاء مستودع البيانات هو ببساطة استخلاص البيانات التشغيلية وإدخالها فى مستودع البيانات؛ حيث تأتى البيانات من مصادر متفرقة ممكن التساؤل فيها.

4- جودة البيانات وسلامتها وتكاملها.

2/9 دورة حياة تطوير مستودع البيانات:

تتضمن مراحل دورة حياة تطوير مستودع البيانات على التالى:

1- **مرحلة التحليل**: تعرف الأسئلة التحليلية، حاجات البيانات، توقيت البيانات الفورى والمتدرج، إنشاء قاموس بيانات أو مستودع واصفات البيانات على مستوى المنظمة، التحليل البعدى الذى يعرف الحقائق والأبعاد والقيم.

2- **مرحلة التصميم**: التى تتضمن مخطط البيانات، تحويل البيانات، التراكم،القيمة المحسوبة مسبقا، معمارية الأجهزة والبرمجيات، التوصل للنمذجة البعدية فيما يتصل بكل من: جدول الحقائق ويمثل الجدول الأصلى فى النموذج البعدى ويشتمل على قياسات الأعمال، جدول القرار وهو أحد مجموعات الجداول المقارنة لجدول الحقيقة، وتتضمن على جداول أبعاد الخصائص المختلفة.

3- **مرحلة استيراد البيانات**: فى هذه المرحلة تعرَّف مصادر البيانات، واستخلاص البيانات المحتاج إليها من النظم المتواجدة فى مجال تحديد البيانات، وتنقية البيانات والتخلص من الأخطاء وعدم التوافق مع المجال.

4- **مرحلة تركيب الأدوات**: مثل أدوات الاستخلاص، وأدوات تنقية وتصحيح البيانات، أدوات التقرير، أدوات تنقيب البيانات، نظم المعلومات الجغرافية... الخ.

5- **مرحلة الاختبار والنشر**: التى ترتبط باختبارات إمكانية الاستخدام، وتركيبات البرمجيات، وتدريب المستخدمين، الأداء المبنى على الاستخدام... الخ.

المراجع

6. محمد محمد الهادى (2004). توجهات توظيف تكنولوجيا المعلومات والاتصالات المتقدمة فى مرافق المعلومات والمكتبات، القاهرة: المكتبة الأكاديمية [الفصل الخامس: نظم قواعد البيانات الهيكلية، ص ص: 129-163؛ الفصل السادس: قواعد بيانات النص وإدارة الوثائق، ص ص: 157-192؛ الفصل السابع: نظم قواعد بيانات الوسائل / الوسائط المتعددة، ص ص: 195-232].

7. .....، (1993). التطورات الحديثة لنظم المعلومات المبنية على الكمبيوتر، القاهرة: دار الشروق [الفصل الثالث: نظم قواعد البيانات، ص ص: 89-135].

8. .....، (1989). نظم المعلومات فى المنظمات المعاصرة، القاهرة: دار الشروق [الفصل التاسع: نظم قواعد البيانات: التخطيط والإدارة، ص ص: 281-356].

9. .....، (يوليو 1982). قواعد البيانات وشبكات المعلومات فى العلوم الاجتماعية، **مجلة المكتبات والمعلومات العربية**، س2، ع 3 ، ص ص: 29-59.

10. .....، (يناير 1981). تصميم وإدارة قواعد البيانات، **المدير العربى**، العدد 73، ص ص: 41-51.

11. Agosta, L. (2000). The essential guide to data warehousing, Englewood-Cliffs, NJ: Prentice-Hall

12. Atkinson, M. et al (1989). The object-oriented database system manifesto, In: Proceedings DOOD'89, Kyoto, Japan, December 1989.

13. Bancilhon, F. et al 91992). Building an object-oriented database system: The story O2. Morgan Publishers

14. Bischoff, J./ and Alexander, T. (1997). Data warehouse: Practical advice from the experts, Englewood-Cliffs, NJ: Prentice-Hall

15. Blakeley, J. A. (1991). DARPAOpen object-oriented database preliminary module specification: Object-query module, Dallas, TX: Texas Instruments, Inc.

16. Booch, G. (1994). Object-oriented analysis and design with applications, 2nd ed. Benjamin/Cummings Publishing Co.

17. Chen, P. P. (1997). The entity-relationship approach to logical database design, QED Publishing Co.

18. Codd, E. F. (1990). The relational model for database management, version 2, New York: Addison-Wesley Publishing Co.

19. Date, C. J. (1994). An introduction to database systems, 6th ed., New York: Addison-Wesley Publishing Co.

20. Devlin, Barry (1997). Data warehouse from architecture to implementation, New York: Longman Inc.

21. Inmon, W. H. et al (1997). Data stores, data warehouse and Zachman framework, New York: McGraw-Hill

22. Jarke, M. et al (2003). Fundamentals of data warehouse, 2nd ed. Springer

23. Kim, W. (1990). Introduction to object-oriented databases, Cambridge, MA: MIT Press

24. Kimball, Ralph (1996). The data warehouse toolkit, New York: John Wiley

25. Meyer, B. (1997). Object-oriented software construction, 2nd ed., New York: Prentice Hall

26. Simon, Alan (1995). Strategic database technology: Management for the year 2000, Morgan Kaufman

27. Stroustrup, B. (1997). The C++ Programming language, 3rd ed., New York: Addison-Wesley

28. Thomsen, Erik (1997). OLAP solutions, New York: John Wiley