Cybrarians Journal
ISSN 1687-2215
دورية إلكترونية فصلية محكمة متخصصة في مجال المكتبات والمعلومات
العدد 19، يونيو 2009

نحو رؤية جديدة لإدارة المكتبات باستخدام تقنية التنقيب عن البيانات (Data Mining)

وسام محمود أحمد درويش
باحثة بمرحلة الدكتوراه، قسم المكتبات والمعلومات
كلية الآداب – جامعة القاهرة
wes_moh@hotmail.com

الاستشهاد المرجعي

وسام محمود أحمد درويش. نحو رؤية جديدة لإدارة المكتبات باستخدام تقنية التنقيب عن البيانات (Data Mining) .- cybrarians journal .- ع 19 (يونيو 2009) .- تاريخ الاتاحة <أكتب هنا تاريخ الاطلاع على البحث> .- متاح في: <أكتب هنا رابط الصفحة الحالية>

تمهيد:

حتى يمكن لأى عمل أن يقدم المطلوب منه لابد ومن تخطيطه وإدارته بشكل سليم، فالإدارة في مجملها (علم – مهارة – موهبة – إدراك – ذكاء – خبرة – رؤية مستقبلية)، والمكتبات مثلها كأي مؤسسة كَبر أو صَغِر حجمها لابد وأن تخضع لهيكل وظيفي ينظم ويخطط الوظائف داخل المكتبة، إضافة لذلك تخضع المكتبة لهيكل وظيفي آخر يحدد موقعها داخل الكيان الأكبر الذي تندرج تحته...وهكذا .

ويعتبر أي موظف داخل المكتبة هو مسئول إدارى في إطار وظيفته وتخصصه والعمل والخدمة التي يقدمها، فمع وجود العلم والخبرة التي تتولد بالممارسة يستطيع الموظف أن يتقن عمله ويديره على خير وجه.

وإذا ذهبنا إلى نطاق أكبر على مستوى رؤساء الأقسام ومديري الإدارات والمديري العموم؛ يتطلب العمل الإداري رؤية أوسع لكل مسئول من حيث الأفراد وتنظيم العمل بينهم وتوزيعهم على المجموعات والخدمات سواء التقليدية وغيـــر التقليدية، فالتخطيط الواعي- ومعرفة ما لديٌ من إمكانات مادية وبشرية- ومقتنيات - وما يحتاجه المستفيدون الفعليون والمحتملون- وتوقعات المستقبل ، هذه هي الرؤية السليمة للإدارة.

ومن هذا المنطلق، كان لابد من التفكير في وجود آليات حديثة تساعد متخذو القرار في الحصول على رؤيات وتوقعات جديدة للمستقبل.

- **تحديات الإدارة الحديثة في المكتبات:**

توالدت الآن الكثير من العوائق والتحديات التي تقف حائل بين المسئولين في المكتبات وبين رغباتهم في تقديم أقصى استفادة ممكنة من المكتبة ، وكثير من هذه العوائق ليست بالجديدة؛ ولكن مع التقدم التكنولوجي وانفتاح الثقافات أصبحت هذه العوائق تأخذ أشكال أخرى هذا من جهة، ومن جهة أخرى أصبح أسلوب المعالجة هو الآخر مختلف وبرؤيات جديدة وفق الآليات العصرية الحديثة وتقنيات الجيل الثالث، ويذكر على سبيل المثال وليس الحصر بعض من هذه التحديات الحديثة و التي أخذت أشكال وتوجهات أخرى (14):

v **العامـــلون**، وهم العمود الفقري للمكتبة ولأي مؤسسة، ومن التحديات الحديثة التي يواجهها المسئولون فيما يتعلق بالعاملين أنه على الرغم من تخرج العديد من المتخصصين في مجال المكتبات والمعلومات في الآونة الأخيرة إلا أنه يفتقد البعض منهم القدرة على الإطلاع ومواكبة أحدث التطورات في التخصص وعلاقته بالتخصصات الاخرى ومن هـــذه التطورات على سبيل المثال تقنيات Web2 وتقنيات الجيل الثالث، وكيفية الاستفادة منها في تقديم أفضل الخدمات والعمليات والأنشطة داخل المكتبة، فالتحدي الجديد هو ليس وجود متخصصين ولكن مدى كفاءة وتطور هؤلاء المتخصصين مع تحديات العصر.

- **المقتنيـــات**، من أهم المشكلات والتحديات التي يواجهها المسئولون تجاه المقتنيات هي التعدد والتشعب في أكثر من جانب يذكر منها [أشكالها – انواعها – لغاتها – أساليب الإتاحة – حقوق ملكيتها الفكرية والمادية] هذا مما يجعل الأمر أكثر تعقيداً في الوقت الحالي من حيث الإختيار والتقييم.

- **الخدمـــات**، إن عائق تطوير الخدمات سواء الموجودة بالفعل داخل المكتبة أو الخدمات المستحدثة من الأمور التي تأخذ أكثر من جانب يأتي منها الجوانب المادية والبشرية المتاحة بالمكتبة ونوعية المستفيدين منها وأشكاليات المقتنيات بظروفها المختلفة كل ذلك وأكثر يجب أن يضعه المسئولون أمامهم فيما يتعلق بالخدمات والأنشطة، هذا من جانب ومن جانب آخر يجب استخدام أحدث الأساليب في عملية التقيم والإختيار وإعداد دراسات جيدة حول مدى أهمية هذه الخدمات وجدوتها.

ومما سبق ذكره، يتضح لنا أنه لا بد من وجود آليات جديدة تساعد هؤلاء المسئولين على الوقوف على القرارات السليمة، وقد كان للأنظمة الآلية المتكاملة دور لابأس به تجاه هذا الصدد ولكن مع وجود الكثير من التطورات على مختلف الجوانب أصبح دور هذه الأنظمة لم يفي بالغرض بعض الشئ وتطلب الأمر وجود أنظمة أكثر ذكاءاً للحصول على معلومات أكثر دقة وتحليلاً وترابطاً تفيد المسئولين ومتخدي القرار.

لماذا التنقيب عن البيانات Data Mining؟

يعد استخدام الآليات الحديثة وتطبيقاتها في مجال المكتبات والمعلومات من أهم التطورات التي تشهدها الساحة الآن سواء على المستوى الأكاديمي والمستوى العملي ..ومن المعروف أن استخدام المكتبات لعلم الإحصاء ليس بالجديد، فقد تم الاستفاة من كثير من التقنيات الإحصائية في تخصصنا من أجل قياس الآداء وكيفية التوزيع الصحيح لميزانية المكتبة على كافة الأنشطة والخدمات . وتعتمد تقنية التنقيب عن البيانات (Data Mining) موضوعنا الأساسي في هذا المقال على الأحصاء بشكل أساسي وخاصة في إعداد

الخوارزميات من أجل استخراج الأنماط المعرفية المستقبلية، ومن هذا المنطلق ليس بالغريب على تخصص المكتبات والمعلومات في استخدام هذه التقنية الحديثة .

ويمكن تعريف تقنية التنقيب عن البيانات (Data Mining) بأنها" *تقنية أو آلية لتحليل قواعد البيانات تحاول إكتشاف الأنماط والعلاقات المفيدة في مجموعة من البيانات لم تكن معروفة من قبل، ويستخدم في هذا التحليل الأساليب الإحصائية المتقدمة وأساليب الذكاء الإصطناعي والشبكات العصبية*"(2)

وعملية اكتشاف المعرفة تتكون من مجموعة خطوات تبدأ من البيانات المجمعة في شكلها الخام حتى تصل إلى شكل جديد من البيانات، وهذه الخطوات هي (2):

1- **Data Selection** (إختيار البيانات): وفي هذه المرحلة يتم إختيار البيانات التى يتم عليها تحليل من مجموعة البيانات.

2- **Data Cleaning** (تنقية البيانات): تختص هذه العملية بحذف البيانات غير المهمة والمكررة.

3- **Data Transformation** (تحويل البيانات): أى عملية توحيد/إدماج البيانات في أشكال محددة مناسبة حتى تكون مهيئة لعملية التنقيب.

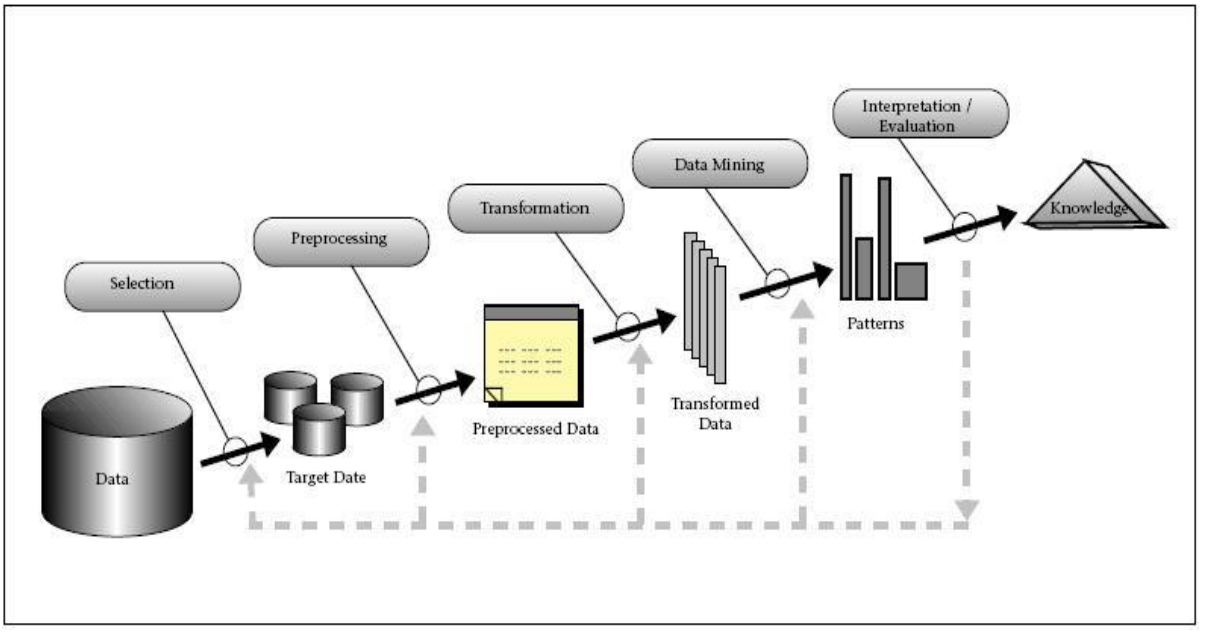
4- **Data Mining** (التنقيب عن البيانات): وتعد الخطوة الجوهرية حيث يتم استخدام التقنيات الذكية لاستنباط أنماط مفيدة من المعلومات.

5- **Pattern Evaluation** (تقييم الأنماط): أي تحديد الأنماط التى تمثل المعرفة وفق المقاييس المعطاة.

6- Knowledge Representation (تمثيل المعرفة): وهي المرحلة الأخيرة حيث يتم اكتشاف المعرفة الجديدة، وفي هذه المرحلة يتم استخدام التقنيات المصورة Visualization لمساعدة المستفيدين على فهم وتفسير النتائج المستخرجة.

ويوضح الشكل رقم (1) المراحل المختلفة لاكتشاف المعرفة، موضحاً مرحلة التنقيب عن البيانات كأحد خطوات عملية اكتشاف المعرفة، ويعد هذا الشكل من أشهر الأشكال المعبرة عن ذلك في أدب الموضوع (6)، (11)، (15) ، (20).

ومن الجدير بالذكر أن هذه التقنية دخلت مجالات أخرى كثيرة مثل التسويق والإدارة والتربية والإقتصاد والسياسة والعلوم والطب والتراث الإسلامي، واستخدمت في تقديم العديد من الأنماط المعرفية الجديدة من المعلومات والتي كشفت عن جوانب لم تكن معروفة من قبل علاوة على استخدامها في التنبؤ بما يمكن حدوثه مستقبلاً إذا تكررت أو تشابهت مثل هذه البيانات (1)، (16) ، (17).



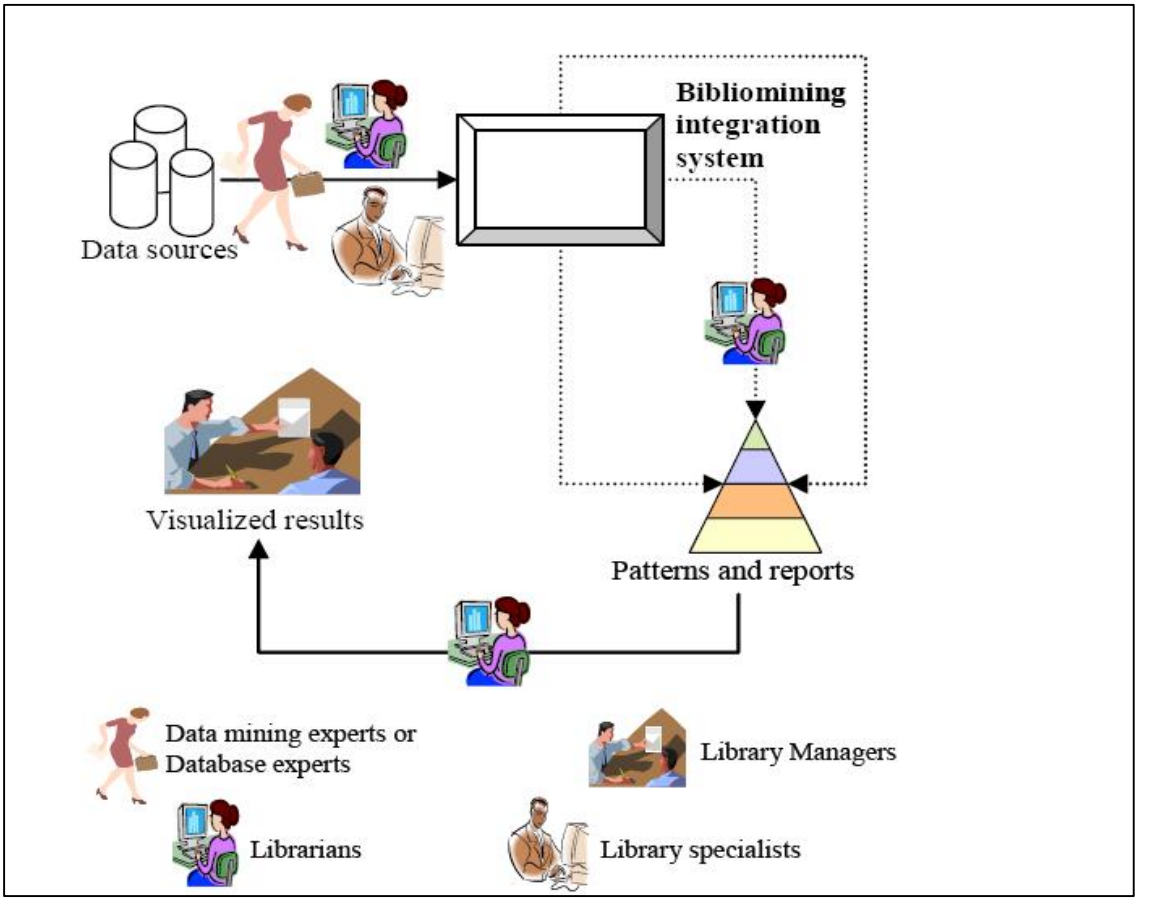
شكل رقم (1)

خطوات عملية إكتشاف المعرفة من قواعد البيانات (KDD)

المصدر: **Fayyad, Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic. From Data Mining to knowledge Discovery in Databases.- AI Magazine.-(Fall 1996).- pp:37-54.**

وفيما يتعلق باستفادة مجال المكتبات والمعلومات لهذه التقنية، فقد تم بالفعل استخدام هذه التقنية مع تخصصنا فيما يتعرف بـــData Mining in Libraries، وقد استخدم مصطلح (Data Mining) في مجال المكتبات والمعلومات للمرة الأولى عام 1998، ونظراً لما حدث من التباس عند الاسترجاع من جانب الباحثين في مجال المكتبات حيث كانت نتائج البحث تسترجع المؤلفات عن مكتبات البرمجيات المتعلقة بالتنقيب عن البيانات Software Libraries for Data Mining، وليس التنقيب عن البيانات للمكتبات Data Mining for Library Data أو Data Mining for Libraries، لذلك صك مصطلح آخر لفض هذا الالتباس عام 2003 مصطلح "التنقيب الببليوجرافي: Bibliomining" الذي يعني بتطبيق الأدوات الإحصائية وأدوات التعرف على الأنماط في كم كبير من البيانات المرتبطة بنظم المكتبات من أجل المساعدة في اتخاذ القرارات، أو تبرير الخدمات المقدمة وتطويرها خاصة في المكتبات الرقمية، وبمعنى آخر فإن التنقيب الببليوجرافي Bibliomininig عبارة عن إعادة رؤية للبيانات من منظور مختلف لتحقيق قيمة مضافة" وقد ارتبط مصطلح Mining بالعديد من المصطلحات الأخرى في مجال المكتبات مثل Web mining، Knowledge Mining، Citation Mining، Text Mining ...إلخ (2)، (3)، (4)، (5)، (10)، (12)، (14)، (19).

ويوضح الشكل رقم (2) نموذج لخطوات عملية التنقيب عن البيانات في مجال المكتبات (Bibliomining) (18).



شكل رقم (2)

خطوات عملية التنقيب عن البيانات في مجال المكتبات (Bibliomining)

المصدر: SHIEH, JIANN-CHERNG. THE INTEGRATION SYSTEM FOR LIBRARIANS' BIBLIOMINING, Asia-Pacific Conference on Library & Information Education & Practice, 2009

علاوة على ذلك قد برزت في السنوات الأخيرة الكثير من الدراسات الحديثة التى اهتمت باستخدام هذه الأداة في علم المكتبات والمعلومات و مدى أهميتها وفاعليتها، ويذكر

في ذلك الكثير من الدراسات والمشروعات التي أعدتها منظمة OCLC في هذا الصدد ويمكن الدخول على الرابط التالي للتعرف على هذه الأنشطة المختلفة :

http://www.oclc.org/research/projects/mining/default.htm

ويهتم الجزء التالي من هذه المقالة؛ بتوضيح الفوائد التي يمكن الحصول عليها من استخدام تقنية التنقيب عن البيانات (Data Mining) في تقديم أوجه جديدة تمكننا من الاستفادة منها في إدارة المكتبات ومراكز المعلومات بنظرة مختلفة لموارد وأنشطة المكتبة، وهذه المجموعة من الفوائد على سبيل المثال وليس الحصر:

1- إدارة مقتنيات المكتبة:

فمن خلال استخدام تقنية التنقيب عن البيانات يمكن للمكتبة إدارة مقتنياتها بشكل جديد في أكثر من جانب يذكر منها (5)، (12):

- **معرفة الثغرات في مقتنيات المكتبة**، حيث من خلال استخدام خوارزميات التنقيب عن البيانات يمكن الحصول على أنماط معرفية جديدة ودقيقة لم تكن معروفة من قبل؛ تبين لنا أوجه القصور في المقتنيات وأوجه الزيادة، مما يصبح أمام متخذي القرار الفرصة في تقييم مقتنياتهم في أكثر من جهة، وهذا يساعد أيضا في فتح الباب أمام المكتبة للمشاركة في المصادر مع المكتبات الأخرى لسد هذه الفجوات لديها ولمساعدة المكتبات الأخرى في مشاركتها في الجوانب الأخرى التي تم تغطيها بشكل جيد.

- **إعادة تقسيم مجموعات المكتبة وتكاملها**، يساعد ذلك متخذي القرار في الحد من تكرار المقتنيات ويحدث ذلك على مستوى الموضوع الواحد و/ أوالموضوعات الشبيه وذات الصلة؛ فمن خلال التحليل الدقيق والذكي لمقتنيات المكتبة وباستخدام الآليات المختلفة للتنقيب عن البيانات يتم استنباط أنماط معرفية (تمكن من

وجـــود علاقات تربط موضوعات ببعضهـــا البعض لـــم يكـــن واضـــح من ذي قبـــل أمـــام المسئولين ومتخـــذي القـــرار بوجود مثل هذه العلاقـــات بينهـــم)، ومن هنا تحدث التكاملية بين الموضوعات والمقتنيـــات؛ فبدلاً من شـــراء مقتنيات جديدة لموضـــوع ما يتم اســـتبدال ذلك بوضع رؤى وتقســـيمات جديدة لمقتنيات موجـــودة بالفعل من الممكـــن أن يتم الاســـتفادة منها وتقديمها للمستفيدين في هذا الموضوع.

2- قواعد بيانات خاصة بالمستفيدين:

من المعروف أنه يتوافر لدى المكتبة العديد من المعلومات التي تتعلق بالمستفيدن، وتعد دراسات سلوك المســـتفيدين تجاه مجموعات المكتبة ســـواء في البيئة التقليدية أو بيئة الشبكات الرقمية، من المعلومات القيمة لتطويرات مبتكرة في كيفية عرض وإتاحة المعلومات بالشكل الذي قد يرغبه المستفيدين، ومن هذه المعلومات الدراسات الاســـتقصاعيه، وبيانات الإعارة، ومرات الولوج الى غير ذلك من المصادر، ولا سيما إذا تم وضع كل هذه البيانات في قاعدة بيانات واحدة واستخدم فيها آليات التنقيب عن البيانات وتم ربطها مع مقتنيات وأنشطة وخدمات المكتبة ، يصبح أمام المسئولين مادة خصبة يمن من خلالها استخراج معلومات تفيدهو في كافة الأوجة بالمكتبة(**21**).

3- تنمية الموارد البشرية:

يوجد بالمكتبة موارد بشـــرية متخصصـــة وغير متخصصـــة، ويمكن تقســـيم العامليـــن غير المتخصصين لأكثـــر من تخصص، وبالتاليي يصبح هناك العديد من التخصصات والمؤهلات والدرجات العلمية داخل المكتبة الواحدة، ومع زيادة أعداد العاملين يزداد الأمر صعوبة أمام المسئولين ومتخذي القرار في الإلمام بجميع جوانبهم المختلفة. ولكن في حالة توافـــر قاعـــدة بيانـــات واحـــدة تشـــتمل على كافة بيانات العاملين بالمكتبة من حيث (أنواعهم – حالاتهم الاجتماعية – مسكنهم – تخصصاتهم – مؤهلاتهم – هواياتهم – ظروفهم الاقتصادية – خبراتهم... إلى غير ذلك من بيانات يمكن تجميعها عن العاملين) مستخدمة في ذلك آليات التنقيب عن البيانات، يمكننا استخراج علاقـــات ورابط قويـــة بين

هذه البيانات واستنباط أنماط معرفية ومعلوماتية تربط كل ذلك ببعضه البعض، وهذه المعلومات الجديدة لم يكن من الواضح اكتشافها أو الحصول عليها إلا من خلال قاعدة التنقيب عن البيانات، مما تسنح الفرصة أمام المسئولين ومتخذي القرار من اكتشاف الموارد البشرية المتاحة لديهم بصورة جديدة ومبتكرة تساعدهم في إعادة توزيعهم داخل الأقسام والأنشطة المختلفة في المكتبة ـــ ليس فقط وفق تخصصاتهم ولكن وفق الأنماط المعرفية الذكية المكتشفة ـــ مما يساعد كل من المكتبة وعامليها على حد السواء (14).

4- خدمات المعلومات المتاحة بالمكتبة:

مما سبق ذكره من حيث إعادة تقسيم المقتنيات وتكاملها وتنمية الموارد البشرية المتاحة بالمكتبة وحُسن توزيعها ودراسات سلوك المستفيدين، يصبح أمام المسئولين الرؤية الواضحة لكيفية تقييم الخدمات المقدمة ومدى جدواها وعلاوة على ذلك التخطيط الصحيح لتقديم خدمات جديدة (5)، (12).

5- اوجه صرف الميزانية:

فمن الطبيعي بعد إعادة الهيكلة الجديدة لإدارة المكتبة بمساعدة آليات هذه التقنية الحديثة، يصبح هناك بعض الترشيد في أوجه صرف الميزانية وتوفيرها لجوانب جديدة كان من الممكن عدم التفكير فيها مع قصور الميزانية لتغطيتها(5).

قائمة المراجع:

أولاً : المراجع العربية:

1- حسن مظفر الرزو . التنقيب المعلوماتي: على موارد المعرفة الإسلامية.- **أحوال المعرفة**.- س (13)، (إبريل، 2008) .

2- وسام محمود أحمد درويش .استخدام تقنية التنقيب عن البيانات (Data Mining) في تطوير المكتبات الرقمية العربية: دراسة تجريبية.- **مخطط لنيل درجة الدكتوراه**.- قسم المكتبات والوثائق والمعلومات، كلية الآداب، جامعة القاهرة. 2008.

ثانياً: المراجع الأجنبية:

3- Bovaneriee, K.- Is Data Mining Right for your Library. **Computers in Libraries**.-Vol(18) No.1 (Nov/Dec,, 1998).- pp:28-31

4-Chen, Chia Chen, Chen, An Pin. Using Data Mining Technology to Provide a Recommendation Service in the Digital Library.- **The Electronic Library**.-Vol(25) No.6 (2007).- pp:711-724

5-Cohen, Daniel J.From Babel to knowledge: data mining large digital collections.-**D-Lib Magazine**.-Vol(12), No.3 (2006)

6-Developing Digital Libraries Using Data Warehousing and Data Mining.-(Jul., 2004).- Avaliable at < ftp://ftp.inf.puc-rio.br/pub/docs/techreports/04_28_baruque.pdf > Cited in:20/5/2008

7-Fayyad, Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic.From Data Mining to knowledge Discovery in Databases.-**AI Magazine**.-(Fall 1996).- pp:37-54

8-Jansen, Bernard J.-Making Sense of Data: A Practical Guide to

Exploratory Data Analysis and Data Mining.-**Information Processing and Management**.-Vol(44), No.2 (2008).- pp:978-979

9-Kaliammal, A., Thamaraiselvi, G. Data Mining: Concepts and Techniques.- **SRELS Journal of Information Management**.-Vol(41), No.4 (Dec., 2004).- pp:339-348

10-Kim, Hyunki, Developing Semantic Digital Libraries Using Data Mining Techniques .-Thesis (Ph.D.).- University of Florida-2005.

11- Larose, Daniel T. Discovering Knowledge in Data: an Introduction to Data Mining.-West Sussex: John Wiley & Sons Inc., 2004.-222 pp

12-Lavoie, Brian, Dempsey, Lorcan, Connaway, Lynn Silipigni. Making Data Working Harder.- **Library Journal**.-Vol(131), No.10 (Jan., 2006).- pp:40-42.

13-Liu, Peng. Information Processing and Management.-**Journal of Enterprise Information Management**.-Vol(21), No.1 (2008).- pp:81-92

14-Nicholson, Scott. The Basis for Bibliomining: Frameworks for Bringing Together Usage-Based Data Mining and Bibliometrics through Data Warehousing in Digital library Services.- **Information Processing & Management**.-Vol(42), No.3 (May, 2006).- pp: 785-804

15-Raghavan, V.V, Deogun, J.S., Sever, H.-Introduction: Knowledge Discovery and Data Mining.- **Journal of the American Society for Information Science**.-Vol.49, No.5 (Apr., 1998).- pp:397-402

16-Ranian, Jayanthi, Malik, Kamna. Effective Educational Process: A Data Mining Approach.- **VINE: The Journal of Information and Knowledge Management Systems**.-Vol(37), No.4 (2007).- pp:502-515

17-Rupnik, Rok, Kukar, Matjaz. Decision support system to support decision processes with data mining..-Journal of Information and Organizational Sciences.-Vol(31), No.1 (2007).- pp:217-232

18- Shieh, Jiann Cherng, . The Integration System for Librarians' Bibliomining, Asia-Pacific Conference on Library & Information Education & Practice, 2009

19-Shieh, Jiann Cherng, Lin, Yung Shun. Bibliomining User Behaviors in the Library .- **Journal of Educational Media and Library Sciences**.- Vol(44) No.1 (Fall, 2006).- pp:35-60

20-Two Crows Corporation. Introduction to Data Mining and Knowledge Discovery.-Third ed.- Maryland: Two Crows Corporation, 2005.-36 pp.- Available at: < http://www.twocrows.com/intro-dm.pdf >

21-Whittle, Martin. Data mining of search engine logs.-**Journal of the American Society for Information Science and Technology**.-Vol(58), No.14 (2007).- pp:2382-2400