

Chapter XXI

Pattern Management: Practice and Challenges

Barbara Catania

University of Genoa, Italy

Anna Maddalena

University of Genoa, Italy

ABSTRACT

Knowledge intensive applications rely on the usage of knowledge artifacts, called patterns, to represent in a compact and semantically rich way huge quantities of heterogeneous raw data. Due to pattern characteristics of patterns, specific systems are required for pattern management in order to model, store, retrieve and manipulate patterns in an efficient and effective way. Several theoretical and industrial approaches (relying on standard proposals, metadata management and business intelligence solutions) have already been proposed for pattern management. However, no critical comparison of the existing approaches has been proposed so far. The aim of this chapter is to provide such a comparison. In particular, specific issues concerning pattern management systems, pattern models and pattern languages are discussed. Several parameters are also identified that will be used in evaluating the effectiveness of

theoretical and industrial proposals. The chapter is concluded with a discussion concerning additional issues in the context of pattern management.

INTRODUCTION

The huge quantity of heterogeneous raw data that we collect from modern, data-intensive applicational environments does not constitute knowledge by itself. A knowledge extraction process and data management techniques are often required to extract from data concise and relevant information that can be interpreted, evaluated and manipulated by human users in order to drive and specialize business decision processing. Of course, since raw data may be heterogeneous, several kinds of knowledge artifacts exist that can represent hidden knowledge. Clusters, association rules, frequent itemsets and symptom-diagnosis correlations are common

examples of such knowledge artifacts, generated by data mining applications. Equations or keyword frequencies are other examples of patterns, relevant, for example, in a multimedia context. All those knowledge artifacts are often called *patterns*. In a more concise and general way, patterns may be defined as compact and rich in semantics representation of raw data. The semantic richness of a pattern is due to the fact that it reveals new knowledge hidden in the huge quantity of data it represents. Patterns are also compact, since they represent interesting correlations among data providing, in many cases, a synthetic, high level description of some data characteristics. Patterns are therefore the knowledge units at the basis of any knowledge intensive application

Due to their specific characteristics, ad hoc systems are required for pattern management in order to model, store, retrieve, analyze and manipulate patterns in an efficient and effective way.

Many academic groups and industrial consortiums have devoted significant efforts towards solving this problem. Moreover, since patterns may be seen as a special type of metadata, pattern management has also some aspects in common with metadata management.

In general, scientific community efforts mainly deal with the definition of a pattern management framework providing a full support for heterogeneous pattern generation and management, thus providing back-end technologies for pattern management applications. Examples of these approaches are the 3W model (Johnson et al., 2000), the inductive databases approach — investigated in particular in the CINQ project (CINQ, 2001) and the PANDA framework (PANDA, 2001; Catania et al., 2004). In the context of inductive databases, several languages have also been proposed supporting the mining process over relational (or object-relational) data by extending the expressive power of existing data query languages with primitives supporting the mining process. Examples of such approaches are MSQL (Imielinski & Virmani, 1999), Mine-Rule (Meo et

al., 1998), DMQL (Han et al., 1996) and ODMQL (Elfeky et al., 2001). On the other hand, industrial proposals mainly deal with standard representation purposes for patterns resulting from data mining and data warehousing processes, in order to support their exchange between different architectures. Thus, they mainly provide the right front end for pattern management applications. Examples of such approaches are: the Predictive Model Markup Language (PMML, 2003), the common warehouse metamodel (CWM, 2001) and the Java Data Mining API (JDM, 2003).

In general, existing proposals can be classified according to the following aspects:

- (a) The chosen architecture to manage patterns together with data.
- (b) The pattern characteristics supported by the data model.
- (c) The type of operations and queries supported by the proposed languages.

As far as we know, even if several proposals exist, no critical comparison of the existing approaches has been proposed so far. We believe that such a comparison would be very useful in order to determine whether the existing approaches are sufficient to cover all pattern requirements and to guide application developers in the choice of the best solution in developing knowledge discovery applications.

The aim of this chapter is to provide such a comparison. We first present a definition of patterns and pattern management. Then, specific issues concerning pattern management systems, pattern models and pattern languages are discussed, pointing out possible alternative solutions in the context of a given scenario. Several parameters relevant to the pattern management context will also be identified and then used to evaluate the effectiveness of various theoretical and industrial proposals. Moreover, relationships between pattern and general metadata management will also be identified, and some existing approaches for

26 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/pattern-management-practice-challenges/28678

Related Content

Eight Key Elements of Successful Self-Funding E-Learning Programs

Yair Levy and Michelle M. Ramim (2005). *Encyclopedia of Information Science and Technology, First Edition* (pp. 995-1000).

www.irma-international.org/chapter/eight-key-elements-successful-self/14375

Factors Influencing Digital Innovation Strategy in Organizations: An Empirical Analysis

Sheryar Khan Tahirkheli and Olusegun Ademolu Ajigini (2022). *Information Resources Management Journal* (pp. 1-23).

www.irma-international.org/article/factors-influencing-digital-innovation-strategy-in-organizations/287900

Application of Cognitive Map in Knowledge Management

Ali Reza Montazemi and Akbar Esfahanipour (2009). *Encyclopedia of Information Science and Technology, Second Edition* (pp. 169-176).

www.irma-international.org/chapter/application-cognitive-map-knowledge-management/13568

Information Technology Investment Evaluation and Measurement Methodology: A Case Study and Action Research of the Dimensions and Measures of IT-Business-Value in Financial Institutions

Johan Nel (2008). *Information Communication Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 3021-3035).

www.irma-international.org/chapter/information-technology-investment-evaluation-measurement/22861

Modelling the Impact of Administrative Access Controls on Technical Access Control Measures

Winfred Yaokumah (2017). *Information Resources Management Journal* (pp. 53-70).

www.irma-international.org/article/modelling-the-impact-of-administrative-access-controls-on-technical-access-control-measures/186888