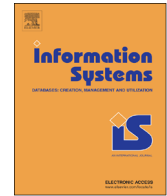




ELSEVIER

Contents lists available at ScienceDirect

Information Systems

journal homepage: www.elsevier.com/locate/infosys

The usage of best practices and procedures in the database community

Victoria Holt, Magnus Ramage, Karen Kear, Nick Heap

The Open University, United Kingdom

ARTICLE INFO

Article history:

Received 29 September 2014
Received in revised form
6 December 2014
Accepted 8 December 2014
Recommended by D. Shasha

Keywords:

Database administration
Database management
Data management
Database lifecycle
Best practice and procedures
Database operations

ABSTRACT

Database management has an important role to play in the management of data assets which are at the heart of every organization. In a fast moving technological era, where data is rapidly expanding, understanding the current best practices and procedures is important for continuous improvement. This paper investigates how databases are actually administered and identifies what practices and procedures are utilized throughout the database lifecycle. The paper highlights the demographics of people who manage database systems and the diverse requirements of database systems given the wide range of software and hardware available. The results of this paper show the breadth of issues relevant to database management. The paper concludes by showing where existing practice and procedures are not optimal, and by highlighting the complexities in the field.

© 2014 Published by Elsevier Ltd.

1. Introduction

Database management is, at its heart, the administrative tasks associated with the storage, modification and retrieval of data held within a database management system (DBMS). Organizations today require impeccable database management in order to maintain a high quality of data, and for that data to be secure and available whenever it is required. The data for governments, banks, financial institutions etc. must also satisfy statutory legal requirements.

It has been well publicized that problems exist with the setting up and successful operation of important databases. An early report by De Blasis [8] highlighted problems with administration, organizational issues, new technology introduction, control and technical configuration. There are many reasons for these problems. The use of appropriate practices and procedures can have a significant impact on the availability, recoverability and quality of data used in the operations

of businesses. The diversity of an organization's domains and strategies can lead to a variety of practices and procedures.

Certain practices used by organizations can be considered "best practice". Best practices are frequently described as those which are recommended for carrying out actions for desirable outcomes (see Fig. 9). Best practices drive operational excellence and effectiveness [13]. Other key terms used here are: processes – "a series of actions or steps taken in order to achieve a particular end"; procedures – "an established or official way of doing something"; and methodology – "a system of methods used in a particular area of activity" (all definitions from Oxford Dictionaries [32]).

This paper presents the findings from a comprehensive survey which investigated to what extent best practices and procedures are utilized by the database community.

2. Background to the research

Database management has evolved over the last five decades, since the first functioning prototype DBMS [21], to become an integral part of most organizations' business. The majority of global organizations today cannot operate without a functioning database. Organizations increasingly

E-mail addresses: holt.victoria@gmail.com (V. Holt),
Magnus.Ramage@open.ac.uk (M. Ramage),
Karen.Kear@open.ac.uk (K. Kear), nick.heap@open.ac.uk (N. Heap).

<http://dx.doi.org/10.1016/j.is.2014.12.004>

0306-4379/© 2014 Published by Elsevier Ltd.

1 realize the value of the data that they hold and are
 3 beginning to draw more benefit from its analysis and
 5 mining [12]. Some examples of usage include shopping
 7 history to predict purchases [11], social media to predict
 9 trends [33], live mapping for disaster aid [16], the storage
 11 of the human genome to aid medical research [7], sensor
 13 data from CERN [34] and other data intensive scientific
 15 discoveries [22,23].

17 Changes in culture and everyday life have brought
 19 about the sharing of more information, and this has
 21 radically changed the usage of databases. New technolo-
 23 gies such as cloud and virtualization enable a different
 25 operating model, one that allows organizations to share
 27 resources, but these new models add to the complexity of
 29 an already complicated activity.

31 The trend of rapidly increasing data volumes is being
 33 driven not only by the requirements of government and
 35 business to store more information, but also by the digitiza-
 37 tion of film and TV and the use of social media. Gantz and
 39 Reinsel [15] have estimated that the volume of all known
 41 data will have grown from 0.8 zettabytes (10^{21} bytes) in 2009
 43 to 35 zettabytes in 2020. This is a mixture of both structured
 45 and unstructured data. Unstructured data, objects that have
 47 little or no identifiable structure, e.g. text, images, audio and
 49 video, were not previously considered within the database
 51 community. However this view has now changed within the
 53 industry as a whole. The Lowell Report [2], a summary of a
 55 gathering of academic database researchers' discussions on
 57 the state of database research, states that "Database needs
 59 are changing, driven by the internet and increasing amounts
 61 of scientific and sensor data" [2, p. 111].

The administration and management of this complex
 area would benefit from a better understanding of the extent
 to which best practices and procedures are utilized by the
 database community. However, that raises the question as to
 whether the adoption of best practice is constrained by the
 many interactions between different the interconnected
 aspects of the management of database systems.

The database management system (DBMS) itself is
 constructed of many components which can be considered
 to form a layered technical system. Typically, disparate
 organizational teams manage each layer independently.
 These teams have different sets of goals, together with a
 variety of approaches, and problems in operation which
 can occur due to the interconnections. This has increased
 the challenges of overall management.

Many organizations demand low cost infrastructure
 without jeopardizing functionality or operational ability.
 Database architecture, design and development are the
 foundation of any well designed DBMS. The database is
 continually evolving and adapting to the demands of the
 users, organizations and the global environment.

The management of the database and the data contained
 within it are often undertaken by different teams. Two
 separate functions were identified by Kahn [25, p. 794] as
 database administration and data administration. However,
 they have many interconnected components. Organizations'
 business requirements for data collection and manipulation
 appear to be driving the requirements for database man-
 agement. The fields are gradually merging, and Mullins [31]
 proposed data administration practices and procedures to

63 address this, arguing that "when database administration is
 65 treated as a management discipline, the treatment of data
 67 within your organization will improve" [31, p. 9]. A manage-
 69 ment discipline is required to manage the successful
 71 operation of the databases. The management methods used
 73 for development and infrastructure projects has also chang-
 75 ed. Historically, development projects have been domi-
 77 nated by Waterfall methodologies, but there has been some
 79 shift towards Agile methodologies to overcome Waterfall's
 81 inflexibility in relation to rapidly changing business require-
 83 ments [37].

85 Changes to the database engine, structure and hard-
 87 ware all require practices and procedures to ensure that
 89 data is protected. The Independent Oracle Users Group
 91 (IOUG) has recently completed several surveys that aid in
 93 understanding the complexities of database systems. The
 95 IOUG report [28] entitled "Managing the Rapid Rise in
 97 Database Growth" identified the importance of database
 99 change management practices.

101 There are many current complexities related to data
 103 management and data administration, and as Aiken et al.
 105 [5] suggest, data management is still evolving. The Clar-
 107 emont report on database research [3, p. 65] highlighted
 109 concerns that are important to the community regarding
 111 the increasing technical scope, processes and keeping track
 113 of the field that is important to the community. Other
 115 surveys previously undertaken provided some insight, and
 117 highlighted the rise of database administration, with an
 119 unclear direction of the future path [5,17–19,27,30].

121 The survey reported here highlights the real world
 123 situation of the database community at present, and shares
 the current practices and procedures of the respondents.

3. The survey

125 The primary goal of the research was to investigate
 127 how databases were actually administered and to identify
 129 what practices and procedures were utilized throughout
 131 the database lifecycle. Further aims were to understand
 133 the demographics of people who manage database sys-
 135 tems, and to investigate how they learned about best
 137 practice and whether any IT frameworks were used. As
 139 the database community is dispersed globally, the survey
 141 has sought to reflect this global nature. The anonymized
 143 data will appear on www.sqltoolkit.co.uk.

3.1. Sampling

145 Non-probability convenience sampling was selected to
 147 allow all those who were willing and able to participate
 149 in the survey. It is impossible to know the size and disper-
 151 sion of the database-management and data-professionals
 153 population. Convenience sampling [9,10,14] is commonly
 155 used during preliminary research to gain a summary of
 157 interesting information. It allows the data to be collected
 159 quickly and inexpensively.

161 The sample was obtained through advertising the
 163 survey via social media such as Twitter, LinkedIn and
 165 Facebook groups, email newsletters and blog posts. A
 167 strategic decision was taken to gather data as widely as
 169 possible across the database population, to include a range
 171

Download English Version:

<https://daneshyari.com/en/article/6858708>

Download Persian Version:

<https://daneshyari.com/article/6858708>

[Daneshyari.com](https://daneshyari.com)