

## **PENGARUH KUALITAS DATA TERHADAP MANFAAT YANG DIPEROLEH DARI IMPLEMENTASI DATA WAREHOUSE**

Munawar

Teknik Informatika, Universitas Esa Unggul  
Jl. Arjuna Utara No. 9 Kebon Jeruk, Jakarta 11510  
an\_moenawar@yahoo.com

### **Abstrak**

Given that the literature indicates that many benefits can be obtained from a DW (data warehouse), but that the threat of failure is also high, therefore more understanding of the quality dimensions that contribute to the success of a DW is required. This study address that need through an empirical investigation of success in data warehousing as measured by DQ (data quality) dimensions involvement in every stage of DW development. This study attempts to confirm DW benefits from reviewed literature in five case studies and three consultants. This study also identifies relationship amongst DQ dimensions in DW development stages and DW success. This study is helpful for DW practitioners, implementers and researchers for understanding of the challenges which DQ dimensions in DW stages are significant to establishing a successful DW development.

**Keywords:** *data warehouse, data quality, data warehouse benefits.*

### **Abstrak**

Dari literatur diketahui bahwa banyak manfaat yang bisa diperoleh dari sebuah data warehouse (DW), namun ancaman atas kegagalan dari implementasi DW juga tinggi. Karena pemahaman lebih jauh tentang dimensi kualitas yang bisa memberikan kontribusi kesuksesan DW sangat diperlukan. Penelitian ini menjawab kebutuhan tersebut melalui investigasi empiris tentang kesuksesan DW yang diukur dari keterlibatan dimensi kualitas data (*data quality* - DQ) di setiap tahapan pembangunan DW. Penelitian ini mengkonfirmasi manfaat DW dari literatur ke dalam 5 studi kasus dan tiga konsultan. Identifikasi atas hubungan antara dimensi kualitas data di semua tahapan pembangunan DW dan tingkat kesuksesan DW juga berhasil dilakukan. Penelitian ini sangat berguna buat para praktisi, implementor dan peneliti untuk memahami tantangan signifikan dimensi kualitas data terhadap kesuksesan pembangunan DW.

**Kata Kunci:** data warehouse, kualitas data, manfaat datawarehouse

## **Pendahuluan**

Saat ini, semakin banyak organisasi yang membangun DW. Hal ini ditunjukkan dengan angka kenaikan pertahunnya 43 % (Winter and Strauch, 2003) dan akan selalu naik setiap tahunnya karena biaya *storage* yang semakin turun dari dari tahun ke tahunnya (Agosta, 2004). Namun sayang, dengan cepatnya pertumbuhan volume data telah menimbulkan problem baru bagi institusi yaitu kualitas data/ DQ (Ranjit and Kawaljeet, 2010), dimana hal ini adalah isu kritis ketika data ditransfer dari satu sistem ke sistem yang lain seperti sifat dasar DW.

DW adalah proyek yang sangat menantang, karena banyaknya masalah teknis dan organisasi yang harus dipertimbangkan. Tujuan dasar pembangunan DW adalah untuk meningkatkan kualitas informasi di organisasi dalam rangka pengambilan keputusan (Watson and Haley, 1997). Manfaat terbesar yang bisa diperoleh dari DW adalah untuk memperbaiki proses bisnis, mendukung pengambilan keputusan strategik, mempersiapkan informasi yang baik dan menghemat waktu bagi pengguna (Watson et al, 2002)

Kualitas data yang baik akan meningkatkan kepercayaan pengguna terhadap DW dan mengoptimalkan manfaat bisnis yang bisa diperoleh dari DW (Kumar, and Thareja, 2013). Banyak peneliti menyatakan bahwa kesuksesan DW bisa diukur dari *net benefit* yang diperoleh dari DW (Hwang and Xu, 2008). Akan tetapi deteksi cacat dan peningkatan kualitas data terkait dengan biaya. Jika target level kualitas terlalu tinggi akan menegasikan manfaat yang bisa diperoleh. Penelitian ini bertujuan untuk memberikan bukti empiris tentang manfaat yang bisa diperoleh dari DW sehingga diharapkan bisa memberikan sumbangsih pada penelitian tentang kesuksesan DW

terkait dengan manfaat DW. Lebih jauh, penelitian ini menyelidiki korelasi antara dimensi kualitas data dengan manfaat DW.

Menurut Rudra dan Yeo (2000) penelitian tentang kesuksesan DW masih langka. Sepanjang peneliti ketahui, hingga saat ini belum ada bukti empiris tentang pembangunan DW di Indonesia. Karenanya, penelitian ini diharapkan bisa memberikan sumbangsih terhadap pembangunan DW dan manfaatnya khususnya di Indonesia.

Disamping perkembangannya yang pesat, DW juga memiliki potensi kegagalan yang tinggi. Lebih dari 50% proyek DW mengalami kegagalan karena rendahnya kualitas data (Friedman, 2004). Rendahnya kualitas data mengakibatkan dampak negatif ke efisiensi organisasi seperti mengurangi keuntungan perusahaan 8-12 % (Haug et al, 2011). Penelitian ini lebih menekankan ke penelaahan korelasi antara dimensi kualitas di semua fase pembangunan DW terhadap kesuksesan DW ditinjau dari sisi manfaat yang bisa diperoleh.

Saat ini makin banyak organisasi yang membangun DW, dengan tingkat pertumbuhan 43% per tahun (Winter and Strauch, 2003) dan akan terus naik di masa depan karena penurunan biaya yang tajam di media penyimpanan/ *storage* (Agosta, 2004). Namun kenyataannya hingga saat ini belum ada penelitian yang mengungkapkan kualitas DW atau kesuksesan DW (Amornbuth, 2015), karena sulit untuk mengukurnya (Bilal Ali, 2014; Hwang and Xu, 2008). Meski demikian, banyak peneliti yang menyatakan bahwa kesuksesan DW bisa diukur melalui manfaat yang bisa diperoleh dari DW (Hwang and Xu, 2008; Mukherjee, 2003; Wixom and Watson, 2001) melalui pertimbangan kualitas data di saat pembangunan DW (Kumar, and Thareja, 2013).

Isu kualitas data dapat terjadi dimanapun tahapan pada fase pembangunan DW (Kumar dan Thareja, 2013). Data sangat dipengaruhi oleh banyak proses yang harus dilalui hingga menuju ke DW. Keseluruhan proses ini akan mempengaruhi kualitasnya. Dari penelitian sebelumnya sudah bisa diidentifikasi dimensi kualitas data apa saja yang harus diakomodir di semua fase pembangunan DW (Munawar, 2016) sebagaimana bisa dilihat pada Tabel 1.

Tabel 1  
Kualitas Data yang Harus Dipertimbangkan di  
Setiap Fase Pembangunan Data Warehouse (Munawar, 2016)

		Data Warehouse Development Phase											
		Requirements					Conceptual	Logical		ETL	Physical		
		User Driven	Goal Driven	External Driven	Data-Driven	Process-driven	Multi Dimensional	Star Schema	OLAP	ETL	Physical Design		
Information Quality	Content	Soundness	Comprehensiveness	✓	✓		✓	✓	✓				✓
			Accuracy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
			Clarity	✓	✓				✓	✓	✓		
			Applicability	✓	✓	✓	✓		✓				
		Relevance	Conciseness	✓	✓								
			Consistency	✓	✓	✓	✓	✓	✓	✓	✓	✓	
			Correctness	✓	✓		✓					✓	
			Currency	✓	✓		✓					✓	
	Access	Process	Convenience	✓	✓			✓	✓	✓	✓	✓	
				Timeliness	✓	✓		✓	✓				✓
				Traceability	✓	✓		✓				✓	✓
				Interactivity	✓	✓							
Infrastructure		Accessibility	✓	✓		✓						✓	
			Security	✓	✓			✓		✓	✓	✓	
			Maintainability		✓		✓			✓	✓	✓	
			Speed	✓	✓			✓		✓			

Setelah DW berhasil dibangun, seharusnya banyak manfaat yang bisa dipetik baik untuk bisnis maupun departemen IT. Banyak penelitian yang mengungkapkan manfaat yang bisa diperoleh dari DW baik yang berwujud maupun yang tidak berwujud sebagaimana ditampilkan pada Tabel 2.

Tabel 2  
Manfaat Bisnis dari DW

Manfaat Bisnis	Sumber
Untuk mengintegrasikan data yang heterogen (internal/external) <i>Heterogeneous data integration</i>	MajidZaman et al, 2012; Torlone, 2008; Watson and Haley, 1998; Watson et al, 2002; Alshawi et al, 2003; Griffin, 1998
Hemat waktu ( <i>Time savings</i> )	Agrawal, 2008; Watson and Haley, 1998; Watson et al, 2002
Informasi yang disajikan bisa lebih banyak dan lebih baik <i>(More and better-quality information)</i>	Rahman et al, 2012; Watson et al, 2002; Alshawi et al, 2003; Watson and Haley, 1998
Keputusan yang lebih baik ( <i>better decisions</i> )	Joseph, 2013; Watson and Haley, 1997; Watson et al, 2002; Griffin, 1998
Perbaikan/ desain ulang proses bisnis. <i>(Improved/re-designed business processes)</i>	Shahzad, 2012; Watson and Haley, 1997; Watson et al, 2002; Watson and Haley, 1998
Mendukung pencapaian tujuan strategik organisasi <i>(Support for accomplishing strategic business objectives)</i>	Al-Debei, 2011; Watson et al, 2002
Satu-satunya sumber yang bisa dipercaya <i>(single source of truth)</i>	Pandey and Mishra, 2014; Watson and Haley, 1997; Watson and Haley, 1998

Penelitian ini dimaksudkan untuk mengetahui korelasi antara dimensi kualitas data di semua fase pembangunan DW dengan kesuksesan DW sebagaimana diindikasikan pada manfaat yang bisa diperoleh.

Penelitian ini untuk menguji manfaat DW yang diperoleh dari literatur ke dalam studi kasus (5 organisasi) dan beberapa konsultan (3 orang). Penelitian ini juga untuk mengidentifikasi hubungan antara dimensi kualitas di semua fase pembangunan DW dengan manfaat yang bisa diperoleh dari DW.

Hasil penelitian ini bisa dimanfaatkan oleh praktisi DW, implementor DW serta peneliti DW guna memahami betapa pentingnya pengintegrasian dimensi kualitas data ke semua fase pembangunan DW agar DW sukses diimplementasikan.

Penelitian ini meliputi beberapa tahapan sebagaimana dijelaskan pada Tabel 3.

Tabel 3  
Tahapan Penelitian

Tahapan	Tujuan
1. Studi literatur	<ol style="list-style-type: none"> <li>Mengetahui aspek kualitas data (DQ) di semua fase pembangunan DW</li> <li>Memahami manfaat DW yang bisa diperoleh</li> <li>Memahami konsep korelasi untuk mengukur hubungan antara dimensi DQ di semua fase pembangunan DW dengan manfaat yang diperoleh dari DW.</li> </ol>
2. Interview dan pengumpulan data	<ol style="list-style-type: none"> <li><i>Exploratory study</i> di organisasi yang sudah mengimplementasikan DW untuk mengetahui manfaat apa yang sudah diperoleh dari DW yang sudah diimplementasikan.</li> <li><i>Expert judgement</i> oleh konsultan DW yang memiliki pengalaman dalam pembangunan DW minimal 3 siklus di sektor industri yang berbeda-beda</li> </ol>
3. Analisis data	<ol style="list-style-type: none"> <li><i>Descriptive qualitative</i> digunakan untuk mengetahui manfaat diimplementasikannya DW.</li> <li>Korelasi spearman digunakan untuk mengukur hubungan antara dimensi kualitas data di semua fase pembangunan DW dengan manfaat yang bisa diperoleh dari DW. Ada 3 indikasi korelasi diantara 2 obyek yaitu korelasi positif (<math>r_s &gt; 0</math>), korelasi negatif (<math>r_s &lt; 0</math>) dan tidak memiliki hubungan (<math>r_s = 0</math>). Khusus untuk korelasi negatif, hal ini menunjukkan hubungan terbalik diantara kedua obyek tersebut. Maksudnya obyek yang satu akan berakibat sebaliknya kepada obyek yang kedua.</li> <li>Pada kasus dimana para responden sepakat untuk setuju atau tidak setuju (tidak ada perbedaan dalam ranking) korelasi spearman tidak bisa digunakan untuk menilai hubungan antar dua obyek. Oleh karena itu validasi dari para responden sangat dibutuhkan untuk mengetahui hubungan antara kedua obyek tersebut. Konsultasi dengan pakar juga dilakukan untuk memvalidasi temuan tersebut</li> </ol>

## Pengumpulan Data

Mengingat besarnya biaya dan lamanya waktu yang dibutuhkan untuk pembangunan DW, bisa dipastikan hanya perusahaan besar sajalah yang mampu membangun DW. Dari sekian banyak perusahaan yang membangun DW di Indonesia, tidak semuanya bisa berpartisipasi dalam penelitian ini. Hal ini disamping karena kesibukan para pengambil keputusan (sehingga tidak memungkinkan untuk berpartisipasi) juga karena sifat DW yang memuat informasi yang sifatnya konfidensial sehingga tidak semua perusahaan bersedia untuk mengungkapkan hal-hal yang terkait dengan DW yang mereka bangun. Oleh karena itu, multi studi kasus (multi case study) adalah pendekatan yang paling tepat untuk pengumpulan data.

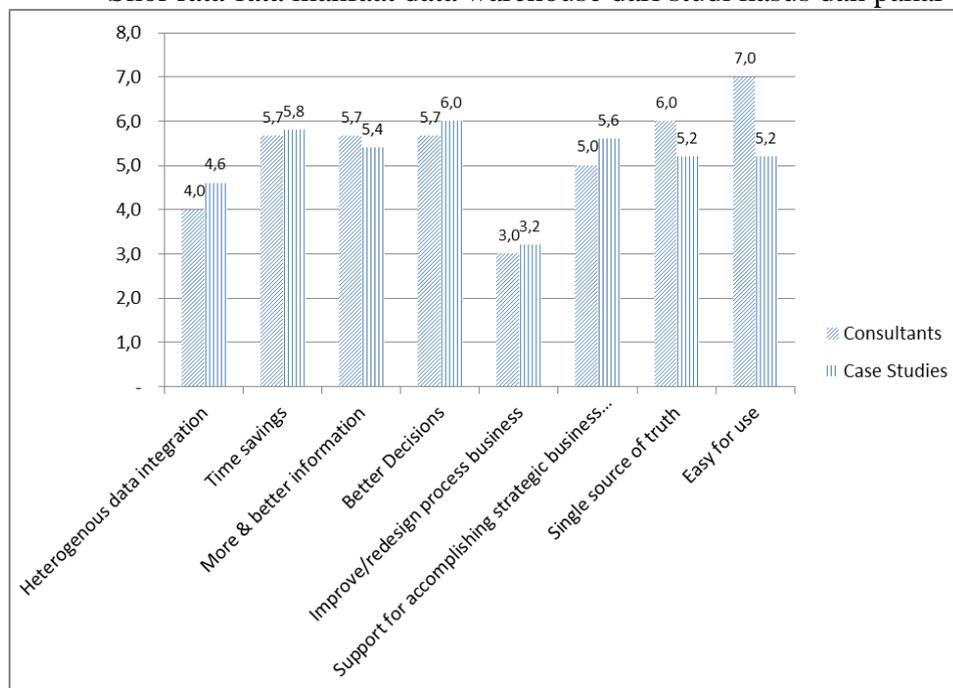
Beberapa studi kasus akan bisa meminimalisir kelemahan satu studi kasus dalam mengeneralisir hasil temuan (Tellis, 1997). Beberapa studi kasus diharapkan akan bisa membantu mengetahui konsistensi dari beberapa kasus yang diteliti (Stake, 1995).

Setelah data dari beberapa studi kasus didapatkan, selanjutnya dilakukan konfirmasi hasil tersebut oleh beberapa pakar DW (konsultan DW yang berpengalaman lebih dari 7 tahun dengan kasus pembangunan DW yang sudah pernah dilakukan lebih dari 3 organisasi yang berbeda sektor industrinya).

Teknik penggalian informasi dilakukan dengan cara interview mendalam (*an in-depth interview*) untuk mendapatkan pemahaman yang teliti dan mendalam tentang *mengapa* dan *bagaimana* suatu kejadian bisa terjadi. Dari sini selanjutnya hipotesis bisa dibuat dan di test ke beberapa studi kasus (Flyvberg, 2006).

Umumnya pertanyaan dimulai dengan pertanyaan terbuka kepada responden. Jawaban mereka akan dicatat dan selanjutnya diteruskan dengan kuesioner dengan menggunakan skor untuk mendapatkan jawaban kuantitatif sebagai manifestasi jawaban deskriptif yang sudah diberikan tentang situasi yang mereka hadapi. Skor ini digunakan untuk menyatakan tingkat setuju/ tidak setuju terhadap elemen penelitian dan juga tingkat kepuasan/ ketidakpuasan dari responden.

Gambar 1  
Skor rata-rata manfaat data warehouse dari studi kasus dan pakar



Interview dibagi menjadi dua group yang mewakili stakeholder yang berbeda yaitu : tim DW dan pengguna bisnis (*business users*). Tim DW mewakili stakeholder yang terkait dengan masalah teknik pembangunan DW, sedangkan pengguna bisnis mewakili pengguna yang menguasai proses bisnis dan tujuan organisasi yang terkait dengan DW. Pertanyaan interview dibedakan antara tim DW dengan pengguna bisnis. Selanjutnya setiap responden diminta untuk mengisi kuesioner

dengan menggunakan 7 skala untuk menunjukkan tingkat setuju/ tidak setuju terhadap elemen penelitian dan juga tingkat kepuasan/ ketidakpuasan dari responden. Coleho and Esteves (2007) menyatakan bahwa skala lebih dari 5 secara umum akan memberikan validitas hasil dan kemudahan penjelasan yang lebih baik dibanding yang menggunakan skala 5.

Selanjutnya data dari hasil kuesioner ini ditampilkan dalam bentuk tabel untuk mendukung temuan yang didapatkan dari hasil interview. Semua temuan dari interview dan kuesioner didiskusikan dan dianalisis secara mendalam dengan cara di review semua responden di organisasi yang berperan dalam penelitian ini dan juga semua konsultan yang terlibat untuk pengecekan kelengkapan dan keakuratannya.

### **Hasil dan Pembahasan**

DW mengintegrasikan berbagai data dari berbagai sumber di semua lini organisasi. Karenanya informasi yang dihasilkan DW perlu berorientasi bisnis sehingga bisa memenuhi semua kebutuhan pengambil keputusan. Dengan banyaknya informasi yang disimpan di DW, kepastian kualitas data menjadi faktor kunci dalam pembangunan DW.

DW yang berkualitas bisa membawa kesuksesan dalam implemantasinya dengan cara memperoleh manfaatnya. Untuk mendapatkan hal tersebut, kualitas data harus dipastikan di setiap tahapnya yaitu: analisis kebutuhan, pemodelan konseptual, logikal, fisik dan ETL.

### **Manfaat Data Warehouse**

Manfaat DW diantara 5 organisasi dan justifikasi oleh 3 pakar hasilnya bervariasi. Berdasarkan hasil interview dengan skala 7 mengungkapkan manfaat yang bisa didapatkan dari DW (Gambar 1 dan Tabel 4). Ada beberapa perbedaan persepsi antara pengguna bisnis dengan pakar dalam menginterpretasikan *ease of use and single source of truth*. Semua pakar sepakat tentang keharusan *ease of use* (kemudahan penggunaan) dalam pembangunan DW (skala 7), akan tetapi pengguna bisnis hanya memberikan nilai 5. Dengan DW yang ramah (*user friendly*) memungkinkan pengguna untuk membuat sendiri laporan yang mereka butuhkan tanpa tergantung kepada orang IT. Namun bagi organisasi pemerintahan, kemudahan penggunaan bukan tujuan utama pembangunan DW. Pengambil keputusan di pemerintahan bukanlah pengguna DW (staf merekalah pengguna DW yang sebenarnya). Mereka hanyalah konsumen dari laporan yang dihasilkan oleh DW. Semua pakar juga sepakat memberikan nilai 6 untuk *single source of truth*, sementara pengguna bisnis hanya memberikan nilai 5, karena banyaknya sistem informasi yang digunakan oleh organisasi-organisasi tersebut. Jadi informasi yang ada di DW bukan satu-satunya sumber untuk mendapatkan informasi.

Tabel 4  
Manfaat Data Warehouse pada Studi Kasus dan Pakar

Manfaat Data Warehouse	Pakar		Studi Kasus	
	Rata	Var	Rata	Var
Heterogeneous data integration	4,0	3,0	4,6	2,3
Time savings	5,7	2,3	5,8	1,2
More & better information	5,7	1,3	5,4	0,8
Better Decisions	5,7	1,3	6,0	1
Improve/redesign process business	3,0	12,0	3,2	3,2
Support for accomplishing strategic business objectives	5,0	3,0	5,6	1,3
Single source of truth	6,0	1,0	5,2	1,2
Easy for use	7,0	0,0	5,2	3,2

Sebelum DW di deploy, karyawan IT dan fungsional di semua organisasi pada studi kasus sering diminta untuk mendukung data melalui serangkaian pendekatan: mengunduh file, input ulang data dan membuat file ekstraksi. Hal yang sangat menyita waktu ini tidak lagi diperlukan setelah DW selesai dibangun. DW yang efektif bisa mengurangi biaya awal dalam pengambilan keputusan baru secara signifikan seperti yang terjadi di rumah sakit. Sebelum ada DW, pengambilan keputusan terkait dengan tagihan pasien asuransi butuh waktu yang relatif lama. Namun setelah ada DW pengambilan keputusan bisa langsung dilakukan setelah proses pembayaran. Dari sisi pengguna, lama waktu untuk akses data, memprosesnya maupun memformat ulang data menjadi format yang sesuai bisa diminimalisir. Karenanya pengambilan keputusan bisa dilakukan secara singkat.

**Korelasi antara dimensi kualitas data di fase analisis kebutuhan dengan manfaat data warehouse**

Identifikasi informasi yang relevan untuk pengambilan keputusan dengan *multi driven (user driven, goal driven, externally driven, data driven, and process driven)* adalah fokus dari analisis kebutuhan dalam pembangunan DW. Dukungan dimensi kualitas data di fase analisis kebutuhan memberikan kontribusi ke manfaat dari DW. Korelasi antara dimensi data kualitas di fase analisis kebutuhan dengan manfaat DW yang diperoleh sangat diperlukan untuk mengukur hubungan diantara keduanya. Secara lebih detail, dengan analisis bivariate korelasi diantara keduanya ditunjukkan pada Tabel 5.

Tabel 5  
Korelasi antara kualitas data di fase analisis kebutuhan dengan manfaat DW

		User driven																Goal driven															
		Heterogenous data integration		Time savings		More and better information		Better Decisions		Improve/redesign process business		Support for accompl strategic business obj		Single source of truth		Ease of use		Heterogenous data integration		Time savings		More and better information		Better Decisions		Improve/redesign process business		Support for accompl strategic business obj		Single source of truth		Ease of use	
		EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS
Information Quality	Content	Relevance	Comprehensiveness	-	0	+	0	+	-	+	+	+	+	+	+	-	+	+	+	+													
		Accuracy	-	-	-	-	+	-	+	-	+	-	+	-	+	0	+	-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		Clarity	-	+	+	+	+	+	+	+	+	+	+	+	+	0	+	+															
		Applicability	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
	Soundness	Conciseness	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		
		Consistency	-	0	-	0	-	-	+	0	+	-	+	-	+	0	+	+	-	0	-	0	-	-	+	0	+	-	+	-	+	+	
		Correctness	-	-	-	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	
	Access	Process	Currency	-	0	-	0	+	-	+	+	+	+	+	+	-	+	+	+	+													
			Convenience	-	-	+	0	+	-	+	+	+	+	+	+	-	+	+	+	+													
		Infrastructure	Timeliness	-	0	0	0	-	-	+	+	+	0	+	-	0	+	+	+	-	-	0	-	-	+	-	+	-	+	+	+	+	
Traceability			0	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+															
Interactivity			-	0	+	0	+	-	+	+	+	+	+	+	+	+	+	+	+														
Infrastructure		Accessibility	-	0	+	0	+	-	+	0	+	+	+	0	+	0	+	+															
		Security Information	-	0	-	0	+	+	0	+	0	0	+	-	+	+	-	-	-	0	-	0	+	+	0	0	0	-	0	-	0	-	
Speed	-	-	+	-	-	-	+	+	+	+	+	+	-	+	+	+	-	-	0	+	0	+	-	+	+	+	+	+	+	+	+		

EX : Expert Judgement  
CS : Case Studies  
  
+ positive correlation  
- negative correlation  
0 no correlation



## Korelasi antara Dimensi Kualitas Data Di Fase Disain Konseptual dengan Manfaat Data Warehouse

Disain konseptual DW (yang direpresentasikan oleh model multidimensional) memange hal-hal yang berkaitan dengan persepsi pengguna atas data. Model multidimensional menggambarkan informasi ini ke dalam fact dan dimensi. Manfaat dari metode ini bisa menampilkan cara berfikir dari perspektif pengguna. Dukungan dimensi kualitas data di fase disain konseptual bisa membawa kesuksesan dalam pembangunan DW sebagaimana tampak pada Tabel 6.

Tabel 6  
Korelasi antara kualitas data di fase disain konseptual dengan manfaat DW

		Conceptual																		
		Heterogenous data integration		Time savings		More and better information		Better Decisions		Improve/redesign process business		Support for accompl strategic business obj		Single source of truth		Ease of use				
Information Quality	Content	Relevance	Comprehensiveness		-	0	+	0	+		+	+	+	+	+	+	+	+	0	
		Accuracy																		
		Clarity																		
		Applicability																		
		Soundness	Conciseness																	
			Consistency																	
	Correctness																			
	Currency		-	+	-	+	-		+	+	+	+	+	+	+	0	0	+		
	Access	Process	Convenience																	
			Timeliness																	
			Traceability																	
			Interactivity																	
		Infrastructure	Accessibility																	
			Security Information																	
Information Maintainability																				
Speed			-	+	+	+	+		+	+	+	+	+	+	+	0	+	+		

Sebagaimana nampak pada Table 6, tidak ada korelasi antara *single source of truth* dan *currency* menurut studi kasus. Akan tetapi para pakar berbeda pendapat tentang hal tersebut dengan menyatakan bahwa ada korelasi positif antara *single source of truth* dan *currency*. Informasi yang selalu *up-to-date/ currency* akan menaikkan tingkat kepercayaan dari *single source of truth*. Detil korelasi antara kualitas data di fase disain konseptual dengan manfaat DW bisa dilihat di Tabel 6.

## Korelasi antara Dimensi Kualitas Data Di Fase Disain Logikal dengan Manfaat Data Warehouse

Disain logikal adalah fase yang paling atraktif karena memberikan efek yang kuat terhadap kinerja sistem. Kinerja DW yang baik adalah salah satu fitur penting yang harus dipastikan keberadaannya. Query untuk analisis adalah salah satu penelitian yang penting yang terkait dengan kinerja DW. Pertimbangan kualitas data di fase disain logikal bisa digunakan untuk memastikan akses yang efisien ke informasi agar DW bisa sukses diimplementasikan. Lebih jauh tentang korelasi antara kualitas data dengan manfaat DW ditampilkan di Tabel 7.

Tabel 7  
Korelasi antara kualitas data di fase disain logikal dengan manfaat DW

		Star Schema								OLAP																											
		Heterogenous data integration	Time savings	More and better information	Better Decisions	Improve/redesign process business	Support for accompli strategic business obj	Single source of truth	Ease of use	Heterogenous data integration	Time savings	More and better information	Better Decisions	Improve/redesign process business	Support for accompli strategic business obj	Single source of truth	Ease of use																				
Information Quality	Content	Relevance	Comprehensiveness	-	-	+	-	+	-	+	0	+	+	+	+	+	+	-	-	-	+	-	+	-	+	0	+	+	+	+	+	+	+	+	-		
		Accuracy	-	-	-	-	+	-	+	-	+	-	+	-	+	-	+	+																			
		Clarity																																			
		Applicability																																			
	Soundness	Conciseness																																			
		Consistency																																			
		Correctness																																			
		Currency	-	-	-	-	-	-	+	-	+	+	+	0	+	0	0	0	-	-	-	-	-	-	+	-	+	+	+	+	0	+	-	0	0	0	
	Access	Process	Convenience																																		
		Timeliness																	-	+	0	+	+	+	+	+	+	+	+	+	+	+	+	0	+	+	
		Traceability																																			
		Interactivity																																			
		Infrastructure	Accessibility	-	-	+	-	+	-	+	-	+	-	+	-	+	-	-	-	+	-	+	-	+	-	+	-	+	-	+	-	+	0	+	-		
		Security Information																																			
		Information Maintainability																																			
		Speed	-	-	+	-	+	-	+	-	+	-	+	0	+	+																					

EX : Expert Judgement  
CS : Case Studies  
  
+ positive correlation  
- negative correlation  
0 no correlation

Disamping korelasi terbalik dengan *heterogeneous data integration*, *currency* juga memiliki korelasi negatif dengan *time savings* dan *more & better information* pada proses star schema. Hal ini menunjukkan semakin banyak dimensi dan *fact* serta semakin terkini informasi yang akan ditampilkan akan mengurangi *time savings* dan jumlah informasi yang bisa ditampilkan.

Dari Tabel 7 bisa dilihat ada perbedaan persepsi antara pakar dengan studi kasus. Kinerja sistem bisa direpresentasikan oleh skema bintang (star schema) dan OLAP. Selama proses ini para pakar dan tim DW di studi kasus melihat fakta dengan persepsi yang berbeda. Di studi kasus, *accessibility* (sebagai contoh) tidak memiliki korelasi dengan *support for accomplishing strategic business objectives* di skema bintang dan OLAP. Namun para pakar justru menilainya memiliki korelasi positif, karena untuk mencapai tujuan bisnis yang strategis bisa didukung oleh tidak terhambatnya cara untuk mendapatkan informasi (*accessibility*) baik di skema bintang maupun OLAP. Korelasi lainnya bisa dilihat secara lebih detil di Tabel 7.

**Korelasi antara Dimensi Kualitas Data Di Fase ETL dengan Manfaat Data Warehouse**

Hampir semua dimensi kualitas data dibutuhkan ETL untuk mendapatkan manfaat DW. Penentuan hubungan antara kualitas data di ETL dengan manfaat DW bisa digunakan untuk mengukur hubungan diantara keduanya. Tabel 8 menjelaskan lebih detil tentang hal tersebut.

Tabel 8  
Korelasi antara kualitas data di fase ETL dengan manfaat DW

			ETL																
			Heterogenous data integration		Time savings		More and better information		Better Decisions		Improve/redesign process business		Support for accompl strategic business obj		Single source of truth		Ease of use		
			EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	
Information Quality	Content	Relevance																	
		Comprehensiveness																	
		Accuracy	-	+	-	+	+	+	+	0	+	0	+	+	+	-	+	0	
		Clarity																	
	Applicability	-	-	+	-	+	-	+	-	+	-	+	-	+	0	+	-		
	Soundness	Conciseness	-	-	+	-	+	-	+	-	+	-	+	-	+	0	+	-	
		Consistency	-	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	
		Correctness	-	0	+	0	+	+	+	-	+	-	+	-	+	-	0	-	
		Currency	-	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	
	Access	Process	Convenience	-	+	+	+	+	+	+	+	+	+	+	+	-	+	+	
			Timeliness	-	-	0	-	+	0	+	-	+	-	+	-	+	0	+	-
			Traceability	-	+	+	+	+	+	+	+	+	+	+	+	+	0	+	+
		Infrastructure	Interactivity																
			Accessibility	-	0	+	0	+	-	+	-	+	-	+	-	+	-	+	+
Security Information			-	-	+	-	+	-	+	-	+	-	+	-	+	-	-	+	
Information Maintainability			-	-	+	-	+	-	+	-	+	-	+	-	+	0	-	-	
Speed			-	0	+	0	+	-	+	+	+	+	+	+	+	+	+	+	

EX : Expert Judgement  
CS : Case Studies  
  
+ positive correlation  
- negative correlation  
0 no correlation

Secara umum, transformasi data ke format standar dan memuatnya ke pemrosesan data di DW membutuhkan korelasi positif antara manfaat DW dengan dimensi kualitas data berikut: *accuracy*, *consistency*, *applicability* dan *traceability*. Hanya saja manfaat *heterogeneous data integration* justru memiliki korelasi negatif terhadap semua dimensi kualitas data di ETL.

Ada beberapa perbedaan persepsi antara para pakar dengan studi kasus sebagaimana bisa dilihat pada Tabel 8. Perbedaan persepsi terjadi pada proses ekstraksi, transformasi dan pemuatan data ke DW. Sebagai contoh adalah *accuracy*. *Accuracy* tidak memiliki korelasi dengan manfaat DW *better decision* pada studi kasus. Sebaliknya para pakar menyatakan memiliki korelasi positif, karena informasi yang cukup presisi dan mendekati kenyataan (*accuracy*) adalah bagian penting dari keputusan yang lebih baik (*better decision*). Lebih jauh tentang hal ini bisa dilihat pada Tabel 8.

## Korelasi antara Dimensi Kualitas Data Di Fase Disain Fisikal dengan Manfaat Data Warehouse

DW biasanya meng-query data dalam jumlah yang sangat besar dari banyak tabel. Tanpa teknik optimasi yang efisien, query tersebut mungkin bisa di eksekusi dalam hitungan jam atau hari. Dampaknya disain fisikal harus bisa menjawab tantangan tersebut.

Hubungan yang timbal balik terjadi antara kinerja dan jumlah informasi yang tersedia. Semakin kompleks pemrosesan informasi yang akan ditampilkan, semakin tinggi kinerja yang dibutuhkan dan sebaliknya. Untuk memenuhi hal tersebut, pemahaman yang mendalam tentang hubungan antara dimensi kualitas data dengan disain fisikal yang terkait dengan manfaat DW sangat dibutuhkan. Korelasi lebih detail tentang hal ini bisa dilihat pada Tabel 9.

Tabel 9  
Korelasi antara kualitas data di fase disain fisikal dengan manfaat DW

		Physical																
		Heterogenous data integration		Time savings		More and better information		Better Decisions		Improve/redesign process business		Support for accompl strategic business obj		Single source of truth		Ease of use		
		EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	EX	CS	
Information Quality	Content	Relevance																
		Comprehensiveness	-	-	+	-	+	-	+	-	+	0	+	-	+	0	+	-
		Accuracy	-	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+
		Clarity																
		Applicability																
	Soundness	Conciseness	-	-	+	-	+	-	+	-	+	-	+	-	+	0	+	-
		Consistency	-	-	-	-	-	-	+	-	+	-	+	-	+	-	+	-
		Correctness																
		Currency	-	-	+	-	+	-	+	-	+	+	+	0	+	-	0	0
		Convenience																
Access	Process	Timeliness	-	-	0	-	-	-	+	-	+	-	+	-	+	0	+	+
		Traceability																
		Interactivity																
	Infrastructure	Accessibility																
		Security Information																
		Information Maintainability																
		Speed	-	0	+	0	+	-	+	+	+	+	+	-	+	+	+	+

EX : Expert Judgement  
CS : Case Studies  
  
+ positive correlation  
- negative correlation  
0 no correlation

Fokus dari disain fisikal adalah pembuatan DW tersentralisasi atau data mart secara individual. Dalam proses ini ada perbedaan persepsi antara para pakar dan studi kasus seperti *consiness*. Pada studi kasus, *consiness* mempunyai korelasi negatif dengan *ease of use*, meskipun para pakar tidak sepakat tentang hal tersebut. Para pakar menyatakan ada korelasi positif antara *conciseness* dan *ease of use*. Kemudahan penggunaan (*ease of use*) bisa dicapai melalui pembuangan informasi yang tidak diperlukan (*consiness*).

Secara umum, ketika diimplementasikan dengan benar, DW bisa memberikan manfaat kepada organisasi dengan informasi yang tepat di saat yang diperlukan dalam pengambilan keputusan. Akan tetapi karena kompleks dan bervariasinya fungsionalitas di dalam DW, maka perlu ada kepastian untuk mengakomodir kualitas data di setiap fase pengembangan DW.

Hal ini tercermin di sebuah bank dan organisasi pemerintah, dimana tidak mudah untuk menterjemahkan informasi yang disajikan oleh DW menjadi hasil bisnis yang positif. Mendapatkan informasi yang diperlukan sangat penting, namun bagaimana menggunakan informasi tersebut untuk tujuan pengambilan keputusan juga sama pentingnya.

Ketika banyak organisasi makin tergantung kepada DW untuk mendukung pengambilan keputusan atas hal-hal yang kritis, ada hal yang masih belum pasti yaitu kesuksesan atas proyek DW. Ada satu kebutuhan untuk menganalisis kesuksesan proyek DW. Penelitian ini bertujuan untuk menjawab hal tersebut melalui investigasi empiris tentang kesuksesan DW yang diukur dengan menggabungkan kualitas data ke semua fase pengembangan DW.

Perbaikan kualitas data dan pendeteksian cacat biasanya terkait dengan anggaran. Analisis biaya dan manfaat seharusnya dilakukan untuk mengoptimalkan manfaat yang bisa didapatkan dengan anggaran tertentu. Di organisasi pemerintahan dan BUMN (pada kasus yang diteliti), anggaran sudah tersedia, namun untuk institusi swasta, anggaran tidak selalu tersedia; karenanya anggaran untuk hal tersebut baru dibagi dan dilakukan setelah perencanaan proyek dilakukan.

Proses bisnis, produk dan jasa, tujuan dan unit organisasi bisa direfleksikan oleh data in DW. Kemampuan untuk menganalisis, merencanakan dan memberikan reaksi atas perubahan terhadap kondisi bisnis secepat mungkin adalah kunci untuk bertahan di dunia bisnis. Perubahan lingkungan bisnis membutuhkan waktu lama dan biaya yang besar sebelum terefleksikan di DW. Oleh karena itu pengembangan DW harus sejalan dengan lingkungan bisnis dan responsif terhadap perubahan di dunia bisnis. Terkait dengan studi kasus yang diteliti, hanya rumah sakit dan perusahaan asuransi yang senantiasa mengembangkan DW mereka agar sesuai dengan perubahan proses bisnis yang terjadi di organisasi mereka.

Pemrosesan informasi biasanya terkait dengan infrastruktur, sedangkan infrastruktur terkait dengan anggaran. Anggaran seharusnya tidak menjadi pembatas bagi organisasi untuk bisa memperoleh manfaat DW, karenanya pemrosesan informasi (*information processing*), anggaran dan manfaat DW berkorelasi terbalik.

Masalah yang paling menantang yang bisa ditemukan dalam kasus yang diteliti adalah kualitas informasi berkaitan erat dengan konten informasi dan akses informasi. Peningkatan kualitas konten informasi seharusnya dikaitkan dengan peningkatan identifikasi atas kebutuhan pengguna. Ini merupakan pekerjaan yang berat karena membutuhkan pemahaman yang mendalam mengenai proses bisnis dan konsep manajerial. Kualitas informasi juga dipengaruhi oleh kualitas akses ke informasi. IT yang mapan bisa mengurangi biaya yang harus dikeluarkan dalam pemrosesan informasi.

Konten informasi dan akses informasi sangat berkaitan dengan pendekatan *user-driven* dan *ETL*. Pendekatan *user-driven* adalah strategi utama untuk menunjukkan kualitas informasi yang disiapkan di DW. Sementara *data-driven* dan *ETL* lebih cocok untuk pemrosesan sumber data menjadi informasi yang cocok untuk pengambilan keputusan. Di *user-driven*, *data-driven* dan *ETL* banyak dimensi yang bisa digunakan untuk memberikan kontribusi atas manfaat DW.

Dalam pemrosesan konten informasi, *timeline* (tepat waktu/ IQ10) dan konsistensi (IQ6) sangat diperlukan untuk kesuksesan DW. Dimensi tersebut menunjukkan hubungan korelasi terbalik yang sangat kuat terhadap manfaat DW. Semakin banyak jumlah data yang harus diproses, semakin lama waktu yang diperlukan untuk memprosesnya dan semakin banyak konsistensi data yang dibutuhkan untuk mendukung proses bisnis yang ada sebagaimana ditunjukkan oleh perusahaan asuransi. Klaim asuransi tidak bisa diproses dengan cepat, meskipun sudah ada 120 juta anggota yang sudah tersimpan di database Oracle.

Akses ke informasi tidak lagi mencukupi bagi sebuah korporasi. Untuk memperoleh *value* (nilai) sebuah bisnis, informasi harus akurat, tepat waktu dan mudah diakses bagi siapapun yang membutuhkannya dalam konteks aktifitas bisnis yang mereka lakukan. DW yang dinamis sangat membutuhkan lingkungan yang dinamis juga. Oleh karenanya perlu dukungan *ETL* yang memiliki kinerja tinggi yang tidak hanya menawarkan *bandwith* yang lebar sebagai input, namun juga mampu menangani perubahan atau perpindahan bermacam-macam data.

Secara umum, ada korelasi terbalik antara semua dimensi kualitas data dan penggabungan data yang heterogen (*heterogeneous data integration/ DW1*) di setiap fase pembangunan DW. Hal ini menunjukkan hubungan timbal balik, dimana semakin heterogen data akan menurunkan nilai kualitas data. Semua dimensi kualitas data memiliki korelasi yang bermacam-macam, sebagaimana diuraikan di bawah ini:

- a. Korelasi positif terjadi di antara dimensi kualitas data dengan manfaat DW pada fase-fase berikut: *user driven*, *data driven*, *process driven*, *conceptual driven*, *star schema*, *OLAP*, *ETL* dan desain fisik.
- b. Berbagai macam korelasi terjadi antara dimensi kualitas data dengan fase *goal-driven*.
- c. Tidak ada korelasi terjadi antara dimensi kualitas data dengan fase *externally-driven*.

Temuan ini diperkuat dengan adanya fakta di bank (pada kasus yang dipelajari) dimana yang mendorong mereka untuk membangun DW adalah pihak luar (*externally-driven* – yaitu bank Indonesia untuk kepentingan pelaporan). Karenanya tidak ada korelasi antara kualitas data di fase *externally-driven* dengan manfaat DW. Akhirnya bisa disimpulkan tidak ada manfaat nyata dari pembangunan DW sebagaimana diungkapkan oleh tim DW di bank tersebut.

Temuan pada penelitian ini mengungkapkan bahwa penggabungan kualitas data di hampir semua fase pembangunan DW menunjukkan dampak yang positif terhadap manfaat DW yang bisa diperoleh. Hal ini nampak pada studi kasus yang dilakukan pada 5 organisasi dimana semakin banyak kualitas data yang disertakan pada fase-fase pembangunan DW menunjukkan kualitas DW yang lebih baik dan sebaliknya sebagaimana yang sudah dibahas didepan.

## **Kesimpulan**

DW yang berkualitas dapat menyebabkan kesuksesan implementasi DW. Kualitas ini tergantung pada kualitas keseluruhan fase pembangunan DW. Hal ini menunjukkan perlunya kualitas data diintegrasikan ke seluruh fase pembangunan DW untuk memperoleh manfaat dari kesuksesan DW.

Integrasi kualitas data ke seluruh fase pembangunan DW memiliki berbagai macam korelasi yaitu (1) korelasi positif yang terjadi di antara dimensi kualitas data dengan manfaat DW pada fase-fase berikut: *user driven*, *data driven*, *process driven*, *conceptual driven*, *star schema*, *OLAP*, *ETL* dan desain fisik. (2) korelasi negatif yang terjadi di antara manfaat DW dengan penggabungan data yang heterogen di setiap fase pembangunan DW (3) tidak ada korelasi yang terjadi di antara dimensi kualitas data di fase *external-driven* dengan manfaat DW.

## **Daftar Pustaka**

- Agosta, L. (2004). Data Warehousing Lessons Learned: A Time of Growth for Data Warehousing, in *DM Review Magazine*, 2004, pp. Retrieved on 29/3/2011, from [http://www.dmreview.com/article\\_sub.cfm?articleId=1012461](http://www.dmreview.com/article_sub.cfm?articleId=1012461).
- Agrawal, D. (2008). The Reality of Real-Time Business Intelligence in Business Intelligence for the Real-Time Enterprise. *Second International Workshop, BIRTE 2008*, Auckland, New Zealand, August 24, 2008, Revised Selected Papers. Lecture Notes in Business Information Processing, pp. 75-88.
- Al-Debei, M.M. (2011). Data Warehouse as a Backbone for Business Intelligence. *European Journal of Economics, Finance, and Administrative Sciences*, 33, 153-166.
- Alshawi, S., Saez-Pujol, I. and Irani, Z. (2003). Data warehousing in decision support for pharmaceutical R & D supply chain. *International Journal of Information Management*, 23, 259-268. [http://dx.doi.org/10.1016/S0268-4012\(03\)00028-8](http://dx.doi.org/10.1016/S0268-4012(03)00028-8).
- Amornbuth, C. (2015). The Relationship of the Quality Data Warehousing to Enhanced Perceived Net Profits and Decision Quality in the Enterprises. *Universal Journal of Management* 3(12): 514-520. DOI: 10.13189/ujm.2015.031206.
- Bilal Ali, Y. A. (2014). Challenges in the Successful Implementation of Data Warehouse. *Journal of Management Research*. ISSN 1941-899X. 2014. Vol 6, No 3.

- Coleho, P.S. and Esteves, S.P. (2007). The choice between a 5-point and 10-point scale in the framework of customer satisfaction measurement. *International Journal of Market Research*, Vol. 49. No. 3, 2007, pp.313-345.
- Flyvberg, B. (2006). Five misunderstandings about case study research. *Qualitative Inquiry* 12, 2, 219-245.
- Friedman, T. (2004). *Data Quality 'Firewall' Enhances Value of the Data Warehouse*. Gartner Reports, Apr. 2004.
- Griffin, R. K. (1998). Data Warehousing. *Cornell Hospitality Quarterly*, 39, 28-35.
- Hwang, M. I., & Xu, H. (2008). A Structural Model of Data Warehousing Success. *Journal of Computer Information Systems*, 49(1).
- Haug, A., Zachariassen, F and Van Liempd, D. (2011). The Costs of Poor Data Quality. *Journal of Industrial Engineering and Management (JIEM 2011)*, 4(2): 168-193.
- Joseph, M.V. (2013). Significance of Data Warehousing and Data Mining in Business Applications. *International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-1, March 2013*.
- Kumar, V. and Thareja, R. (2013). A Simplified Approach for Quality Management in Data Warehouse. *International Journal of Data Mining & Knowledge Management Process (IJDMP)* Vol.3, No.5, September 2013.
- Mukherjee, D. (2003). An empirical investigation of critical factors that influence data warehouse implementation success.
- MajidZaman, E., Quadri, S.M.K., and But, M. A. (2012). Information Integration for Heterogeneous Data Sources. *IOSR Journal of Engineering* Apr. 2012, Vol. 2(4) pp: 640-643. ISSN: 2250-3021.
- Munawar. (2016). Fase-Fase yang Umum Digunakan dalam Pembangunan Data Warehouse dan Kualitas Data yang Harus Dipertimbangkan di Setiap Fasenya. *Jurnal Ilmu Komputer*.
- Pandey, A. and Mishra, S. (2014). Moving from Traditional Data Warehouse to Enterprise Data Management : a Case Study. *Issues in Information Systems* Volume 15, Issue II, pp. 133-140, 2014.
- Rahman, N., Marz, J., and Akhter, S. (2012). An ETL Metadata Model for Data Warehousing. *Journal of Computing and Information Technology - CIT* 20, 2012, 2, pp. 95–111. doi:10.2498/cit.1002046.
- Ranjit Singh and Kawaljeet Singh. (2010). A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing. *IJCSI International Journal of Computer Science Issues*. Vol. 7, Issue 3, No 2 May 2010. ISSN : 1694-0784.
- Rudra, A and Yeo, E. (2000). Issues in User Perceptions of Data Quality and Satisfaction in Using a Data Warehouse - An Australian Experience, *Proceedings of the 33rd Hawaii International Conference on System Sciences*, IEEE 2000, pp. 1-7.

- Shahzad, M.K. (2012). Improving Business Processes using Processoriented Data Warehouse. Doctoral Dissertation in Computer and Systems Sciences School of Information and Communication Technologies KTH - Royal Institute of Technology, Stockholm, Sweden.
- Stake, R. E. (1995). *The Art of Case Study Research*, London: SAGE.
- Tellis, W. (1997). Introduction to case study. *The Qualitative Report*, 3 (2), 1-11.
- Torlone, R. (2008). Two Approaches to the Integration of Heterogeneous Data Warehouses. *Distrib. Parallel Databases* 23(1), pp. 69–97.
- Watson, H.J. and Haley, B.J. (1997). Data warehousing: a framework and survey of current practices, *Journal of Data Warehousing* 2 (1) (1997) 10–17.
- Watson, H. J., Goodhue, D. L. and Wixom, B. H. (2002). The benefits of data warehousing: why some organizations realize exceptional payoffs. *Information and Management*, vol. 39, pp. 491-502.
- Watson, H. and Haley, B. (1998). A Structural Model of Data Warehousing Success. *Journal of Data Warehousing*, 2, 10-17.
- Winter, R and Strauch, B. (2003). A method for demand-driven information requirements analysis in data warehousing. In *Proceedings of Hawaii International Conference on System Sciences*, Hawaii, 1359-1365.
- Wixom, B. H. and Watson, H. J. (2001). An Empirical Investigation of the Factors Affecting Data Warehousing Success. *MIS Quarterly*, vol. 25, pp. 17-41.