

Contents

1	Introduction	1
1.1	Overview	1
1.2	Global Perspective on Biodiversity Information	1
1.3	Malaysian Perspective of Biodiversity Information	2
1.3.1	Physical Repositories	3
1.3.2	Virtual Repositories	4
1.3.3	Problems Concerning Biodiversity Information in Malaysia	4
	References	5
2	Preliminary Study	7
2.1	Introduction	7
2.2	First Phase Implementation	7
2.2.1	Implementation of DiGIR Using an Algae Database	7
2.2.2	Installation of a Simple Portal on a Remote Machine	10
2.3	Second Phase Implementation	12
2.3.1	DiGIR Software Setup and Configuration	12
2.3.2	Provider Setup	14
2.3.3	Running the DiGIR Portal	15
2.3.4	DiGIR Implementation Results	16
2.4	Discussion	20
2.5	Conclusion	20
	References	21
3	Literature Review	23
3.1	Introduction	23
3.2	Existing Database Integration System	23
3.2.1	GenoMax	23
3.2.2	Kleisli	24
3.2.3	Discovery Link	25
3.2.4	SPICE	27
3.2.5	Distributed Generic Information Retrieval	28
3.2.6	Dimensions of Variations for Existing Systems	30

3.3	Related Components and Technologies for Database Integration	30
3.3.1	Databases	30
3.3.2	Database Integration Database Management Systems and Relational Database Management Systems	32
3.3.3	Biodiversity Data Formats	33
3.4	Technical Aspects of Database Integration	39
3.4.1	Integration Approaches	39
3.4.2	Web Pages and Interfaces	39
3.4.3	Database Integration with the Web	46
3.4.4	Markup Languages for Database Integration System	47
3.4.5	Scripting Languages	52
3.4.6	Query Languages and Database Connectivity	54
3.4.7	Web Services	56
3.5	Conclusion	57
	References	58
4	Methodology	63
4.1	Introduction	63
4.2	Methodology for Biodiversity Data Format and Development of Relational Biodiversity Databases	63
4.2.1	Requirements Analysis for a Biodiversity Data Format	64
4.2.2	Development of Relational Biodiversity Databases	65
4.3	Methodology for Building the Proposed Database Integration System	67
4.3.1	Selection of Development Tools	68
4.3.2	Selection of Integration Approach	68
4.3.3	Selection of Relational Biodiversity Databases for CABIR	69
4.3.4	Development of CABIR	70
4.4	System Testing Methodology	72
4.5	Conclusion	73
	References	74
5	Biodiversity Databases	75
5.1	Introduction	75
5.2	Biodiversity Data Format	75
5.3	Relational Biodiversity Databases	75
5.3.1	South East Asia Fruit Flies Database	78
5.3.2	Host-Parasite Database	78
5.3.3	Malaysian Indigenous Fern Database	79
5.3.4	Herbarium of Zingiberaceae Database	83
5.4	Conclusion	83

6 Proposed Solution 85

6.1 Introduction 85

6.2 Hierarchical View of CABIR 85

6.3 Proposed Database Integration System: CABIR. 85

6.3.1 System Architecture 85

6.3.2 System Features 93

6.3.3 CABIR and FileMaker Databases 94

6.3.4 Image Retrieval 94

6.3.5 Flexible Data Format. 96

6.3.6 Link to Remote Cross Platform Providers 98

6.3.7 Link to Heterogeneous Database
Management Systems 98

6.3.8 Error Messages. 103

6.4. System Design 103

6.4.1 Main Page 104

6.4.2 Search Page 104

6.4.3 Results of Search 105

6.4.4 Searching all Records Within a Database. 105

6.5 Results of System Testing 108

6.5.1 Functionality and Usability 108

6.5.2 Reliability and Efficiency. 108

6.5.3 Maintainability and Portability 109

6.6 Conclusion 111

References 111

7 Concluding Remarks 113

7.1 Introduction 113

7.2 CABIR in Biodiversity Databases 113

7.3 CABIR as a Simple and Dynamic Database
Integration System. 114

7.4 CABIR and Data Security 114

7.5 A Summary of Comparison Between CABIR and DiGIR. 114

7.6 CABIR as a Model Database Integration System. 115

7.7 CABIR Limitations 115

7.8 Future Focus. 117

7.9 Conclusion 117

References 118

**Appendix A: Survey for Developing Standard for Biodiversity
Databases** 119

Appendix B: Darwin Core V2 Searchable Concepts 121