

# **The Development of Open Government Data**

This page intentionally left blank

# **The Development of Open Government Data: Connecting Supply and Demand Through Portals**

BY

**DI WANG**

*Wuhan University, China*

**DEBORAH RICHARDS**

*Macquarie University, Australia*

**AYSE AYSIN BILGIN**

*Macquarie University, Australia*

And

**CHUANFU CHEN**

*Wuhan University, China*



United Kingdom – North America – Japan – India – Malaysia – China

Emerald Publishing Limited  
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2022

Copyright © 2022 Di Wang, Deborah Richards, Ayse Aysin Bilgin and Chuanfu Chen.  
Published under exclusive licence by Emerald Publishing Limited.

**Reprints and permissions service**

Contact: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

No part of this book may be reproduced, stored in a retrieval system, transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the publisher or a licence permitting restricted copying issued in the UK by The Copyright Licensing Agency and in the USA by The Copyright Clearance Center. Any opinions expressed in the chapters are those of the authors. Whilst Emerald makes every effort to ensure the quality and accuracy of its content, Emerald makes no representation implied or otherwise, as to the chapters' suitability and application and disclaims any warranties, express or implied, to their use.

**British Library Cataloguing in Publication Data**

A catalogue record for this book is available from the British Library

ISBN: 978-1-80262-316-1 (Print)

ISBN: 978-1-80262-315-4 (Online)

ISBN: 978-1-80262-317-8 (Epub)



**ISOQAR**  
REGISTERED

Certificate Number 1985  
ISO 14001

ISOQAR certified  
Management System,  
awarded to Emerald  
for adherence to  
Environmental  
standard  
ISO 14001:2004.



INVESTOR IN PEOPLE

# Table of Contents

List of Figures and Tables	<i>vii</i>
About the Authors	<i>xiii</i>
Abbreviations	<i>xv</i>
Preface	<i>xvii</i>
Acknowledgments	<i>xix</i>
<b>Chapter 1 Introduction</b>	<b>1</b>
<b>Chapter 2 Key Concepts and Literature Review</b>	<b>17</b>
<b>Chapter 3 An Evaluation of the Supply Side of OGD</b>	<b>39</b>
<b>Chapter 4 Understanding Citizens' Demands for OGD and OGD Utilization</b>	<b>65</b>
<b>Chapter 5 A Comparison of the Supply Side and Demand Side of OGD Portals</b>	<b>99</b>
<b>Chapter 6 The Usability of OGD Portals</b>	<b>121</b>
<b>Chapter 7 Citizens' Acceptance and Utilization of OGD Portals: An Experiment Using a Virtual Agent</b>	<b>151</b>
<b>Chapter 8 Discussion</b>	<b>223</b>
<b>Chapter 9 Conclusion</b>	<b>231</b>

*vi Table of Contents*

Bibliography	241
Appendix A: Complete Survey for Stage 2	257
Appendix B: Dialogues of the Virtual Agent for Stage 4	265
Appendix C: Complete Instruments for Stage 4	313
Index	335

# List of Figures and Tables

Figure 1.	Relationship between Open Data, OGD, and PSI.	2
Figure 2.	Research Roadmap.	10
Figure 3.	Structure and Flow of the Selected Literature Review Topics.	18
Figure 4.	Open Government Data Lifecycle.	21
Figure 5.	Roadmap of Concepts in Chapter 2 and Later Studies.	37
Figure 6.	AHP Hierarchy Structure (See Table 7 for A1–A15 Descriptions).	54
Figure 7.	Comparison of Evaluation Scores by Portal.	58
Figure 8.	Number of Different Format Datasets provided by Portals.	58
Figure 9.	Research Model.	69
Figure 10.	Citizens' Acceptance of Data-Acquisition Methods.	86
Figure 11.	Reported Usage Frequency and Preferences of Keyword Search and Browse.	86
Figure 12.	Differences in Acceptance of Regular Recommendations between Genders.	87
Figure 13.	Frequency and Percentage of Help Functions.	88
Figure 14.	Frequency of People's Preference of Virtual Agents' Appearance.	89
Figure 15.	Distribution of Different Gender Participants' Preferences of Virtual Agent's Role.	90
Figure 16.	Distribution of Different Aged Participants' Preferences of Virtual Agent's Appearance.	91

Figure 17.	Distribution of Different Education Background Participants' Preferences of Virtual Agent's Gender.	91
Figure 18.	The Refined Research Model.	97
Figure 19.	Modified DOI Research Model for Connecting OGD Demand and Supply.	101
Figure 20.	Modified Research Model With Conclusions.	116
Figure 21.	Screenshot of the Shanghai OGD Portal.	148
Figure 22.	Integrated Research Model.	155
Figure 23.	Experimental Design.	157
Figure 24.	Home Page of the Imitated Portal.	158
Figure 25.	Interface for the Conversational Virtual Agent.	160
Figure 26.	Experimental Procedure.	164
Figure 27.	Homepage of the Imitated Portal With the Virtual Agent.	168
Figure 28.	Methods of Knowing About OGD and OGD Portals.	174
Figure 29.	Types of Data Used.	174
Figure 30.	Usage Frequency of OGD and OGD Portals.	174
Figure 31.	Purposes of Use of OGD.	175
Figure 32.	Frequency Distribution Histogram of Playing Computer Games.	175
Figure 33.	Process for Building DOI Models.	214
Figure 34.	Relationship Map of the Variables in DOI Models.	218
Figure 35.	Process for Building Model to Predict Citizens' Possibility of Using the OGD Portal.	219
Figure 36.	Model for Predicting Citizens' Possibility of Using the OGD Portal.	221
Figure A1.	An Example of an OGD Portal.	261
Table 1.	Structure of Objectives and Research Questions.	8
Table 2.	Comparison of Technology Acceptance Models and Theories.	29
Table 3.	Comparison of OGD Portal Evaluation Frameworks.	42
Table 4.	Evaluation Aspects of Frameworks and Principles.	43
Table 5.	Exclusion of Evaluation Elements.	46

Table 6.	Categories for Data Collection.	50
Table 7.	Local OGD Portal Evaluation Framework.	51
Table 8.	Province-Level OGD Portals in China Included for Evaluation.	56
Table 9.	Evaluation Results.	57
Table 10.	Data Accessibility by Category.	59
Table 11.	Data Quality by Category.	60
Table 12.	Correlation Analysis for Contexts and Number of Datasets.	61
Table 13.	Descriptions of OGD Attributes.	71
Table 14.	Socio-Demographic Characteristics.	74
Table 15.	Likelihood Ratio Tests of Citizens' Knowledge of OGD.	75
Table 16.	Reliability and Adequacy Tests of all Scales.	76
Table 17.	Descriptive Analysis of Citizens' Demands for OGD Subjects.	77
Table 18.	<i>T</i> -Tests of Citizens' Knowledge of OGD and Their Demands of OGD Subjects.	79
Table 19.	One-Way ANOVA of Age, Education, Occupation, and OGD Subjects.	81
Table 20.	Descriptive Statistics of OGD Attributes.	81
Table 21.	One-Way ANOVA of Age, Education, Occupation, and OGD Attributes.	82
Table 22.	<i>T</i> -Tests of Citizens' Knowledge of OGD and OGD Attributes.	83
Table 23.	Chi-Square Tests of OGD Utilization Purpose.	85
Table 24.	Chi-Square Tests of Citizens' Preferences of Virtual Agents' Appearance.	90
Table 25.	<i>T</i> -Tests of OGD Utilization Purposes and Demands for OGD Subjects.	94
Table 26.	<i>T</i> -Tests of Data-Acquisition Methods and OGD Attributes.	95
Table 27.	Pearson Correlation Analysis of Data Acquisition and OGD Attributes.	96

Table 28.	Pearson Correlation Analysis of Need of Help and OGD Attributes.	97
Table 29.	Comparison Model of OGD's Supply Side and Demand Side.	102
Table 30.	Description of Cases for Supply Side.	107
Table 31.	Reliability and Adequacy Tests of the Model.	109
Table 32.	Online Visibility of OGD Portals.	110
Table 33.	Comparative Results for Supply Side and Demand Side in Relative Advantage.	111
Table 34.	Comparative Results for Supply Side and Demand Side in Compatibility.	112
Table 35.	Comparative Results for Supply Side and Demand Side in Complexity.	113
Table 36.	Correlations Between Observability, Relative Advantage, and Compatibility on Demand Side.	114
Table 37.	Correlation Analysis of Relative Advantage and Compatibility.	115
Table 38.	List of Selected OGD Portals.	124
Table 39.	Selection of Usability Principles.	126
Table 40.	Criteria for Usability Principles.	133
Table 41.	Fleiss's Kappa for Principles and Criteria.	139
Table 42.	Average Values for Principles and Criteria.	140
Table 43.	Kruskal–Wallis <i>H</i> Test of OGD Portals.	142
Table 44.	Portals' Average Scores by Principle and One-Way ANOVA Results.	144
Table 45.	Portals' Average Scores by Criterion and One-Way ANOVA Results.	145
Table 46.	Socio-Demographics of the Participants.	170
Table 47.	Participants' Knowledge of OGD and OGD Portals.	173
Table 48.	Description of Big Five Personality Dimensions and Trust Tendency.	177
Table 49.	Reliability and Adequacy Tests of Scales.	178
Table 50.	Results of Shapiro–Wilk Normality Test.	180
Table 51.	Descriptive Analysis of Completion of Tasks.	182

Table 52.	Mann–Whitney <i>U</i> Test of Completion of Tasks Between the Two Groups.	183
Table 53.	Descriptive Analysis of TAM Importance Scale.	184
Table 54.	Descriptive Analysis of TAM Perceived Ease of Use Scale.	185
Table 55.	Analysis of PEOU and Socio-Demographic Variables.	187
Table 56.	Analysis of PEOU and OGD and OGD Portal Experience Variables.	189
Table 57.	Descriptive Analysis of TAM Perceived Usefulness Scale.	192
Table 58.	Pearson’s Correlation Analysis of Importance and PEOU.	193
Table 59.	Pearson’s Correlation Analysis of Importance and PU.	196
Table 60.	Pearson’s Correlation Analysis of PEOU and PU.	198
Table 61.	Descriptive Analysis of DOI for OGD Portal.	200
Table 62.	Analysis of Demographic Background and DOI.	201
Table 63.	Descriptive Analysis of DOI for the Virtual Agent.	202
Table 64.	Analysis of Demographic Background and DOI.	203
Table 65.	Analysis of Participants’ Past Experience With OGD and Trustworthiness.	206
Table 66.	Descriptive Analysis of Rapport.	206
Table 67.	Analysis of Socio-Demographic Background and Rapport.	208
Table 68.	Chi-Square Tests of Demographic Background and Two-Task Accuracy Rate.	209
Table 69.	<i>T</i> -Tests of TAM Scales and Two-Task Accuracy Rate.	210
Table 70.	Mann–Whitney <i>U</i> Tests of DOI and Trustworthiness Scales and Two-Task Accuracy Rate.	211
Table 71.	<i>T</i> -Tests of Rapport Scales and Two-Task Accuracy Rate.	212
Table 72.	Mann–Whitney <i>U</i> Tests of DOI for Virtual Agent and Two-Task Accuracy Rate.	213

*xii List of Figures and Tables*

Table 73.	Cutoff Points and Categorization of DOI Variables.	214
Table 74.	Accuracy Rates of Models for DOI Variables.	215
Table 75.	Models for DOI Variables.	216
Table 76.	Research Objectives and Research Questions of This Book.	232
Table B1.	Chinese Dialogue CSV Files for Introduction.	268
Table B2.	Chinese Dialogue CSV Files for FAQ.	288
Table B3.	Chinese Dialogue CSV Files for Data Application.	305
Table B4.	Chinese Dialogue CSV Files for Feedback.	310

## About the Authors

**Di Wang** is a Postdoc Researcher from School of Information Management, Wuhan University. She received her PhD in Computing from Macquarie University and PhD in Information Resources Management from Wuhan University. Her research interests include big data analytics, open data policy, open data resources, and data services. She published in journals including *Government Information Quarterly* and *Library Hi Tech* and presented at conferences including JCDL, ICADL, iConference, etc. She has been awarded several research grants from the National Science Foundation of China, the National Social Science Foundation of China, and China Postdoctoral Science Foundation.

**Deborah Richards** is a Professor in the Department of Computing at Macquarie University. Following 20 years in the IT industry during which she completed a BBus (Comp and MIS) and MAppSc (InfoStudies), she completed a PhD in Artificial Intelligence on the Reuse of Knowledge at the University of New South Wales and joined academia in 1999. While she continues to work on solutions to assist decision-making and knowledge acquisition, for the past decade, her focus has been on intelligent systems, agent technologies, and virtual worlds to support human learning, health, and well-being. She has over 300 refereed publications and been awarded numerous research and teaching grants from the Australian Research Council and other research funders.

**Ayse Aysin Bilgin** is an Associate Professor, active in the statistical analysis and design of studies in health and statistics education. Her research in health developed to include adolescent health and recently studies related to breast cancer, MS, and Alzheimer's disease. She is one of the pioneers on applying machine learning to health data sets. She was awarded "Excellence in research: Five Future-shaping Priorities (Healthy People)" by Macquarie University. Ayse has won several learning and teaching awards for her outstanding contributions to student learning such as ALTC citation for "Outstanding Contributions to Student Learning." Her pedagogical research interests are focused on statistics education, such as learning approaches in statistics, impact of learning spaces to students' learning. She is the President of International Association for Statistical Education (IASE) (July 2021–July 2023).

**Chuanfu Chen** is a Distinguished Research Professor and Dean of the Graduate School, Wuhan University. From 2005 to 2012, he was Dean of the School of

Information Management of Wuhan University and was the Deputy President of Library Society of China from 2004 to 2012. He earned his MLS in 1986, and his PhD in Law in 2001, both from Wuhan University. His research focuses on library development, public access to information resources, big data governance, and copyright. He has published in many journals including *ASLIB Proceedings*, *College & Research Libraries*, *Electronic Library*, *Journal of Library Science of China*, and *Scientometrics*, among others. He has also been awarded numerous research grants from the National Social Science Foundation and the Natural Science Foundation of China.

# Abbreviations

AHP	Analytic Hierarchy Process
API	Application programming interface
CR	Consistency ratio
DOI	Diffusion of Innovation theory
FAQ	Frequently asked question
GDP	Gross domestic product
HCI	Human–computer interaction
ICL	Intention – complexity
ICP	Intention – compatibility
ICT	Information and communication technology
IOB	Intention – observability
IRA	Intention – relative advantage
ITR	Intention – trialability
MM	Motivational model
OGD	Open government data
PBC	Perceived behavioral control
PEOU	Perceived ease of use
PSI	Public sector information
PU	Perceived usefulness
RQ	Research question
TAM	Technology acceptance model
TIPI	Ten-item personality inventory
TPB	Theory of planned behavior
TRA	Theory of reasoned action
TTS	Text to speech
UTAUT	Unified theory of acceptance and use of technology

This page intentionally left blank

# Preface

Open government data (OGD) have developed rapidly in these years due to various benefits that can be derived through transparency and public access. However, researchers emphasize lack of use instead of lack of disclosure as a key problem in OGD's present development. Previous studies look into this issue either from the supply side, focusing on data quantity and quality, or from the demand side, focusing on factors that affect users' acceptance of OGD, but seldom consider both sides at the same time. This book makes a comparison of the supply and demand sides of OGD and explores possible directions for the future development of OGD portals based on the discovered mismatches lying between the two sides.

The core purpose of this book is to improve OGD utilization by balancing the supply side and demand side of OGD according to the demands of citizens through the development of OGD portals. To achieve this objective, four connected studies were designed and carried out. The first study built an evaluation framework for understanding the development of the supply side of OGD by evaluating existing Chinese province-level OGD portals. Sequentially, with citizens as primary users on the demand side and the major beneficiaries of OGD, the second study focused on a survey conducted to analyze citizens' awareness and utilization of OGD portals. A third study compared the supply and demand sides of OGD based on Diffusion of Innovation (DOI) theory, using the data collected in the previous two studies. A final study tested the proposed usability criteria for building an OGD portal in helping users to use the data on the portal by carrying out a between-subjects experiment. All case studies in these four parts were carried out in China.

This research finds that Chinese OGD portals are in an early stage of development. Citizens have limited awareness of OGD and OGD portals. Significant correlations are recognized among citizens, and their demands and utilization of OGD. Mismatches lie between the supply and demand sides of OGD. Following the proposed usability criteria for building an OGD portal could improve citizens' proper utilization of OGD. Future directions for developing OGD are identified.

The key contribution of this book to the present literature is the theoretical and practical understanding of OGD and its user, as well as proposing directions for OGD portals' future development in order to encourage citizens' OGD utilization.

This book originates from the PhD research study of Dr Di Wang in Macquarie University, Australia, under the supervision of Prof. Deborah Richards, Assoc. Prof. Ayse Aysin Bilgin, and Prof. Chuanfu Chen.

# Acknowledgments

First and foremost, I would like to express my great gratitude to my supervisor, Prof. Deborah Richards. Your immense knowledge and research experiences guided me through my whole research. This book would not be completed without your support and advice.

I would also like to express my sincere appreciation to my associate supervisor, Assoc. Prof. Ayse Aysin Bilgin, and my adjunct supervisor, Prof. Chuanfu Chen, for your support, motivation, and suggestions during all the time of my research and writing of this book.

I would like to acknowledge the contributions of my parents, Mr Lin Wang and Mrs Lifen Ma for your continuous loving support. I could not have completed this work without your encouragement.

Besides my supervisors and parents, I would like to thank the people who have offered generous assistance with my studies, particularly Ms Meredith Taylor, who helped me with building the virtual agent for the experiment, and Ms Xu Chen, who helped me with setting OGD portals for the experiment.

I extend my thanks to Macquarie University, which provided the MQ Research Training Program (MQRTP), and the China Scholarship Council, which has provided financial assistance during my research. I would also like to express my appreciation for the kind support of the School of Information Management, Wuhan University, that has allowed me to complete this book.

Finally, I would like to thank everybody who was important and supportive to the successful realization of my first book, as well as expressing my apology that I could not mention each personally one by one.

Di Wang  
Wuhan, China, December 2021