# **Lecture Notes in Computer Science**

16046

Founding Editors

Gerhard Goos Juris Hartmanis

# **Editorial Board Members**

Elisa Bertino, *Purdue University, West Lafayette, IN, USA*Wen Gao, *Peking University, Beijing, China*Bernhard Steffen , *TU Dortmund University, Dortmund, Germany*Moti Yung , *Columbia University, New York, NY, USA* 

The series Lecture Notes in Computer Science (LNCS), including its subseries Lecture Notes in Artificial Intelligence (LNAI) and Lecture Notes in Bioinformatics (LNBI), has established itself as a medium for the publication of new developments in computer science and information technology research, teaching, and education.

LNCS enjoys close cooperation with the computer science R & D community, the series counts many renowned academics among its volume editors and paper authors, and collaborates with prestigious societies. Its mission is to serve this international community by providing an invaluable service, mainly focused on the publication of conference and workshop proceedings and postproceedings. LNCS commenced publication in 1973.

Robert Wrembel · Gabriele Kotsis · A Min Tjoa · Ismail Khalil Editors

# Database and Expert Systems Applications

36th International Conference, DEXA 2025 Bangkok, Thailand, August 25–27, 2025 Proceedings, Part I



Editors
Robert Wrembel D
Poznan University of Technology
Poznań, Poland

A Min Tjoa D Vienna University of Technology Vienna, Wien, Austria Gabriele Kotsis

Johannes Kepler University Linz

Linz, Austria

Ismail Khalil D
Johannes Kepler University Linz
Linz, Austria

ISSN 0302-9743 ISSN 1611-3349 (electronic) Lecture Notes in Computer Science ISBN 978-3-032-02048-2 ISBN 978-3-032-02049-9 (eBook) https://doi.org/10.1007/978-3-032-02049-9

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

## **Preface**

We present the proceedings of the 36th International Conference on Database and Expert Systems Applications (DEXA 2025). DEXA was established in 1990 under the name International Conference on Database and Expert Systems Applications, which has persisted until today. The conference has been running annually without a break for over three decades, serving as a premier international forum for researchers, practitioners, and industry experts in the fundamental fields of data modeling, databases and data storage systems, data engineering, data analytics, data science, and recently - machine learning and artificial intelligence, for standard and big data.

This year, DEXA was held on 25–27 August, 2025 in Bangkok, Thailand. The conference received 123 submissions. From this set, 35 were accepted as regular papers (giving an acceptance rate of 28%).

DEXA 2025 is proud to have accepted also 22 short papers. They offer a vital platform for presenting innovative projects and preliminary results, novel ideas, ongoing research, or concise technical contributions that may not yet be mature enough for a full paper but have significant potential for future development and promise vibrant discussions. The inclusion of short papers encourages broader participation and facilitates the timely dissemination of emerging research.

The selection of all papers was based on evaluations by the Program Committee members. Each paper was single-blindly evaluated by three members. Here we express our gratitude to the PC members of DEXA 2025, for their timely and thorough evaluations.

As in the past, the DEXA 2025 proceedings consist of two volumes with regular and short papers. The accepted papers cover a variety of research topics on both theoretical and practical aspects. The papers cover among others the following topics: (1) large language models, (2) data quality, (3) applications of machine learning and artificial intelligence, (4) classification techniques, (5) image processing, analytics, and vision systems, (6) recommender techniques, (7) data integration techniques, (8) optimization methods, (9) graph applications, (10) data analytics methods, (11) security and privacy, and (12) benchmarks and surveys.

This year, DEXA introduced so-called short invited talks, with the goal to present trends in data and knowledge engineering in a less formal setting. Four such talks were accepted for the conference.

Also, this year, for the first time, the best papers will be published in a special issue of the Data & Knowledge Engineering (DKE, Elsevier) journal, entitled *Integrating Machine Learning and Data Engineering for Advanced Data Science*. This special issue will also include the best papers from the DAWAK 2025 conference. Taking the opportunity, the PC-chair would like to thank the DKE Editor-in-Chief, Carson Woo, for his approval of the special issue. Special gratitude goes to Ismail Khalil - a Steering Committee member of DEXA/DAWAK and the main organiser of these events. His invaluable

### vi Preface

help in all tasks related to organizing DEXA 2025 materialized in these two volumes of the proceedings and the event itself.

August 2025,

Robert Wrembel Gabriele Kotsis A Min Tjoa Ismail Khalil

# **Organisation**

# **Program Committee Chair**

Robert Wrembel Poznań University of Technology, Poland

### **Publicity Chairs**

Aziz Nanthaamornphong Prince of Songkla University, Thailand Putu Wuri Handayani University of Indonesia, Indonesia

### **Steering Committee**

Gabriele Kotsis Johannes Kepler University Linz, Austria A Min Tjoa Vienna University of Technology, Austria

Lukas FischerSoftware Competence Center Hagenberg, AustriaBernhard MoserSoftware Competence Center Hagenberg, Austria

Christine Strauss University of Vienna, Austria

Ismail Khalil Johannes Kepler University Linz, Austria

# **Program Committee Members**

A Min Tjoa Vienna University of Technology, Austria

Abdelkader Hameurlain IRIT, Toulouse University, France

Abdessamad Imine Loria, France

Adam Przybylek Gdańsk University of Technology, Poland Adriana Marotta Universidad de la República, Uruguay

Aida Omerovic SINTEF and NTNU, Norway Allel Hadiali LIAS/ENSMA, France

Andreas Ekelhart Secure Business Austria, Austria

Anne Kayem University of Exeter, UK
Bala Srinivasan Monash University, Australia

Bartosz Bebel Poznań University of Technology, Poland

Bettina Fazzinga University of Calabria, Italy

Brahim Ouhbi ENSAM, Morocco Cedric Du Mouza CNAM, France

Christian Thomsen Aalborg University, Denmark

Dawid Wiśniewski Poznań University of Technology, Poland Deborah Dahl Conversational Technologies, USA

Ela Pustulka FHNW University of Applied Sciences and Arts,

Switzerland

Elio Masciari University of Naples Federico II, Italy

Erich Neuhold University of Vienna, Austria

Eunika Mercier-Laurent University of Reims Champagne Ardenne and

IFIP. France

Flavio Ferrarotti Software Competence Centre Hagenberg, Austria Flavius Frasincar Erasmus University Rotterdam, The Netherlands

Florence Sedes IRIT - University of Toulouse, France Franck Morvan RIT - University of Toulouse, France

Giovanna Guerrini University of Genova, Italy

Gheorghe Cosmin Babeş-Bolyai University, Romania Silaghi Hamidah Ibrahim Universiti Putra Malaysia, Malaysia

Hendrik Decker Ludwig Maximilian University of Munich,

Germany

Hiroyuki Toda Yokohama City University, Japan

Idir Amine Amarouche USTHB, Algeria

Ionut Iacob Georgia Southern University, USA

Isao Echizen University of Tokyo, Japan Ismael Navas-Delgado University of Málaga, Spain

Ivan Izonin Lviv Polytechnic National University, Ukraine

Ivanna Dronyuk Jan Dlugosz University, Poland

Javier Nieves Azterlan, Spain

Jean-Paul Kasprzyk University of Liège, Belgium Jérôme Darmont Université Lyon 2, France Jianwei Zhang Iwate University, Japan

Johann Gamper Free University of Bozen-Bolzano, Italy

Jorge Lloret University of Zaragoza, Spain

Josef Küng Johannes Kepler University Linz, Austria Jun Miyazaki Tokyo Institute of Technology. Japan Kamonluk Suksen Chulalongkorn University, Thailand Karim Benouaret Université Claude Bernard Lyon 1, France

Lars Moench University of Hagen, Germany

Laura Erhan Free University of Bozen-Bolzano, Italy

Laurent d'Orazio IRISA, France

Lenka Lhotska CVUT, Czech Republic
Luca Caviglione IMATI-CNR, Italy
Manfred Hauswirth TU Berlin, Germany
Manolis Gergatsoulis Ionian University, Greece

Marcin Paprzycki IBS PAN and WSM, Poland Marinette Savonnet University of Burgundy, France

Markus Endres Munich University of Applied Sciences, Germany

Massimo Guarascio ICAR-CNR, Italy

Maude Manouvrier Université Paris Dauphine - PSL, France Michal Kratky VSB-Technical University of Ostrava, Czech

Republic

Michael Sheng Macquarie University, Australia Mizuho Iwaihara Waseda University, Japan

Mustafa Atay Winston-Salem State University, USA

Nazha Selmaoui University of New Caledonia, New Caledonia

Noura Faci Université Lyon 1, France

Olivier Teste IRIT, France

Pavlo Radiuk Khmelnytskyi National University, Ukraine
Paweł Misiorek Poznań University of Technology, Poland
Peiquan Jin University of Science and Technology of China,

China

Petra Asprion FHNW University of Applied Sciences and Arts,

Switzerland

Rachid Anane Coventry University, UK
Riad Mokadem University of Toulouse, France

Riccardo Albertoni CNR-IMATI, Italy

Samira Maghool Università degli Studi di Milano, Italy

Sergio Ilarri University of Zaragoza, Spain Soon Chun City University of New York, USA

Srinath Srinivasa International Institute of Information Technology

Bangalore, India

Stéphane Jean University of Poitiers, ISAE-ENSMA, LIAS,

France

Sven Groppe University of Lübeck, Germany

Talel Abdessalem Télécom Paris, France

Toshiyuki Amagasa University of Tsukuba, Japan
Traian Marius Truta Northern Kentucky University, USA

Vincenzo Deufemia University of Salerno, Italy

Vitaliy Yakovyna University of Warmia and Mazury in Olsztyn,

Poland

Wojciech Macyna Wrocław University of Technology, Poland Yan Zhu Southwest Jiaotong University, China Yang-Sae Moon Kangwon National University, South Korea

### **External Reviewers**

A. K. M. Tauhidul Islam Informatica, USA

Amna Rizvi University of Sydney, Australia

Anand Kumar Amazon, USA

Andre Kashliev Eastern Michigan University, USA

Davide Costa Altilia.ai, Italy

Eleftherios Kalogeros Ionian University, Greece

Fayçal Saidani Université Mouloud Mammeri de Tizi Ouzou,

Algeria

Feng Yu Youngstown State University, USA

Francesco Granata Altilia.ai, Italy

Gautier Filardo Centre de Recherche de la Gendarmerie

Nationale, France

Luca De Grandis Altilia, Italy

Maryam Mozaffari Free University of Bozen-Bolzano, Italy

Matthew Damigos Ionian University, Greece
Muhammad Umair University of Sydney, Australia
Thilina Lokuruge University of Sydney, Australia

Vinu Venugopal International Institute of Information Technology,

Bangalore, India

# **Organisers**





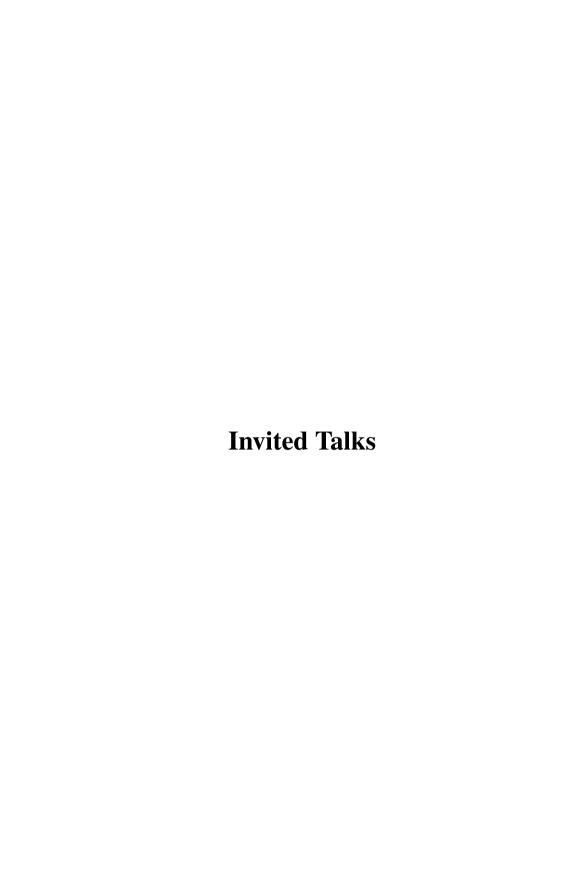


# From Data Silos to Data Mesh: A Case Study in Financial Data Architecture (Industrial Talk)

#### Mariusz Sienkiewicz

Director of Supervisory Data Analysis Center, Polish Financial Supervision Authority, Poland

Abstract. Successful data analytics implementation requires seamless access to both data and related metadata. In many organizations, analytics challenges arise from Data Silos, which impede cross-functional access to data and knowledge sharing across the organization. This talk presents practical insights from a data architecture transformation project conducted at a large institution with over 1,400 employees and overseeing over 2,000 market entities. The organization faced significant analytical and operational challenges due to the presence of Data Silos—isolated repositories associated with specific business areas. To address these limitations, the institution initiated a transition to a Data Mesh architecture to improve data availability and enhance analytical capabilities. This talk explains the rationale behind the persistence of silos, evaluates alternative architectural models, and justifies the choice of Data Mesh based on organizational context. Key elements of the transformation include developing a data management framework, implementing a data catalog, creating a data lake to provide data input flexibility, and establishing a common analytics platform based on Data Domains. While the project is still ongoing, the talk describes the methods being implemented and shares early results, key learnings, and practical recommendations for institutions undertaking similar architectural transitions.



# Blending Contextual Data with Heterogeneous Time Dimensions for Improved Time Series Analysis

### Anton Dignös

Free University of Bozen-Bolzano, Italy

Abstract. In modern industrial settings, sensors continuously generate vast amounts of time series data critical for automation and process optimization. However, analyzing this data in isolation limits its effectiveness, as it often lacks integration with contextual factors that influence outcomes but are not directly observable. While traditional data fusion techniques aim at combining multimodal data such as images or videos, contextual factors in industrial environments frequently differ not in modality but in temporal structure. We identify four distinct time dimensions - constant, time series, events, and intervals - that commonly characterize contextual data in these settings. By transforming diverse time structures into a unified format, we enable the application of conventional machine learning techniques, enhancing the depth and accuracy of industrial data analysis. This talk presents a case study and initial work on a foundational approach for systematically integrating such temporally heterogeneous contextual factors into time series analysis.

# A Hybrid Data Model to Support Transportation Analytics of Emergency Service Vehicles

#### Carson K. Leung

University of Manitoba, Winnipeg, Canada

Abstract. Using a single type of database solution to support real-world applications is becoming more and more challenging because of the volume and variety of data. For instance, the data collected for the transportation industry comprise both structured and unstructured data. Using solely a single type of database solution—relational database system-only or graph database-only—to store and manage data can be challenging. As real-world applications ask even more complex questions related to data, the database solution should be able to facilitate answering these questions in a reasonable time. Hence, in this talk, I present a hybrid model, which integrates data to support transportation analytics. The model consists of relational databases and non-relational databases (namely, graph databases), pooling their strengths to support the demands of the modern application. I also demonstrate this hybrid data model as a practical solution with a case study on improving emergency services—such as emergency medical services (EMS)—response times by having the support of the presented platform.

## Contents - Part I

**Industrial Keynote** 

From Data Silos to Data Mesh: A Case Study in Financial Data Architecture Mariusz Sienkiewicz	3
Invited Talk	
Blending Contextual Data with Heterogeneous Time Dimensions for Improved Time Series Analysis	23
Saifullah Burero, Anton Dignös, Jerry W. Sangma, and Johann Gamper	
A Hybrid Data Model to Support Transportation Analytics of Emergency Service Vehicles	35
Large Language Models	
Automated Archival Descriptions with Federated Intelligence of LLMs  Jinghua Groppe, Andreas Marquet, Annabel Walz, and Sven Groppe	53
Entropy-Guided Probing for Predicting LLM Hallucinations with Knowledge Graph Features  Ushtar Ali, Steven Lynden, Akiyoshi Matono, and Toshiyuki Amagasa	68
Towards Automating RDF Extraction for Archaeological Knowledge	

Graphs with LLMs .....

Processing and Large Language Models .....

Table Annotation Utilizing Large Language Model and Knowledge Graph . . . . .

Ali Hariri, Stéphane Jean, and Mickaël Baron

Ying Zhang and Mizuho Iwaihara

Ontology-Based Forest Fire Management Using Complex Event

Improving Software Security Through a LLM-Based Vulnerability

Syeda Sadia Alam, Mst Shapna Akter, and Alfredo Cuzzocrea

Ritesh Chandra, Sonali Agarwal, and Sadhana Tiwari

83

98

113

122

SysResolve: Study on In-Context LLM Generation of Resolution Scripts  Harsh Borse, Utkalika Satpathy, Mainack Mondal, and Bivas Mitra	130
Data Quality	
A Novel Unsupervised Anomaly Detection Method Based on TCN-LSTM-CMA Autoencoder	139
Behaviour Modelling and Wayfinding Error Detection in Low Mountain Hiking	154
Explainable Time Series Anomaly Detection by Dynamic Mode Decomposition	169
Exploring Quantum Bootstrap Sampling for AQP Error Assessment:  A Pilot Study  Feng Yu and Raya Jahan	184
AI-Driven Semantic Data Quality Assessment and Scoring for Relational Databases  Antony Seabra, Claudio Cavalcante, Nicolaas Ruberg, and Sergio Lifschitz	199
Network Anomaly Detection Using Gramian Angular Field Transformation and Vision Transformer  Jaroslaw Kobiela	207
Machine Learning/Artificial Intelligence Applications	
Identifying Multimodal Sarcasm Based on Incongruous Knowledge Capturing and Contrastive Learning	215
Ensemble ToT and Its Application to Automatic Grading	230
Improving Prompt-Based Learning Framework for Mental Health Aspect Detection from Social Media	245

Contents – Part I	xxii
DInos: A Deep Reinforcement Learning Approach to Generalizable Autoscaling in Stateless Cloud Applications	260
Influential Slot and Tag Selection in Billboard Advertisement  Dildar Ali, Suman Banerjee, and Yamuna Prasad	270
Speech-Scenario Generation Based on the Philosophy of a Prominent Leader Within a Small Community	291
VarCGAN: Variational Cyclic Generative Adversarial Network For Music Genre Style Transfer  Pooja Singh, Dhruv Mishra, and Ankita Khandelwal	301
Innovative Framework for Early Estimation of Mental Disorder Scores to Enable Timely Interventions	322
A Hybrid Approach to Estimating AI Carbon Emissions	329
A Data Product Classification by Technical and Machine Learning Aspects Laura Schuiki, Ulf Schreier, Holger Schwarz, and Bernhard Mitschang	338
Classification Techniques	
Discovering Voting Power for Ensemble Methods	347
Classifying Public and Private Documents Using Context-Based Predictions  Abrar Hasin Kamal and Anne V. D. M. Kayem	363
Author Index	381

# **Contents – Part II**

image Processing, Analytics, and vision Systems	
Relationship Analysis of Image-Text Pair in SNS Posts  Takuto Nabeoka, Yijun Duan, and Qiang Ma	3
Enhancing Segmentation of Irregular Microstructural Elements Using Extended Channel Information and Transfer Learning  Łukasz Marcjan, Sandra Gajoch, Dorota Wilk-Kołodziejczyk, Marcin Małysza, Krzysztof Jaśkowiec, and Grzegorz Gumienny	19
Deep-RVT: A Residual Vision Transformers for Human Action Recognition Sayda Elmi, Morris Bell, and Sai Karthik Navuluru	34
Recommender Techniques	
Food Recommendation With Balancing Comfort and Curiosity Yuto Sakai and Qiang Ma	51
ONFOODS: A Substitute Recommendation System in Food Recipes  Maryam Mozaffari, Anton Dignös, Oswald Lanz, Dominik Matt, Gabriele Pasetti Monizza, Matthias Gauly, and Johann Gamper	66
Inspire Me with Your Questions: Repurposing Historical Questions for New Documents  Yifan Liu, Yixuan Cao, and Ping Luo	80
Data Integration	
MRF-JOIN: Differentially Private Vertical Data Synthesis via Federated Marginal Join on Shared Attributes  Marin Matsumoto, Tsubasa Takahashi, Shun Takagi, and Masato Oguchi	99
Efficient Source Selection for Federated SPARQL Queries Using Adjacent Predicate Information	115
Yudai Ogura, Tadashi Masuda, and Toshiyuki Amagasa	113

# xxvi Contents – Part II

via Multi-stage Cascading Information Fusion  Jianwei Zhang, Shota Sato, Yuta Sasaki, and Yuhki Shiraishi	130
Unified Schema-Driven Graph Polystore: Achieving Transparency in Multi-model Integration and Migration	136
Optimisation Methods	
Group Trip Planning Query Problem with Multimodal Journey	147
A Model-Based Approach for Simple Construction and Efficient  Evaluation of Dataframes	163
Energy and Performance Evaluation of Serverless and Serverful Models on Spark for Database Join Operations  Phan-An-Truong Tran, Laurent D'orazio, Thuong-Cang Phan, and Le Gruenwald	178
Graph Applications	
The Missing Link: Joint Legal Citation Prediction Using Heterogeneous Graph Enrichment  Lorenz Wendlinger, Simon Alexander Nonn, Abdullah Al Zubaer, and Michael Granitzer	197
Graph Patterns in Fine-Grained Access Control for Graph-Structured Data Daniel Schmid, Aya Mohamed, Dagmar Auer, Bahara Muradi, and Josef Küng	212
An Efficient Point-of-Interest Placement Method Based on Betweenness Centrality Ryuta Shiraishi, Ryusei Ohtani, Yuko Sakurai, and Satoshi Oyama	228
Analytics	
Analytics Modelling over Multiple Datasets Using Vector Embeddings  Andreas Loizou and Dimitrios Tsoumakos	237

Contents – Part II	XXV11
Towards IoT-Based Smart Mobility Framework for Proactive Road Stress  Detection in Individuals with ASD	254
A Divisive Unsupervised Feature Selection Approach for Explainable Remaining Useful Life Prediction	270
Data Storytelling to Unlock the Communicative Power of Digital Twins  Faten El Outa, Hugo Breuillard, and Guillaume Dechambenoit	287
Queueing Theory for Verifying the Utilization Rate of an Image Processing System  Jaqueline Donin Noleto Noleto, Thiago Germano do Nascimento, and Pedro Henrique Malheiros Costa Martins	293
Effect of Frequency Features of ELA Maps on the Detection Performance of Image Manipulation Based on DCT and FFT Basis Features	299
ALPHA: A Multi-Attention Enhanced YOLO Framework for Robust Photovoltaic Defect Detection  Bechir Ben Tekfa, Amira Mouakher, and Naeem Ayoub	305
Security/Privacy	
Secure Approach for Blockchain-Based Anonymous Attribute-Based Searchable Encryption Scheme for Data Sharing  Dhruv Kalambe, Nish Shah, Payal Chaudhari, and Priyanshi Manglani	313
Incremental k-Anonymization for Continuously Growing Big Databases Akifumi Kurumatani, Hiromasa Yoshimoto, and Kazuo Goda	329
Post Quantum Cryptographic Schemes and Libraries Selection	338
Benchmarks and Surveys	
Workload-Based Clustering of Large Number of Database-as-a-Service Instances	347

# xxviii Contents - Part II

Accelerating Python Code with Parallel I/O	360
Benchmarking Embedding Techniques for Modeling User Navigation Behavior on Task-Oriented Software  Ikram Boukharouba, Florence Sèdes, Benoit Verhaeghe, and Christophe Bortolaso	367
The Wrecking SQL Incremental Validation Methodology  Ruanitto Docini, Eduardo C. de Almeida, and Luiz S. Oliveira	376
A Survey of Control Technologies for Autonomous Underwater Vehicles Janette Christin Kaspar	383
Author Index	389