Curriculum Vitae[a] - Dr Fabio Mercorio

CONTACT
INFORMATION

University of Milano - Bicocca
Viale dell'Innovazione 10, Room 2042 - U7 Building -
Milan - Italy
Date of Birth: November 6th, 1983
Nationality: Italian

Voice: (+39) 320 8581712
it.linkedin.com/in/mercorio

E-mail: fabio.mercorio@unimib.it
web: www.crisp-org.it/mercorio

CURRENT
POSITION

Assistant Professor   **10/2016 - present**

Department of Statistics and Quantitative Methods, **University of Milan - Bicocca**

– Research Topics: Knowledge Discovery in Databases, Data Analysis and Data Quality, Graph-Databases, AI Planning and Model-checking

Research Collaborator   **09/2010 - present**

CRISP Research Centre, **University of Milan - Bicocca**

– Research Topics: Labour Market Intelligence, BI and Big Data Analytics, Data Analytics

EMPLOYMENT
HISTORY

Visiting Researcher   **09/2015 - 08/2016**

King's College London, **Department of Informatics**

– Research Topics: Heuristic Search Algorithms, Pattern Databases

Research Fellow in Computer Science   **11/2011 - 09/2016**

Department of Statistics and Quantitative Methods / CRISP Research Centre, **University of Milan - Bicocca**

– PI: Prof. Mario Mezzanzanica
– Research Topics: Data Management and Knowledge Discovery, Data Quality and Cleansing, AI Planning and Model-checking

EDUCATION

Ph.D. in Computer Science   **03/2012**

At **Department of Computer Science** , **University of L'Aquila**

– Supervisor: Dr. Giuseppe Della Penna
– Co-Supervisor: Dr. Daniele Magazzeni (King's College London)
– *Thesis:* Model Checking for the Analysis and Control of Complex and Non-Deterministic Systems http://hdl.handle.net/10281/36432

Master degree in Computer Science   **10/2008**

At **Department of Computer Science** , **University of L'Aquila**

– Supervisor: Dr. Giuseppe Della Penna
– Co-Supervisor: Dr. Daniele Magazzeni (King's College London)
– Thesis title: Un Planner per domini PDDL+ basato su tecniche di Model Checking *(A PDDL+ planner based on Model Checking Techniques)*
– Score: 110/110 *cum laude*

   – **National Scientific Qualification** for the position of Associate Professor in Computer Science, March 2018

   – **FFABR Grants for research productivity** provided by Italian Ministry of research Finanziamento annuale individuale delle attivita base di ricerca [Grants provided on a competitive basis aimed at funding research activities]

   – **Best paper award** at the Third International Workshop, Human Computer Interaction - Knowledge Discovery, see [33][17]

   – **Best paper award** at the Third International Conference on Data Technologies and Applications, see [32],[16]

   – Scholarship for Academic Excellence (Abruzzo Region, Italy), from 2002 until 2008

   – Italian Ministry of University and Research, PhD Fellowship, from 2009 until 2011

   – CRISP Research Centre, PostDoc Fellowship, from 2011 until 2014

I am currently involved in the following research projects.

Real-time Labour market information on skill requirements.
EU project aiming at collecting and classifying Web Job vacancies on a well-established international standard classifier (ESCO), extracting the requested skills from the data. I am now applying my background and expertise on AI techniques for reasoning with the collected job vacancies (e.g., matching skills from the data with the predefined ones, by defining similarity functions between occupations on the basis of skill similarities, and exploiting graph databases for scaling over million nodes) (2014-2016).

Development of the Observatory of the Lombardy Labour Market.
Granted by Lombardy Region. Automatic Cleaning of tens million records from Administrative DBs through both AI and Model Checking techniques. The project, as a challenge, has required to create and handle huge datasets (e.g., the cleansed instances of the source dataset) through the realisation of an automatic and scalable model that has been then directly connected to the external database for retrieving up to million records. (2011-2014).

Smart City: The Decision Theater.
Granted by Miur (Italian Ministry of Research) lead by SAS Institute in collaboration with several big ICT companies (e.g., Italtel, Telecom Italia, Selex Sistemi Integrati). The project aims at developing a platform of services for making PAs smarter. He contributed to the proposal definition phase. The project is expected to start 2016 until 2018.

Reviewing for Scientific Journals

   – Knowledge-Based Systems - Elsevier.
   – Applied Soft Computing - Elsevier.
   – AICommunications - IOSpress.
   – Future Generation Computer Systems - Springer.

Program Committee Membership

   – PC member of the The 28th International Conference on Automated Planning and Scheduling (ICAPS-18)
   – PC member of the 40th AAAI-18 Main track + Student Program + Doctoral Consortium
   – PC member of the 32nd ACM SAC 2018
   – PC member of the 7th International Conference on Data Technology and Application (DATA 2018)

- PC member of the The 27th International Conference on Automated Planning and Scheduling, Pittsburgh, USA, (ICAPS-17)
- PC member of the Workshops Program Committee for 25th International Joint Conference on Artificial Intelligence (IJCAI-17)
- PC member of the 39th AAAI-17 (Student Program PC member)
- PC member of the 31th ACM SAC-SWA 2017
- PC member of the 6th International Conference on Data Technology and Application (DATA 2017)
- PC member of the 22nd European Conference on Artificial Intelligence (ECAI-16)
- PC member of the 25th International Joint Conference on Artificial Intelligence (IJCAI-16)
- PC member of the 26th International Conference on Automated Planning and Scheduling (ICAPS-16)
- PC member of the Thirtieth International Conference on Artificial Intelligence (AAAI 2016) Student Abstract and Poster
- PC member of the 31th ACM/SIGAPP Symposium On Applied Computing (ACM SAC 2016)
- PC member of the 5th International Conference on Data Technology and Application (DATA 2016)
- PC member of IMMM the Sixth International Conference on Advances in Information Mining and Management (IMMM 2016)
- PC member of the 30th ACM/SIGAPP Symposium On Applied Computing (ACM SAC 2015)
- PC member of the twenty-ninth International Conference on Artificial Intelligence (AAAI 2015) Student Abstract and Poster
- PC member of KomIS - the special session on Knowledge Discovery meets Information Systems, 2015
- PC member of IMMM the International Conference on Advances in Information Mining and Management (IMMM 2015)
- PC member of the ICAPS Workshop on Model Checking and Automated Planning (MOCHAP 2015)
- PC member of the 6th Italian Planning and Scheduling Workshop (IPS 2015)
- PC member of SoData 2015 - The First Summer School on BI and Big Data Analytics (Capri, Italy, June 10-12 2015)
- PC member of the 4th International Conference on Data Technology and Application (DATA 2015)
- PC member of the 3rd International Conference on Data Technology and Application (DATA 2014)
- PC member of KomIS - the special session on Knowledge Discovery meets Information Systems, 2014
- PC member of the ICAPS Workshop on Model Checking and Automated Planning (MOCHAP 2014)
- PC member of the 4th International Conference on Advances in Information Mining and Management (IMMM 2014)
- PC member of the 5th Italian Planning and Scheduling Workshop (IPS 2013)

Event Organisation
- Co-chair of soDATA - Summer School on BI and Big Data Analytics, June 2017 at Capri Island (Italy) (TbA)
- Co-chair of KomIS - Knowledge Discovery meets Information Systems, Special session at DATA 2017 (TbA)
- Co-chair of ICAPS-16 Doctoral Consortium - June 12th 2016 London (UK)
- Co-chair of soDATA - Summer School on BI and Big Data Analytics, June 10-12 2015 at Capri Island (Italy)
- Co-chair of KomIS - Knowledge Discovery meets Information Systems, Special session at DATA 2015
- Co-chair of KomIS - Knowledge Discovery meets Information Systems, Special session at DATA 2014

Research Collaboration
He established some fruitful research collaborations with the following organisations/centres/research teams

  – Multimedia Information System group, lab ITEM-CINI University of Napoli "Federico II", headed by prof. Antonio Picariello (Since 2016 to present)
  – L'Aquila Model Checking and Planning Group, University of L'Aquila, headed by Prof. Giuseppe Della Penna (Since 2009 to present)
  – CRISP Research Centre, University of Milan-Bicocca, headed by Prof. Mario Mezzanzanica (Since 2010 to present)
  – Information Retrieval Lab, University of Milan-Bicocca, headed by Prof. Gabriella Pasi (Since 2014 to present)
  – AI Planning Group, King's College London, headed by Prof. Daniele Magazzeni (Since 2014 to present)
  – The Big Data Group - The European Network on Regional Labour Market Monitoring (Since 2015 to present)

TEACHING AS LECTURER

Current teaching as lecturer. Each course was assessed by projects, written examinations and oral examinations.

2015/2016. University of Milano-Bicocca
  - Module on *SQL and RDBMs.*
    Responsible for the whole course (6 lectures - 24 hours)
    Master on Data Management for Clinical Research.
  - Module on *SQL, Data Integration and tools for ETL.*
    Responsible for the whole course. *scheduled for the second half 2016.*
    Master on Business Intelligence and Big Data Analytics.
  - Module on *Graph Databases (Lab.)*
    Responsible for the whole course. *scheduled for the second half 2016*
    Master on Business Intelligence and Big Data Analytics.
  - Module on *Tools for Data Quality and Cleaning.*
    Responsible for half od the course. *scheduled for the second half 2016*
    Master on Business Intelligence and Big Data Analytics.

(PAST)

I designed and taught the following courses:

2014/2015. University of Milano-Bicocca
  - Module on *SQL and RDBMs.*
    Master on Data Management for Clinical Research.
    Responsible for the whole course (6 lectures - 24 hours)
  - Module on *SQL, Data Integration and tools for ETL.*
    Responsible for the whole course (10 lectures - 40 hours).
    Master on Business Intelligence and Big Data Analytics.
  - Module on *Graph Databases (Lab.)*
    Responsible for the whole course (5 lectures - 18 hours).
    Master on Business Intelligence and Big Data Analytics.
  - Module on *Tools for Data Quality and Cleaning.*
    Responsible for half od the course (5 lectures - 20 hours)
    Master on Business Intelligence and Big Data Analytics.

2013/2014. University of Milano-Bicocca
  - Module on *SQL and RDBMs.*
    Responsible for the whole course (6 lectures - 24 hours)
    Master on Data Management for Clinical Research.
  - Module on *SQL and Data Quality*
    Responsible for half od course (9 lectures - 30 hours).
    Master on Business Intelligence and Decision Support Systems.

2012/2013. University of Milano-Bicocca
 - Module on *SQL and RDBMs.*
   Responsible for the whole course (6 lectures - 24 hours)
   Master on Data Management for Clinical Research.
 - Module on *SQL and Data Quality*
   Responsible for half od course (9 lectures - 30 hours).
   Master on Business Intelligence and Decision Support Systems.

2011/2012. University of Milano-Bicocca
 - Module on *SQL and Data Quality*
   Responsible for half od course (9 lectures - 30 hours).
   Master on Business Intelligence and Decision Support Systems.

TEACHING AS
TEACHING
ASSISTANT

Current teaching as Teaching Assistant

2015/2016. University of Milano-Bicocca, Dept. of Statistics and Quantiative Methods
 - Bachelor Module on *Informatics - Python.*
   Class taught by Prof. Dario Pescini. (12 lectures - 24 hours).
 - Bachelor Module on *Lab. Informatics - Python.*
   Class taught by Dr. Mirko Cesarini. (11 lectures - 32 hours).

(PAST)

I collaborated in teaching the following courses:

2014/2015. University of Milano-Bicocca, Dept. of Statistics and Quantiative Methods
 - Bachelor Module on *Informatics - Python.*
   Class taught by Prof. Mario Mezzanzanica. (12 lectures - 24 hours).
 - Bachelor Module on *Lab. Informatics - Python.*
   Class taught by Dr. Mirko Cesarini. (11 lectures - 32 hours).

2013/2014. University of Milano-Bicocca, Dept. of Statistics and Quantiative Methods
 - Bachelor Module on *Informatics - Python.*
   Class taught by Prof. Mario Mezzanzanica. (12 lectures - 24 hours).
 - Bachelor Module on *Lab. Informatics - Python.*
   Class taught by Dr. Mirko Cesarini. (11 lectures - 32 hours).

2012/2013. University of Milano-Bicocca, Dept. of Statistics and Quantiative Methods
 - Bachelor Module on *Informatics - Python.*
   Class taught by Prof. Mario Mezzanzanica. (12 lectures - 24 hours).
 - Bachelor Module on *Lab. Informatics - Python.*
   Class taught by Dr. Mirko Cesarini. (11 lectures - 32 hours).

2011/2012. University of Milano-Bicocca, Dept. of Statistics and Quantiative Methods
 - Bachelor Module on *Informatics - Python.*
   Class taught by Prof. Mario Mezzanzanica. (12 lectures - 24 hours).
 - Bachelor Module on *Lab. Informatics - Python.*
   Class taught by Dr. Mirko Cesarini. (11 lectures - 32 hours).

2010/2011. University of L'Aquila, Dept. of Computer Science
 - MSc Module on *Formal Methods for Verification of Complex Systems*
   Class taught by Dr. Daniele Magazzeni. (2 lectures - 8 hours).

2009/2010. University of L'Aquila, Dept. of Computer Science
 - MSc Module on *Formal Methods for Verification of Complex Systems*
   Class taught by Dr. Daniele Magazzeni. (2 lectures - 8 hours).

Co-supervised theses:

– Aurora Zeller-Celso. *Smart mobility and services innovation: an overview of the mobility in Milan* (jointly supervised with prof. Mario Mezzanzanica). March 2015
– Fulvio Pucci. *Graph Databases: Application on Twitter: the case or Expo2015* (jointly supervised with prof. Mario Mezzanzanica). January 2015
– Giulia Galbiati. *Data Quality Indicators and Techniques for the Analysis of Labour Market Data Flows* (Jointly supervised with Prof. Mario Mezzanzanica). September 2013.
– Nicola Pesenti. *Validation of a Data Quality Process* (Jointly supervised with Prof. Mario Mezzanzanica). March 2012.
– Francesco Scoccia. *Planning with Genetic Algorithms* (Jointly supervised with Dr. Giuseppe Della Penna). December 2010.
– Francesco Antonelli. *Automatic Control of a Building Lifting System* (Jointly supervised with Dr. Giuseppe Della Penna). March 2010.
– Luigi Grassano. *A model-checking based approach for Planning Activities of a Batch Chemical Production Plant* (Jointly supervised with Dr. Giuseppe Della Penna). October 2009.

Artificial Intelligence: domain-independent planning, temporal continuous planning, planning in mixed discrete-continuous domains, planning in hybrid domains.

Knowledge Discovery: Data Quality and Cleansing, Business Intelligence, Graph Databases, Machine Learning, Information Systems.

Formal Methods: model checking techniques, HW/SW verification, control theory, graph algorithms.

My research interests are mainly in **Artificial Intelligence Planning** and **Knowledge Discovery**, both addressed in my PhD Thesis [43], with particular attention to the application of Data Analysis and Control through Formal Methods. In all my research activities, I've started to give my contribution to the respective communities by publishing on top tier International Conferences (IJCAI, ICAPS, IDA) and Journals (Information Processing & Management, ACM Journal of Data and Information Quality, AICommunications, Applied Intelligence). My contributions as well as a description of my current research activities are sum- marised in the following.

Artificial Intelligence Planning

I contributed in addressing **planning** and **universal planning** for systems with **continuous change** through model-checking techniques. Indeed, many concrete planning problems in which planning can be applied involve a mixture of continuous (possibly nonlinear) and discrete behaviours, also known as *hybrid systems*. To deal with such kind of systems, recently the AI Planning Community has defined the PDDL+ language which is also able to capture continuous processes and events.
As a first step, I was co-developer of the UPMurphi planner, as presented in [42], which performs planning by exploiting the planning-as-model-checking paradigm. Then, I defined and implemented a formal mapping between the PDDL+ and UPMurphi semantics, then applying UPMurphi to a number of real-world planning applications (see [13, 14],[39, 41, 40] for details).
The main issue in planning through model checking is the state explosion phenomenon as the region of reachable states is often huge. To mitigate this problem, I worked to enrich the UPMurphi planner with a disk-based algorithm, namely the V-UPMurphi tool (see [43],[13] for details) which allows it to use the disk for storing both the system dynamics and the final solution, using OBDDs to compress the resulting plans through the C++ CUDD library.
My research interest also investigates the planning in **non-deterministic domains**. Indeed, while in classical planning a plan always reaches a goal, in non-deterministic domains an action may have several outcomes, which could make the plan unable to reach a goal. The concept

of **strong plan** fills this gap by ensuring that a plan always achieves a goal regardless of the non-determinism. In this field, I have been working on the development and analysis (proving correctness, completeness and complexity) of an algorithm which, as a novel contribution, (1) considers a $\mathbb{R}^+$ cost function looking for cost-optimal strong plans; (2) extends the synthesis of strong planning also to hybrid and/or nonlinear domains (see [37], for details). I implemented the cost-optimal strong planning algorithm on the top of V-UPMurphi tool, namely the SUPMurphi tool, applying it to some motivating non-deterministic domains [8]. A complete summary of this part of my work has been published in [12].

In 2016, as visiting researcher at Kings College London, I have been involved in the realisation of DiNo, the first heuristic domain independent planner - built on top of UPMurphi - able to cope with the full PDDL+ semantics [24, 23].

Knowledge Discovery.

Business and governmental applications, web applications, ongoing relations between citizens and public administrations generate a lot of data, many of which can be considered as *weakly-structured data* (i.e., data whose evolution can be expressed as a pathway on a graph). Such data are often used in several decision making activities as evaluation of active policies, resource allocation, service design and improvement.

[**Data Quality and Analysis**]. Given my background on Model Checking, as a first contribution I addressed the problem of the **data quality verification** of databases as a **model checking problem**, indeed the former composes one of the main step of the KDD process. In this settings, a solution for the latter (if any) represents a data quality problem for the former. I extended the application of model checking by defining the Robust Data Quality Analysis, a technique to support and improve existing cleansing activities, applying it to the *CO System* of the Italian Labour Market Domain studied at the CRISP Research Centre (see [38, 34],[11]). This work has been awarded as best paper at [33] while an extended version appears in IPM journal [9]. Since Model Checking is well suited to explore huge state spaces, our approach has been recently used for performing a sensitivity analysis on statistical indicators on the basis of the cleansing procedure applied (see, e.g. [36], [7]).

[**Data Cleaning**]. Recently, I inherited the concept of AI universal planning by defining the *universal cleanser*: a repository of *all the feasible* actions able to cleanse a dirty data. It is automatically synthesised off-line on the basis of the UPMurphi algorithm (see[18] for an overview). As a characteristic, it is *domain-dependent*, i.e. it can deal only with database data conforming to the model that generated it, but *data-independent* since, once computed, it can be used to cleanse any dataset conforms to the model [35]. More recently, the idea of Universal Cleansing has been expressed as a pure planning problem within the AI Community [34, 30][10]. On the other hands, I investigated that an efficient modelling of the data evolution as a pathway on a graph allows using AI planning techniques for improving the data cleansing accuracy, as shown in [32]. This work has been awarded as best paper ad DATA 2014 international conference. In 2015 I started working on an International research project aimed at classifying Web Job vacancies over an international and well-established classification system for retrieving both occupation descriptions and skills (preliminary results in [25]).

[**Big Data for Decision Making**]. In 2015 Ive started working on an International research project on *Big Data for Labour Market Intelligence*, that requires to deal with all the 4 Vs of the Big Data Framework: namely, the Volume due to the huge Web dataset sizes; the Velocity due to the pace at which data are generated, the Variety as consequence of the use of dierent Web source data models, and Veracity caused by the quality issues that arise from the source datasets). Here, the research tender granted by the EU Agency for European Vocational and Educational Agency (CEDEFOP) aimed at classifying Web Job advertisements over an international and well-established classification system for retrieving both occupation descriptions and skills. Specifically, I have been involved on classification task (i.e., use machine learning to classify job advertisements) and the data model as well, i.e., building a knowledge base through the Neo4j graph database to reason over order of hundreds million nodes, then extracting skills from the data using semantic similarity metrics (preliminary results in [6],[19, 25, 21, 26]).

[**Towards the Market.**] Finally, as a researcher on AI at CRISP research center, since 2015 Ive been working on WollyBI (`http://www.wollybi.com/en/`): A commercial solution that exploits Machine Learning and semantic analysis techniques to perform Web labour market data analysis.

RESEARCH
SOFTWARE
DEVELOPED

I was co-developer and co-maintainer of the following research tools:

- **UPMurphi [a C++ model checking based universal planner]** The Universal Planner Murphi is a tool for planning and universal planning for linear and nonlinear continuous PDDL+ models with processes and events. UPMurphi is based on the Discretise and Validate approach [42] and it has been designed to automatically interface with the VAL plan validator. The tool translates PDDL+ models into UPMurphi models and then outputs a collection of optimal plans according to the PDDL+ syntax. Moreover, UPMurphi is able to validate (using the standard PDDL+ VAL validator) all the generated plans. *Related Publications:[43],[42, 39, 40, 41].* Download from GitHub
- **V-UPMurphi [handles huge state spaces through a disk-based algorithm]** It is a computational engine to enhance the ability of UPMurphi to synthesise plans and universal plans for (possibly nonlinear) Discrete Time Hybrid Systems defined with a PDDL+ model. V-UPMurphi exploits the use of the disk during the exploration of the dynamics. It implements on disk the state space reduction techniques inherited from Murphi (as bit-compression, hash-compaction) and allows one to pause and resume the system analysis process, using the disk to both explore and store the graph of the system dynamics. *Related Publications:[12][43]*
- **SUPMurphi [implements the strong planning algorithm dealing with non-deterministic systems]** It uses the Murphi description language to model non-deterministic behaviour of systems' actions and, thanks to the use of disk-based algorithm, it is able to address strong planning problems. Then, it allows one to validate the final strong plan (if any) on the system graph. *Related Publications:[37][12][43]*
- **DiNo [The first heuristic PDDL+ planner].** It is built on top of UPMurphi and exploits the PDDL+ to UPMurphi mapping developed during the PhD studies (see, [42], [13]). As a novel characteristics, DiNo defines and implements a new heuristic function: the Temporal Relaxed Planning Graph (SRPG), that allows the synthesis of heuristic plan in a more efficient way wrt UPMurphi. *Notice: As visiting researcher, I have been involved in the realisation of DiNo by Planning Group researchers of King's College London researchers Related Publications:[23, 24, 22].* Download from GitHub
- **GraphDBLP [The first GraphDatabase for Computer Scientists].** It is the graph-database instance of the well-known DBLP, the Computer Science Digital Library. It employs Neo4j as a baseline for building a researchers knowledge graph. It allows performing graph-traversal queries looking for communities, research topics and shortest paths among researchers. It also uses word embeddings to model similarities between research topics. *Related Publications:[4].*

He is involved in the realisation of the following commercial solutions:

- **WollyBI:** `www.wollybi.com` A commercial tool that exploits machine learning for classifying multilingual Web Job vacancies (gathered from web sources) over the standard ESCO taxonomy, thus supporting the decision making activities of several labour market stakeholders.

FOREIGN
LANGUAGES

– Italian (mother tongue)
– English (proficient user - C1)

*   *Journal Articles*

[1] Flora Amato, Aniello Castiglione, Fabio Mercorio, Mario Mezzanzanica, Vincenzo Moscato, Antonio Picariello, and Giancarlo Sperl. Multimedia story creation on social networks. *Future Generation Computer Systems (to appear)*, pages –, 2018.

[2] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. Classifying online job advertisements through machine learning. *Future Generation Computer Systems (to appear)*, 2018.

[3] Lovaglio, Cesarini, Mercorio, and Mezzanzanica. Skills in demand for ict and statistical occupations: Evidence from web-based job vacancies. *Statistical Analysis and Data Mining*, 11(2):78–91, 2018.

[4] Mario Mezzanzanica, Fabio Mercorio, Mirko Cesarini, Vincenzo Moscato, and Antonio Picariello. Graphdblp: a system for analysing networks of computer scientists through graph databases. *Multimedia Tools and Applications*, pages 1–32, 2018.

[5] Sperli', Amato, Mercorio, Mezzanzanica, Moscato, and Picariello. A social media recommender system. *International Journal of Multimedia Data Engineering and Management (IJMDEM)*, 9(1):36–50, 2018.

[6] Roberto Boselli, Mirko Cesarini, Stefania Marrara, Fabio Mercorio, Mario Mezzanzanica, Gabriella Pasi, and Marco Viviani. Wolmis: a labor market intelligence system for classifying web job vacancies. *Journal of Intelligent Information Systems*, Sep 2017.

[7] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and Fabio Mercorio. A model-based approach for developing data cleansing solutions. *The ACM Journal of Data and Information Quality*, 5(4):1–28, March 2015. ISSN 1936-1955.

[8] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, and Fabio Mercorio. Synthesis of cost-optimal strong plans in non-deterministic domains. *Journal on Artificial Intelligence Tools*, 24(6), 2015. ISSN 0218-2130.

[9] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and Fabio Mercorio. A model-based evaluation of data quality activities in KDD. *Information Processing & Management*, 51(2):144–166, 2015. ISSN 0306-4573.

[10] Roberto Boselli, Mario Mezzanzanica, Mirko Cesarini, and Fabio Mercorio. Towards data cleansing via planning. *Intelligenza Artificiale*, 8(1), 2014. ISSN 1724-8035.

[11] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and Fabio Mercorio. Longitudinal data consistency verification using formal methods. *International Journal of Information Quality*, 3(3):185–206, 2014. ISSN 1751-0457.

[12] Fabio Mercorio. Model checking for universal planning in deterministic and non-deterministic domains. *AI Communications*, 26(2):257–259, 2013. ISSN 0921-7126.

[13] Giuseppe Della Penna, Daniele Magazzeni, and Fabio Mercorio. A universal planning system for hybrid domains. *Applied Intelligence*, 36(4):932–959, 2012. ISSN 0924-669X.

[14] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, and Fabio Mercorio. Resource-optimal planning for an autonomous planetary vehicle. *International Journal of Artificial Intelligence & Applications (IJAIA)*, 1(3):15–29, 2010. ISSN 0976-2191.

*   *Book Chapters*

[15] Mario Mezzanzanica and Fabio Mercorio. Big data enables labor market intelligence. In *Encyclopedia of Big Data Technologies*, pages 1–11. Springer International Publishing, Cham, 2018.

[16] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. Accurate data cleansing through model checking and machine learning techniques. In Markus

Helfert, Andreas Holzinger, Orlando Belo, and Chiara Francalanci, editors, *Data Management Technologies and Applications*, volume 178 of *Communications in Computer and Information Science*, pages 62–80. Springer International Publishing, 2015. ISBN 978-3-319-25935-2.

[17] Roberto Boselli, Mario Mezzanzanica, Mirko Cesarini, and Fabio Mercorio. A policy-based cleansing and integration framework for labour and helthcare data. In *Knowledge Discovery and Data Mining, LNCS 8401*, pages 141–168. Springer, 2014. ISBN 978-3-662-43967-8.

*   *Referred Conference Papers*

[18] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. An AI planning system for data cleaning. In *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2017, Skopje, Macedonia, September 18-22, 2017, Proceedings, Part III*, pages 349–353, 2017.

[19] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. Using machine learning for labour market intelligence. In *Machine Learning and Knowledge Discovery in Databases - European Conference, ECML PKDD 2017, Skopje, Macedonia, September 18-22, 2017, Proceedings, Part III*, volume 10536 of *Lecture Notes in Computer Science*, pages 330–342. Springer, 2017.

[20] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, Mario Mezzanzanica, and Alessandro Vaccarino. A pipeline for multimedia twitter analysis through graph databases: Preliminary results. In *DATA 2017 - the International Conference on Data Technologies and Applications*, 2017.

[21] Stefania Marrara, Gabriella Pasi, Marco Viviani, Mirko Cesarini, Fabio Mercorio, Mario Mezzanzanica, and Marco Pappagallo. A language modelling approach for discovering novel labour market occupations from the web. In *2017 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2017)*, pages 1026–1034, 2017.

[22] Wiktor Piotrowski, Maria Fox, Derek Long, Daniele Magazzeni, and Fabio Mercorio. PDDL+ planning with temporal pattern databases. In *The AAAI-17 Workshop on Symbolic Inference and Optimization (SymInfOpt-17)*, 2017.

[23] Wiktor Mateusz Piotrowski, Maria Fox, Derek Long, Daniele Magazzeni, and Fabio Mercorio. Heuristic planning for hybrid systems. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February 12-17, 2016, Phoenix, Arizona, USA.*, pages 4254–4255, 2016.

[24] Wiktor Mateusz Piotrowski, Maria Fox, Derek Long, Daniele Magazzeni, and Fabio Mercorio. Heuristic planning for PDDL+ domains. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI 2016, New York, NY, USA, 9-15 July 2016. (An earlier version of this paper was presented at AAAI-16 Workshop on Planning for Hybrid Systems - PlanHS-16)*, pages 3213–3219, 2016.

[25] Flora Amato, Roberto Boselli, Mirko Cesarini, Fabio Mercorio, Mario Mezzanzanica, Vincenzo Moscato, Fabio Persia, and Antonio Picariello. Challenge: Processing web texts for classifying job offers. In *Semantic Computing (ICSC), 2015 IEEE International Conference on*, pages 460–463, Feb 2015. ISBN 9781479979356.

[26] Flora Amato, Roberto Boselli, Mirko Cesarini, Fabio Mercorio, Mario Mezzanzanica, Vincenzo Moscato, Fabio Persia, and Antonio Picariello. Classification of web job advertisements: A case study. In *SEBD 2015 - The 23rd Italian Symposium on Advanced Database Systems*, pages 144–151, 2015. ISBN 9781510810877.

[27] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. Applying the ahp to smart mobility services: A case study. In *Proceedings of the 2nd Special Session in Knowledge Discovery meets Information Systems (KomIS) at DATA 2015*, 2015. ISBN 978-989-758-103-8.

[28] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, and Fabio Mercorio. UP-Murphi released: PDDL+ planning for hybrid systems. In *Proceedings of the 2nd Workshop on Model Checking and Automated Planning (MOCHAP-2015)*, pages 35–39, 2015.

[29] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. Data quality on kdd: a real-life scenario. In *SEBD 2014 - The 22nd Italian Symposium on Advanced Database Systems*, pages 378–385, 2014. ISBN 9781634391450.

[30] Roberto Boselli, Mario Mezzanzanica, Mirko Cesarini, and Fabio Mercorio. Planning meets data cleansing. In *The 24th International Conference on Automated Planning and Scheduling (ICAPS 2014)*, pages 439–443. AAAI, 2014. ISBN 978-1-57735-660-8.

[31] Daniele Magazzeni, Fabio Mercorio, Balbir Barn, Tony Clark, Franco Raimondi, and Vinay Kulkarni. Temporal planning for business process optimisation. In *Proceedings of the ICAPS-2014 Workshop on Scheduling and Planning Applications (SPARK 2014)*, 2014.

[32] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and Fabio Mercorio. Improving data cleansing accuracy: A model-based approach. In *DATA 2014 - the International Conference on Data Technologies and Applications (**best paper awarded**)*. SciTePress, 2014. ISBN 978-989-758-035-2.

[33] Roberto Boselli, Mirko Cesarini, Fabio Mercorio, and Mario Mezzanzanica. Inconsistency knowledge discovery for longitudinal data management: A model-based approach. In *SouthCHI13 special session on Human-Computer Interaction & Knowledge Discovery, Lecture Notes in Computer Science, vol. 7947 (**Best paper awarded**)*. Springer, 2013. ISBN 978-3-642-39145-3.

[34] Roberto Boselli, Mario Mezzanzanica, Mirko Cesarini, and Fabio Mercorio. Can planning meet data cleansing? In *The 5th Italian Workshop on Planning and Scheduling at AIxIA2013*, pages 63–66, 2013.

[35] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and FabioMercorio. Automatic synthesis of data cleansing activities. In *DATA 2013 - the International Conference on Data Technologies and Applications*, pages 138–149. SciTePress, 2013. ISBN 978-989-8565-67-9.

[36] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and Fabio Mercorio. Data quality sensitivity analysis on aggregate indicators. In Markus Helfert, Chiara Francalanci, and Joaquim Filipe, editors, *DATA 2012 - the International Conference on Data Technologies and Applications*, pages 97–108. SciTePress, 2012. ISBN 978-989-8565-18-1.

[37] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, Fabio Mercorio, and Enrico Tronci. Cost-optimal strong planning in non-deterministic domains. In *Proceedings of the 8th International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, pages 56–66. SciTePress, 2011. ISBN 978-989-8425-74-4.

[38] Mario Mezzanzanica, Roberto Boselli, Mirko Cesarini, and Fabio Mercorio. Data quality through model checking techniques. In *Intelligent Data Analysis (IDA), Lecture Notes in Computer Science vol. 7014*, pages 270–281. Springer, 2011. ISBN 978-3-642-24799-6.

[39] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, and Fabio Mercorio. A PDDL+ benchmark problem: The batch chemical plant. In *Proceedings of the The 20th International Conference on Automated Planning and Scheduling (ICAPS 2010)*, pages 222–225, Toronto, Canada, 2010. AAAI Press. 978-1-57735-449-9.

[40] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, and Fabio Mercorio. Planning for autonomous planetary vehicles. In *Proceedings of the The Sixth International Conference on Autonomic and Autonomous Systems*, pages 131–136, Cancun, Mexico, 2010. IEEE. ISBN 978-1-4244-5915-5.

[41] Fabio Mercorio. Planning for continuous domains. In *The AI*IA Doctoral Consortium, Brescia (Italy) December 1-3*, 2010.

[42] Giuseppe Della Penna, Benedetto Intrigila, Daniele Magazzeni, and Fabio Mercorio. UP-Murphi: a tool for universal planning on PDDL+ problems. In *Proceedings of the 19th International Conference on Automated Planning and Scheduling (ICAPS 2009)*, pages 106–113, Thessaloniki, Greece, September 2009. AAAI Press. ISBN 978-1-57735-406-2.

[43] Fabio Mercorio. *Model Checking for the Analysis and Control of Complex and Non-deterministic Systems.* PhD thesis, Department of Computer Science - University of L'Aquila - Italy (public available at http://hdl.handle.net/10281/36432), 2012.

This document is composed of 12 pages.