



WILEY

# Interdisciplinary Reviews

DATA MINING AND  
KNOWLEDGE DISCOVERY

## A Guide for Authors

Welcome from the Editors in Chief .....	1
Contact information .....	1
Authors' checklist .....	2
1. Background: the Wiley Interdisciplinary Reviews.....	3
2. WIREs Data Mining and Knowledge Discovery .....	3
3. Editorial Board .....	4
4. Preparing your manuscript .....	5
5. Online submission instructions .....	10
6. What happens to your manuscript.....	10
Appendix I: Article list by subject category.....	12
Appendix II: Copyright Permission Request Form.....	17

## WELCOME FROM THE EDITORS IN CHIEF

I am delighted to welcome you to the circle of authors of *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery (WIREs DMKD)*. This is an important and timely endeavor that will do much to assist the communication of high-quality information in our field.

As Editor in Chief, my goals are to maximize the quality and relevance of the publication's content and to ensure, with the help of the international advisory board (see page 4), that we continue to receive world-class submissions that will keep *WIREs DMKD* in the forefront of the field. Ultimately, our success depends on our authors. I urge you, therefore, to study the guidelines in this document, follow the prescribed procedures for manuscript preparation and delivery, and adhere as closely as possible to your delivery deadlines.

*WIREs DMKD* is an online serial publication that will evolve into a fully integrated, dynamic reference source. As the major topics are fleshed out with your contributions and those of other notable scholars, John Wiley & Sons intends to capture this content in a comprehensive print reference work, the *Wiley Encyclopedia of Data Mining and Knowledge Discovery*. We believe this novel dual-format publication will be an exciting and important contributor to research and scholarly discourse in our field.

Thank you again for agreeing to participate. I look forward to a long and fruitful relationship.

Witold Pedrycz  
Editor in Chief  
*Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*

## CONTACT INFORMATION

All correspondence relating to manuscript preparation, proofs and other related matters should be addressed to:

**Cassie Strickland**  
Associate Editor  
John Wiley & Sons, Inc.

Phone: 619.269.6636  
E-mail: [dmkd@wiley.com](mailto:dmkd@wiley.com)

For questions concerning the scope of your contribution or other issues of scientific substance, please contact:

**Professor Witold Pedrycz**  
Editor in Chief  
Electrical Engineering Department  
University of Alberta  
Alberta T6G2V4  
Canada

E-mail: [pedrycz@ee.ualberta.ca](mailto:pedrycz@ee.ualberta.ca)

## AUTHORS' CHECKLIST

### AS YOU PREPARE TO WRITE YOUR ARTICLE

Have you:

- studied the instructions in your letter of acceptance confirmation concerning the type, topic, and scope of your article?
- taken note of the specified delivery date for your article? (It is important that you adhere to this, or let us know if you will need more time.)
- studied the complete article list (see Appendix I) to get a sense of the overall scope of the journal?
- studied the step-by-step instructions in this document (see Section 4) concerning the preparation of your article?
- studied the detailed guidelines in this document on style and format of figures, references, etc. (see Sections 4.5 and 4.6)?
- downloaded the WIREs article template and read the instructions in its margins?

### AS YOU PREPARE TO SUBMIT YOUR ARTICLE

Have you:

- reviewed the online submission instructions (see Section 5)?
- included a brief Abstract, Related Articles, and Further Reading/Resources list?
- included numbered in-text citations for the references?
- provided bibliographic references in the specified format outlined in this Guide (see Section 4.5.6)?
- provided figures as separate files rather than embed them in your article template (see Section 4.6)?
- clearly numbered and labeled all illustrations and figures, and provided legends for each?
- obtained and sent to us all necessary permissions (including permission for electronic publication) for any tables, figures, or text reproduced from previously published material? To acquire permissions, use the *Permission Request Form* (Appendix II of this guide) or request permission from registered publishers online, via the Copyright Clearance Center's "Rightslink" page (<http://www.copyright.com>); please keep copies of all correspondence for your records.
- provided the names and contact information (including e-mail addresses) for 4–5 potential reviewers you feel are familiar with your field and could review your article?
- filled out the Contributor Agreement (sent in the letter of acceptance confirmation), printed, signed, and scanned it for uploading during submission? (To fax or mail, please contact [dmkd@wiley.com](mailto:dmkd@wiley.com).)

## 1. BACKGROUND: THE WILEY INTERDISCIPLINARY REVIEWS

The *Wiley Interdisciplinary Reviews* (WIREs) represent a major new publishing franchise for John Wiley & Sons. Their editorial goal is to emphasize the importance of interdisciplinarity in science and to support cross-disciplinary collaborative efforts in research and education. The WIREs are not journals *per se*; rather, they are hybrid publications that combine the most powerful features of traditional reference works and review journals in a compelling format designed to exploit the full potential of the online medium. The key features of the WIREs are as follows.

- They focus on high-profile, well-funded research areas at the interfaces of the traditional disciplines.
- The WIREs emphasize collaborative and integrative approaches to scientific research, presenting cutting-edge science from a multidisciplinary perspective.
- They will be launched as serial publications so that they can benefit from full abstracting and indexing and, especially, impact factors.
- Meticulous conceptual development of each of the WIREs will result in a highly structured, comprehensive coverage of the field in question.
- The WIREs will adopt a common “templated” editorial format and structure that will maximize quality and consistency within and between the works.
- To raise their visibility and drive online usage, the WIREs will initially be made available *free of charge* to institutional subscribers (and hence to individual end-users at academic, government, and corporate institutions).
- This unique editorial concept is designed to promote author participation, long-term loyalty, and community spirit in interdisciplinary research.

## 2. WIRES DATA MINING AND KNOWLEDGE DISCOVERY

### 2.1 AIMS OF THE PUBLICATION

*Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* is a major new scientific publication that will support the information needs of researchers in this field and help to shape its future development. Its goals are to

- present the current state of the art of Data Mining and Knowledge Discovery through an ongoing series of commissioned reviews written by leading researchers
- capture the crucial interdisciplinary flavor of this field by including articles that address the key topics from the differing perspectives of data mining and knowledge discovery including application areas in technology, business, healthcare, education, government and society and culture
- capture the rapid development of Data Mining and Knowledge Development through a systematic program of content updates
- encourage new participation in this field by presenting its achievements and challenges in an accessible way to a broad audience.

*WIREs DMKD* will be fully indexed in the major abstracting services, and will be assigned an impact factor in the same way as a journal. Unlike the review journal literature, however, *WIREs DMKD* will offer a comprehensive, coherent, well-structured coverage of the field. It will also be updated in a systematic fashion so that its content remains as current as possible.

## 2.2 STRUCTURE AND SCOPE

The following top-level category structure is proposed for *WIREs Data Mining and Knowledge Discovery*:

**Algorithmic Development**  
**Application Areas**  
**Commercial, Legal and Ethical Issues**  
**Fundamental Concepts of Data and Knowledge**  
**Technologies**

The Editors have developed the *WIREs DMKD* article list based on this structure. For a current list of article titles, please see Appendix I on Page 12.

## 2.3 READERSHIP

*WIREs DMKD* will be designed in such a way that different subsets of the content will be useful to upper-level undergraduate and postgraduate students, to teaching and research professors in academic programs, to scientists and research managers in industry; moreover, *WIREs DMKD* will include review and background information useful to scientists entering the field of Data Mining and Knowledge Discovery.

## 3. EDITORIAL BOARD

### 3.1 EDITOR IN CHIEF

***Witold Pedrycz, Professor and Director of Computer Engineering at the University of Alberta, Canada***

Witold Pedrycz has served since 1998 as Professor and Director of Computer Engineering at the University of Alberta, Canada. His research interests include pattern recognition, data mining, and granular computing as well as fuzzy systems engineering, neural networks and evolutionary computing. He has authored eleven research monographs and published numerous journal articles and conference papers in these fields. Pedrycz is a coeditor of the *Handbook of Fuzzy Computation* and the *Handbook of Data Mining and Knowledge Discovery* (both OUP) and coauthor of *Fuzzy Systems Engineering* and *Handbook of Granular Computing* (both Wiley). He currently serves as an Associate Editor of *IEEE Transactions on Systems, Man and Cybernetics*, *IEEE Transactions on Neural Networks*, and *IEEE Transactions on Fuzzy Systems*. He is also an Editor-in-Chief of *Information Sciences*. He is a past President of IFSA and past President of NAFIPS. Pedrycz is a Fellow of IEEE and a Fellow of IFSA.

### 3.2 EDITORIAL ADVISORY BOARD

**Andrzej Cichockin** – Japan: RIKEN Brain Science Institute  
**Oscar Cordon** – Spain: European Center for Soft Computing  
**Frank Hoffmann** – Germany: Technische Universität Dortmund  
**Mehmed Kantardzic** – USA: University of Louisville  
**Rudolf Kruse** – Germany: Otto-von-Guericke Universität Magdeburg  
**Boris Mirkin** – United Kingdom: Birkbeck College  
**Sushmita Mitra** – India: Indian Statistical Institute  
**Peter Lenk** – University of Michigan

## 4. PREPARING YOUR MANUSCRIPT

### 4.1 REVIEW YOUR DESIGNATED ARTICLE TOPIC AND SCOPE DESCRIPTION

Please review your designated topic and any additional scope notes provided in your letter of invitation. If you have any questions or concerns about the topic or scope, please contact us immediately for further advice. One month before your article is due, we will send an email asking you to provide us with an outline for your article. This is to ensure that your ideas for the scope of the article are in line with the editors'.

### 4.2 STUDY THE COMPLETE ARTICLE LIST

Appendix I is a current list of planned article titles/subjects, grouped by topic/subtopic. Please study this list to understand the overall scope of *WIREs DMKD* and to see how your article fits into the larger organizational scheme.

### 4.3 VERIFY YOUR DESIGNATED ARTICLE TYPE

*WIREs DMKD* articles will be assigned a specific article type based on their intended level and readership. Each article type will have an associated template that specifies its length, overall structure, and style; authors will be required to use the appropriate template when preparing their articles. The following article types are included in *WIREs DMKD*:

---

#### OPINIONS

Opinions provide a forum for thought-leaders, hand-picked by the editors, to provide a more individual perspective on the field in question.

**Average extent = 2,000-4,000 words, ≤ 5 figures/tables, 30-60 references, ~5 pages.**

---

#### OVERVIEWS

Overviews will provide a broad and relatively non-technical treatment of important topics at a level suitable for advanced students and for researchers without a strong background in the field.

**Average extent = 5,000-8,000 words, 10-16 figures/tables, 50-100 references, 10-14 pages.**

---

#### ADVANCED REVIEWS

Advanced Reviews, aimed at researchers and advanced students with a strong background in the subject, will review key areas of research in a citation-rich format similar to that of leading review journals.

**Average extent = 4,000-6,000 words, ≤ 10 figures/tables, 50-75 references, ~9 pages.**

---

#### FOCUS ARTICLES

Focus articles are short articles, sometimes included within a larger article, that describe specific real-world issues, examples, implementations, etc. These articles will be technical in nature.

**Average extent = 2,500-4,000 words, ≤ 7 figures/tables, 40-60 references, ~5 pages.**

## 4.4 USE THE TEMPLATE

A Word and LaTeX template for writing your article are available in your ScholarOne Manuscripts Author Center. Please use one of these templates to prepare your article as this will help ensure that it will meet the necessary requirements. In the Word template, consult the 'Comment' notes in the right-hand margin for an explanation of each section in the article. If at any time you do not see these Comments, select View → Print Layout. Further help is available by emailing [dmkd@wiley.com](mailto:dmkd@wiley.com).

## 4.5 REFER TO THE FOLLOWING INSTRUCTIONS ON TEXT FORMAT AND STYLE

### 4.5.1 VERSIONS OF THE MANUSCRIPT

Text should be in DOC (preferred) or RTF format; TeX/LaTeX files are also acceptable.

The manuscript you submit should be the final version that you wish to be sent for peer review. Please do not send partial versions or drafts, unless we specifically request that you do so.

### 4.5.2 ABSTRACT

Please provide us with a concise ( $\leq 250$  words) abstract for your article. This is particularly important for the online version of the publication. Note that you will also be asked to copy the abstract into ScholarOne/Manuscript Central, the online article submission system, for use in corresponding with potential peer reviewers.

### 4.5.3 HEADING STRUCTURE

A maximum of three levels of headings will be used in *WIREs DMKD*. Do not use acronyms in headings. Do not use the heading *Introduction*. The material that immediately follows the Abstract is the introduction, without a heading. Do use each level of heading uniformly throughout the manuscript. Do create headings that:

1. help the reader find information quickly;
2. are descriptive yet specific;
3. are compatible in phrasing and style; and
4. are concise (less than 50 characters).

### 4.5.4 SIDEBARS

If any are appropriate, you are encouraged to include sidebars ("boxed" information that is relevant to but separate from the main text), especially to highlight contemporary interdisciplinary themes. Each sidebar should be a maximum of 250 words. Do not include more than two sidebars.

### 4.5.5 NOTES

While footnotes are generally discouraged in scientific publications, the template will allow you to create them. You should use these notes sparingly and make every effort to include necessary information in the main body of text.

#### 4.5.6 REFERENCE LIST

The reference list at the end of your article provides substantiating references for factual statements in your main text. Please follow these guidelines:

- Each article cited in the reference list must have an accompanying numerical in-text citation (e.g. “It has been demonstrated experimentally<sup>5</sup> that...”). The references should be numbered and listed in the order in which they are cited in the text, *not* in alphabetical order.
- Please format your reference list in accordance with the style used to cite references in PubMed (see examples below).
- For journal articles where there are more than ten authors of a paper, list the first ten authors’ names followed by *et al.* For each author listed, supply all initials and last name.
- For a book, give the complete title, edition (if other than first), the publisher and city, the year of publication, and the page number(s). Please also give inclusive pagination (include the first and last page numbers of a paper), to indicate to the reader whether the paper is a long or a short one.
- References in the bibliography should be numbered.
- When citing a specific passage, table, or figure in a print reference, give the exact page number where the item appears; for example, “Reference 11, p. 196.”
- Should you wish to use Endnote, please follow the style of BioMed Central journals or *Epilepsy Currents*.

Many publishers are now assigning a unique Digital Object Identifier (DOI) to journal articles, book chapters, etc. This identifier, once assigned, persists through the lifetime of the object and can be used to find the object on the Internet, even if it is moved to another URL. See the URL [www.doi.org](http://www.doi.org) for more information. If a DOI is assigned to the material that you cite, include the DOI in the reference.

Some sample references in the correct format follow:

##### SAMPLE REFERENCE FOR AN ARTICLE IN A JOURNAL

Rebetez M. Public expectation as an element of human perception of climate change *Climatic Change* 2006, 32:495-509.

##### SAMPLE REFERENCE FOR AN ARTICLE IN AN EDITED BOOK

Meyer WB. Americans and their weather. Oxford University Press, Oxford, 2000, 278.

##### SAMPLE REFERENCE FOR AN ARTICLE IN A CONFERENCE PROCEEDINGS

Boswell DA, Andrews EB. Managing Urbanization. In Harris, F, ed. *Global Environmental Issues*. John Wiley & Sons, Chichester; 2004 153-196.

##### SAMPLE REFERENCE FOR A WEBSITE

<http://www.cdc.noaa.gov/ENSO/enso.current.html> (accessed July 3 2008)



---

#### 4.5.7 FIGURE CAPTIONS/TABLES

Figure captions and tables should be included at the end of your manuscript. Do not embed them within the text, and do not submit them as separate files. You will also be asked to copy the figure captions into the ScholarOne/Manuscript Central online submission system when you load the image files for your figures.

---

#### 4.5.8 FURTHER READING/RESOURCES

Many readers will use *WIRES DMKD* as a first reference source to orient themselves in the subject. The purpose of the further reading/resources list is to provide additional sources of general information, should they wish to inquire further. If you are aware of books, articles, or academic websites that may be useful to the reader and do not appear in your reference list, you may add a separate list of these after the reference list. Recent journal articles may also be included if they are not too specialized. If you wish, you may comment on the references in this section. References in the further reading/resources list are not to be numbered, but otherwise should follow the style outlined for bibliographic references.

---

#### 4.5.9 RELATED ARTICLES

Related articles are other *WIRES DMKD* articles that will serve as valuable ‘see also’ links for readers. Please select up to 3 article titles from the list provided (Appendix I and also linked to the Word template) and add these to the table at the end of the article template.

---

### 4.6 REFER TO THE FOLLOWING INSTRUCTIONS ON PREPARATION OF FIGURES

It is essential that your artwork be submitted simultaneously with your text and that it be complete and in the correct format. Like the text, artwork goes through several production stages before it is placed onto a page, and it travels separately from the text. **Upload each figure as a separate image file.** Some detailed preparation guidelines are given below.

---

#### 4.6.1 FIGURE PREPARATION AND FORMATTING

- All artwork should be viewable through a freely available plug-in.
- Cite all figures in the text (e.g., “see Fig. 1”) and save each figure separately. Do not embed figures in the text or include captions in figures. Label figures using 8-point Helvetica regular or Times Roman fonts; for subscripts and superscripts use a 6-point font. Use of full-color figures is encouraged, and you will *not* be charged for their use.
- Create line drawings using Illustrator, Freehand, Canvas, CorelDraw, or Adobe Photoshop, and save as an EPS or TIF file. **Line drawings must be a minimum resolution of 600 dpi.** The following programs are *not* acceptable as art-rendering tools: MacDraw, PowerPoint, ClarisWorks, Harvard Graphics, Freelance, Persuasion, Paintbrush, Micrografix Drawing, Word, Excel, FrameMaker, or Ventura.
- Halftones (photographs) should be saved as TIF files **with a minimum resolution of 300 dpi.** For color figures, convert from RGB to CMYK. The following file formats are *not* acceptable: BMP, WMF, JPEG, and GIF.

- For screen captures, select RGB or Index mode; capture at 100% of image size. Save image as a PCX, TIF, or BMP file, with a minimum resolution of 72 or 96 dpi.

*Submitting figure files that do not conform to these standards may delay publication of your article.*

#### 4.6.2 ART SIZING

Be consistent in the use of line weight and type style. All figures should be of proportionate size with one another. If figures must be reduced, remember that the type will also be reduced, so it should be created proportionately so that it is clearly readable at the smaller size. The ideal final type size is 8 to 9 points.

#### 4.6.3 LINE WEIGHTS

Use 0.5 to 0.75 point as the basic line weight for drawing, with heavier or lighter line weights for emphasis and clarity according to the following scheme (do not use "hairline" rules, as they do not reproduce well):

- Graph axes: 0.5 point (also used for boxes in flow diagrams and to show flow or direction lines)
- Graph borders: 0.75 point.

### 4.7 ACQUIRE PERMISSION FOR ALL PREVIOUSLY PUBLISHED MATERIALS

You must have permission to use any material from a copyrighted source. Redrawing an illustration is not enough—even if you use someone’s illustration only as a basis for your own, you must obtain permission to make a new version. *It is your responsibility as the author to obtain permission to use copyrighted materials in your article.* You must either write to the copyright owner with a request for permission (using the *Permission Request Form*, Appendix II of this Guide) or apply for permission online via the Copyright Clearance Center’s “Rightslink” page (<http://www.copyright.com>). Make sure to apply for print *and* electronic rights. Please let us know if you are unable to obtain permission for electronic delivery. When only a small amount of material is to be used, granting of permission is usually a formality, but publishers need to know how much will appear. A few publishers charge a fee for granting such permission; this fee can usually be negotiated, but payment of the fee is your responsibility. *John Wiley & Sons will not pay permissions fees on your behalf.* If you do not wish to pay a particular permissions fee, you may prefer to modify your article so that the copyrighted material is not required.

Please include the necessary credit lines in the appropriate places in your article and send us the completed permission request forms when you submit your manuscript. The language should be the exact language used by the copyright owner, or, if nothing is specified, should include the title, author’s name, previous publisher, and the date of copyright. Since credit lines are part of the manuscript, it is your responsibility to supply them. Credit lines will often accompany figures and illustrations; they should be included at the appropriate place in the figure legend or text.

### 4.8 REFER TO THE FOLLOWING INSTRUCTIONS REGARDING MULTIMEDIA ITEMS

You may wish to include multimedia elements (such as video clips or animations) in your article. The following general rules apply:

- The item should be in a popular media format that can be played with freely available software on both PCs and Macs (e.g., Quicktime (\*.mov) or generic Windows media player formats (\*.wmv or \*.mpg)).
- The file size should be limited to <10 megabytes.
- The item must of adequate quality for web publication.
- The item must genuinely add scientific or educational value to the material.

All items must have a caption and, where necessary, a permission statement backed up by a formal permission letter. Multimedia materials cannot be uploaded on the ScholarOne/Manuscript Central submission site, but must be emailed separately to [dmkd@wiley.com](mailto:dmkd@wiley.com).

## 5. ONLINE SUBMISSION INSTRUCTIONS

Articles commissioned for *WIREs DMKD* will be submitted and peer-reviewed using the ScholarOne Manuscripts article management system. When you were invited to contribute an article, a user account was created for you. If you do not know your User ID and/or Password, you can use the 'Password Help' function on the log in page at <http://mc.manuscriptcentral.com/dmkd>. To submit your manuscript online, log in and click on the 'Author Center' icon to begin your submission. Detailed instructions on using ScholarOne Manuscripts can be found by clicking the 'User Tutorials' button in the 'Resources' box on the right-hand side of the log in page.

**Please be sure to study the instructions given at the site carefully**, and then let the system guide you through the submission process. You will be able to exit and re-enter the system at any stage before finally submitting your work. All submissions are kept strictly confidential. You can log in periodically and check your Author Center to monitor the movement of your manuscript through the review process.

## 6. WHAT HAPPENS TO YOUR MANUSCRIPT

### 6.1 REVIEW AND REVISION

Our editorial process starts with the review of your manuscript by qualified experts in the field and by the *WIREs DMKD* editors. We will pass reviewers' comments on to you and request that you make any necessary changes. In addition, we will discuss with you any editorial changes that may be necessary.

In some cases the reviewers will have no comments, and the manuscript will be processed for composition. In other cases, reviewers' and our own editorial comments will give rise to further correspondence. The Editors reserve the right to cut, to request more information from the author, to revise, and in some cases, to add publicly available material. If the Editors' changes are extensive, you will have an opportunity to review the changes before the manuscript is sent for production.

### 6.2 GALLEY PROOFS

In due course you will receive the galley proofs. By this point we will have discussed any major changes with you, but smaller changes are sometimes made in the editorial office as we proceed. At this stage, you may still make any corrections of typographical errors that are necessary (but not changes in style or content). Changes

at the galley stage are costly and are properly reserved for correction of printer's errors. Any changes in galleys beyond corrections of typographical errors will be made at the discretion of the editor and may be overruled for the sake of expediencies of schedule and expense.

There will be hundreds of contributing authors, and any contributor who returns proofs late has the potential to delay the publication of the entire work. Please proofread carefully, and give the job a high priority. We do not usually send page proofs to authors (page proofs come in quickly from the typesetter and must be cleared in short order), so the galley proof will be your only opportunity to make corrections.

### 6.3 AUTHOR SERVICES

Once your article goes into production, you will be eligible for Wiley-Blackwell Author Services. Benefits include:

- the ability to track your manuscript through the production process to publication,
- email notification when your article is in Early View,
- free access to your article for yourself, corresponding authors, and up to 10 colleagues, and
- a preferential author's discount of 25% on Wiley books.

For more information and to register, please go to <http://authorservices.wiley.com>.

## APPENDIX I: ARTICLE LIST BY SUBJECT CATEGORY

### ALGORITHMIC DEVELOPMENT

#### **Association Rules**

- 001: Mining uncertain data
- 002: A priori algorithm and its extensions
- 003: FP algorithm
- 004: Association rule mining for regression problems
- 005: Evolutionary and metaheuristics-based approaches for association rules extraction and optimization
- 006: Distributed association rules
- 007: Parallel and distributed computing for cyber security
- 008: Privacy preserving distributed association rule mining
- 009: Incremental algorithms of associative rules
- 010: KD interestingness measures based on unexpectedness
- 011: State of the art in privacy-preserving data mining
- 014: Association rule hiding methods
- 015: High-accuracy privacy-preserving mining
- 016: Sequence mining
- 017: Time-constrained sequence mining
- 018: Constraints in sequence

#### **Biological Data Mining**

- 020: Mining biological data
- 021: Sequence analysis in biological data mining
- 022: Alignment
- 023: Genomics in the context of biological data mining
- 024: Analyzing genomic data
- 027: Structural genomics
- 026: Protein DNA interactions
- 025: Biomedical informatics and pharmacogenomics
- 028: Host-virus interaction
- 029: Effect of post-transcriptional RNA or genomic approaches for transcription
- 030: Proteomics in the context of biological data mining
- 031: Proteomics for biomedical research
- 032: Protein structure prediction and folding
- 033: Drug design
- 034: Docking
- 035: Gene expression and regulation
- 036: Biclustering
- 037: Classification
- 038: Feature selection
- 039: Clustering
- 040: Tutorial on biological networks
- 041: Signal transduction pathway
- 042: Gene interaction/regulation

- 043: Metabolic pathway
- 044: Protein-protein interaction networks
- 045: RNA in the context of data mining
- 046: Non-coding RNA
- 047: Role of gene regulation in disease
- 048: Micro RNA

#### **Hierarchies and Trees**

- 323: Decision trees: fundamentals, lessons learned and state of the art
- 324: Real-world problems solved by decision trees
- 326: A review on the use of classification trees for bioinformatics
- 327: Decision trees and software engineering
- 049: Real world applications of random forests and regression forests
- 050: Random forests for medical applications
- 051: Random forests for bioinformatics applications
- 052: Evolutionary design of decision trees
- 053: Evolutionary design of decision trees for medical applications
- 054: Evolutionary design of fuzzy decision trees

#### **Model Combining**

- 055: Ensemble based systems in decision making
- 056: Algorithm independent techniques
- 057: Ensemble methods in machine learning
- 058: Bagging and boosting
- 059: Multiple classifier fusion
- 060: Bagging and random forests
- 061: Boosting
- 062: Boosting and associative classifiers
- 063: Stacked generalization
- 064: Mixture of experts
- 065: Mixture of experts for case-based reasoning
- 066: Multiple model regression estimation

#### **Multimedia**

- 067: Video data mining
- 068: Speech data mining
- 069: Image-driven data mining

#### **Spatial and Temporal Data Mining**

- 070: Temporal data mining
- 071: Temporal pattern mining
- 072: Time series mining
- 073: Data stream mining
- 074: Concept drift and change detection
- 075: Spatial data mining

#### **Structure Discovery**

- 077: State-of-the-art on objective-function based clustering

078: Cluster validation  
 079: Clustering with supervision mechanisms  
 080: State-of-the-art on clustering large data  
 081: Clustering ensembles  
 082: Clustering distributed data  
 083: Privacy preserving  
 084: Ubiquitous data mining  
 085: Coclustering/biclustering  
 086: K-Means clustering  
 087: Spectral clustering  
 088: Scalable clustering  
 089: Variable selection in clustering  
 090: Clustering temporal data  
 091: Evolutionary clustering  
 092: Mining time-series data  
 093: Particle swarm and ant colony  
 094: Genetic algorithms for crisp and fuzzy clustering  
 095: Fuzzy clustering algorithms

#### **APPLICATION AREAS**

##### **Business and Industry**

096: CRM  
 097: CEM  
 098: Marketing strategy and tactics  
 099: Product line and portfolio management  
 100: Sales force management  
 101: Marketing efficiency and effectiveness evaluation  
 102: Incorporating cost and return information into data mining applications  
 103: Cost quantification in classification applications

##### **Data Mining Software Tools**

104: Integrating genetic fuzzy systems in data mining tools  
 105: A software system for spatial data mining  
 106: Data mining with open source software  
 107: Tools for privacy preserving distributed data mining  
 108: Data mining tools  
 109: Brighthouse: an analytic data warehouse for adhoc queries

##### **Education and Learning**

110: Student performance evaluation by mining datastreams  
 111: Use of bayesian networks and other learning algorithms in student evaluation  
 112: Behavioral model-based DM methods for cheating detection  
 113: Query languages and metasearch  
 114: Document clustering using frequent word meaning sequences  
 115: Information retrieval and data mining  
 116: Multilingual text mining

117: Text mining from annual reports  
 118: Text mining and visualization tools

##### **Government and Public Sector**

328: Data mining for homeland security: recent advances  
 329: Homeland security considerations in data mining

##### **Health Care**

119: Scheduling of emergency physicians - data mining approach  
 120: Predictive data mining in clinical medicine  
 121: Temporal reasoning and medical data mining  
 122: Gene mining  
 123: Confidentiality in medical data mining  
 124: Protein data mining  
 125: Text mining for ubiquitous medicine  
 126: Epidemiological studies and data mining  
 127: Data mining and discovery of risk patterns  
 128: Data mining and medical claims  
 129: Data Mining and Left Heart Syndrome  
 130: Medical images and data mining  
 131: Multivariate image mining  
 132: Diagnostic image mining

##### **Science and Technology**

133: Mining of chemical data bases  
 134: Data mining in e-mail forensics  
 135: Data mining in crime scene analysis  
 135: Data mining and crime analysis  
 138: Intrusion detection and cyber security  
 139: Fault detection  
 140: Sensor information management  
 141: Data mining avionics  
 142: Process monitoring and data mining  
 143: Forecasting technology and part obsolescence  
 144: Perceptual optimization of complex visualizations  
 145: Data mining for hierarchical model creation  
 146: Data mining for monitoring dangerous infections  
 147: Data mining of source code to improve bug finding techniques  
 148: Grid computing  
 149: Ambient intelligence

##### **Society and Culture**

330: Applying web usage mining for personalizing hyperlinks in web-based adaptive educational systems  
 331: Personalized online sales using web usage data mining  
 332: Integrating AHP and data mining for product recommendation based on customer lifetime value  
 333: Mining changes in customer buying behavior for collaborative recommendations

## COMMERCIAL, LEGAL AND ETHICAL ISSUES

### **DMKD Regulatory Issues**

- 150: Intellectual property
- 151: Data and results ownership
- 152: Liability risk
- 153: National security controls
- 154: Government disclosure requirements
- 155: Privacy-preserving data mining systems

### **Ethical Considerations**

- 156: Whistle blowing
- 157: Legal uses of DMKD for regrettable ends
- 158: Green technology

### **Social Considerations**

- 159: Privacy and data mining
- 160: Privacy-preserving data mining systems
- 334: An efficient sanitization algorithm for balancing information privacy and knowledge discovery
- 335: Privacy preserving data mining with malicious participants
- 161: Fairness and integrity in data mining
- 162: The privacy interest in seclusion as affected by data mining
- 163: Personal security
- 164: Liberty/convenience

## FUNDAMENTAL CONCEPTS OF DATA AND KNOWLEDGE

### **Data Concepts**

- 171: Methods and techniques for discovering taxonomies of behavioral process models
- 172: Taxonomies for web-based mining
- 173: Taxonomies in sequence mining
- 174: Taxonomy adaptation for group profiling
- 175: Web mining patterns
- 176: Patterns mining on the web
- 177: Mining dynamic patterns
- 178: Data quality mining
- 179: Information enhancement for data mining
- 180: Ubiquitous data stream mining
- 181: Resilient data stream mining
- 182: Distributed data mining in scientific applications
- 183: Parallel data mining on graphics processors

### **Key Design Issues in Data Mining**

- 189: Cost-effective data mining
- 190: Scalable data mining
- 191: Sustainable data mining
- 192: Logical implication and causal dependency
- 193: New strategies for privacy protection on semantic web
- 194: Secure data mining on the grid
- 195: Standards and services for data mining
- 196: Similarity measures for sequential data

### **Knowledge Representation**

- 184: Knowledge-based data mining
- 185: Knowledge discovery and knowledge-based techniques in medicine: towards more relevance oriented data mining
- 186: Background knowledge for intelligent tutoring systems
- 187: Ontologies and pattern bases
- 188: Using multiple ontologies to integrate complex biological data

### **Motivation and Emergence of Data Mining**

- 165: Trends in data mining
- 166: Data mining industry is maturing
- 167: Data mining: challenges and realities with real world data
- 168: Domain driven data mining
- 169: ADMIRE data mining strategy
- 170: Empowering scientific discovery

## TECHNOLOGIES

### **Association Rules**

- 313: Fundamentals of association rules in data mining and knowledge discovery
- 314: Subgroup discovery
- 315: Frequent pattern mining
- 316: Fuzzy association rules
- 317: Software tools for association rule generation
- 318: Statistically sound exploratory pattern discovery
- 319: Methodologies and technologies for rule-based system design
- 320: Learning concepts represented by sets of rules
- 321: Fuzzy rules and associations
- 322: Rule induction for medical applications

### **Classification**

- 306: Feature extraction, document clustering and classification
- 307: Multiway clustering and feature extractions, spectral clustering
- 308: Gene classification, gene expression data processing and exploratory data analysis
- 309: Applications of tensor (multi-way array) factorizations and decompositions in data mining
- 310: Use of probabilistic graphical models for data mining and knowledge discovery
- 311: Multiple classifier systems: fundamentals and methods
- 312: Use of diversity measures for the design of multiclassifiers

### **Computational Intelligence**

- 214: Knowledge discovery
- 215: An overview on the use of neural networks for data mining tasks

216: Self-organizing maps and data mining  
 217: Neural based learning classifiers  
 218: Fuzzy neural networks  
 220: Data sampling  
 221: Fuzzy machine learning and data mining  
 222: Fuzzy pattern mining  
 223: Fuzzy logic and soft computing for intelligent information retrieval  
 224: Fuzzy approaches for web mining  
 225: Knowledge discovery with granular computing  
 226: Fuzzy temporal association rules  
 227: Fuzzy rule based classification systems: models and applications  
 228: Fuzzy approaches for gene expression data mining  
 229: Fuzzy clustering  
 230: Evolutionary algorithms  
 231: Evolutionary computation for data mining and knowledge discovery  
 233: Swarm intelligence applications for data mining and knowledge discovery  
 234: Evolutionary computation and swarm intelligence-based clustering methods  
 235: Evolutionary computation for subgroup discovery  
 236: Evolutionary computation for preprocessing tasks in DMKD  
 237: Evolutionary computation for feature selection and extraction in DMKD  
 238: Multiobjective evolutionary optimization  
 239: Learning classifier systems and DMKD  
 240: Genetic programming in data mining for drug discovery  
 241: Microarray data mining with evolutionary computation  
 242: Knowledge discovery in data mining via an evolutionary algorithm  
 243: Mining association rules from databases with continuous attributes using genetic network programming  
 244: Evolutionary rule induction  
 245: Genetic-fuzzy mining with appropriate choice of fitness functions  
 246: Genetic fuzzy classifiers  
 247: Fitness modeling in evolutionary algorithms  
 248: Genetic fuzzy systems, fuzzy knowledge extraction by evolutionary algorithms  
 249: Evolving fuzzy systems for data streams  
 250: Data mining tools

#### **Computer Architectures for Data Mining**

197: Data mining and life sciences applications on the grid  
 198: Cloud computing and data mining

199: Accelerating data mining workloads: current approaches and future challenges in system architecture design  
 200: High performance data mining  
 201: New computer architecture for data mining  
 202: Rule mining in sensors' networks

#### **Data Preprocessing**

207: Cost-sensitive data preprocessing for CRM  
 208: Missing data and data imputation  
 209: Sampling in relational data mining  
 210: Distance preserving dimensionality reduction  
 211: Relational data preprocessing techniques  
 212: Data fusion and data mining  
 213: Benchmarking attribute selection techniques for discrete class data mining

#### **Statistical Fundamentals**

325: Hierarchically optimal classification tree analysis: applications in medicine and allied health disciplines  
 336: The role of statistical models in data mining  
 251: Bayesian trees  
 252: Classification and regression trees  
 253: Distributions: mechanisms and validation  
 254: Resampling methods  
 255: Smoothing and nonparametric density and regression estimation  
 256: Experimental design  
 257: Generalized linear models  
 258: Linear and logistic regression  
 259: Function data and function estimation  
 260: Model selection and validation  
 261: Penalized maximum likelihood and cross validation  
 262: The role of statistical inference in data mining  
 263: Outlier detection  
 264: Robust statistics  
 265: Network sampling

#### **Structure Discovery and Clustering**

266: Structures  
 267: Criteria  
 268: Methods of hierarchical clustering  
 269: Partitive clustering (K-Means Family)  
 270: Clustering criteria  
 271: Dynamic clustering  
 272: Conceptual clustering  
 273: Clustering mixed data  
 274: Robust clustering  
 275: Mean shift paradigm  
 276: CLARANS family  
 277: Choosing the number of clusters  
 278: High dimensions  
 279: Subspace clustering  
 280: Mixture of distributions



281: Gaussian (number-of-clusters in a gaussian mixture model)  
282: Non-gaussian (including number of clusters)  
283: Density-based  
285: Conceptual  
286: Fuzzy clustering  
287: Rough clustering  
288: Spectral clustering  
289: Consensus  
290: Biclustering  
291: Graph theoretic  
292: Additive clusters  
293: Clustering by Kohonen Network  
297: Ordering  
298: Robinsonian

299: Pagerank and Co  
300: Collaborative filtering  
301: Validation  
302: External  
303: Internal  
304: Ontology-based external validation of clustering  
305: Semi-supervised structure discovery

#### **Visualization**

203: Visual data mining  
204: Visualizing predictors in neuroimaging  
205: Text mining visualization and social media  
206: Collaborative virtual visualization

## APPENDIX II: COPYRIGHT PERMISSION REQUEST FORM



Figure/table no. in  
original manuscript

Figure/table no. in  
*WIREs DMKD*

**DATE:**

**FROM:** Name .....

**TO:** Name .....

Address .....

Address .....

.....

.....

.....

.....

Dear Sir or Madam

I am preparing for publication an original work entitled

**[Article Title]**, by **[Author Name]** (**[Article ID]**)

to be published by John Wiley & Sons Inc., 111 River Street, Hoboken, New Jersey 07030,

which will appear as part of a work entitled: **WILEY INTERDISCIPLINARY REVIEWS: DATA MINING AND KNOWLEDGE DISCOVERY**

edited by **Witold Pedrycz**

I request your permission to include the following in the work referred to above, in all subsequent editions of the work and in all derivative works based on the work, in any and all media of expression now known or later developed and in all languages, to be published by Wiley or its licensees throughout the world

[insert figure/table details here]

The usual form of acknowledgement is to quote the author(s) or photographer and publication title of the original material or source. John Wiley & Sons Inc. will include the words:

“Reproduced by permission of [the owner of the publishing rights] ”

Please would you therefore confirm the name of the copyright holder to be quoted as granting permission.

Please indicate agreement by signing and returning a copy of this letter. In signing, you warrant that you are the sole owner of the rights granted and that your material does not infringe upon the copyright or other rights of anyone. If you do not control these rights, I would appreciate your letting me know to whom I should apply.

Yours sincerely

[Your name]

---

---

We hereby grant permission for the use of the material requested above.

Date ..... Signed .....

Copyright Holder...../ Publisher .....