

# **An Investigation Of Web Crawler Behavior Characterization**

Getting the books **An Investigation Of Web Crawler Behavior Characterization** now is not type of inspiring means. You could not unaccompanied going once books accretion or library or borrowing from your connections to entrance them. This is an extremely simple means to specifically acquire guide by on-line. This online statement **An Investigation Of Web Crawler Behavior Characterization** can be one of the options to accompany you as soon as having extra time.

It will not waste your time. understand me, the e-book will totally atmosphere you additional business to read. Just invest tiny epoch to contact this on-line broadcast **An Investigation Of Web Crawler Behavior Characterization** as competently as review them wherever you are now.

**Covid-19 and Psychological Disorders: Impacts and Therapeutic Interventions  
From Molecular Basis to Social** - Elnara Shafiyeva 2022-09-29

*Frontiers of WWW Research and Development -- APWeb 2006* - Xiaofang Zhou 2006-01-09

This book constitutes the refereed proceedings of the 8th Asia-Pacific Web Conference, APWeb 2006. More than 100 papers cover all current issues on WWW-related technologies and new advanced applications for researchers and practitioners from both academic and industry.

**Security and Privacy in Communication Networks** - Yan Chen 2009-10-13

This book constitutes the thoroughly refereed post-conference proceedings of the 5th International ICST Conference, SecureComm 2009, held in September 2009 in Athens, Greece. The 19 revised full papers and 7 revised short papers were carefully reviewed and selected from 76 submissions. The papers cover various topics such as

wireless network security, network intrusion detection, security and privacy for the general internet, malware and misbehavior, sensor networks, key management, credentials and authentications, as well as secure multicast and emerging technologies.

**Workload Modeling for Computer Systems Performance Evaluation** - Dror G. Feitelson 2015-03-23

A book for experts and practitioners, emphasizing the intuition and reasoning behind definitions and derivations related to evaluating computer systems performance.

Handbook of Research on Pattern Engineering System Development for Big Data Analytics - Tiwari, Vivek 2018-04-20

Due to the growing use of web applications and communication

devices, the use of data has increased throughout various industries. It is necessary to develop new techniques for managing data in order to ensure adequate usage. The Handbook of Research on Pattern Engineering System Development for Big Data Analytics is a critical scholarly resource that examines the incorporation of pattern management in business technologies as well as decision making and prediction process through the use of data management and analysis. Featuring coverage on a broad range of topics such as business intelligence, feature extraction, and data collection, this publication is geared towards professionals, academicians, practitioners, and researchers seeking current research on the development of pattern

management systems for business applications.

*Inspiring Stewardship* - Didier Cossin  
2016-05-10

Transform from leader to steward to drive growth and make an impact  
*Inspiring Stewardship* helps leaders drive sustainable development and growth by balancing short- and long-term objectives amidst competing interests. By examining a range of exceptional individual and organisational stewards alongside insights from quantitative studies, this book unlocks the contributions you can make to enhance your influence and secure your legacy. Profiles of leaders from Ratan Tata to Bill Gates illustrate what stewardship really means, and the attributes and characteristics of steward leaders. Are they

fundamentally different? Where do they get their exceptional drive and resilience? Why do they shoulder such huge responsibility? Risks, rewards, motivations and actions are explored in depth to give you a real-world view of stewardship, and to show you how to embody these principles in your own everyday life. The notion of stewardship has never been more relevant to business than it is today, as managers struggle to balance present and future interests. This book provides clarity, direction and guidance for more purpose-driven action. Understand the concepts and principles of stewardship Discover characteristics of steward leaders across cultures Learn how stewards have shifted the course of their organisations Examine the capabilities that deepen

organisational resilience Today, ownership of firms is increasingly fragmented, investment structures are more and more complex and the time horizons of investors, executives and owners do not always coincide. Inspiring Stewardship helps you break through the tension and drive sustainable growth through adaptability and good leadership.

**Encyclopedia of Criminal Activities and the Deep Web** - Khosrow-Pour D.B.A., Mehdi 2020-02-01

As society continues to rely heavily on technological tools for facilitating business, e-commerce, banking, and communication, among other applications, there has been a significant rise in criminals seeking to exploit these tools for their nefarious gain. Countries all over the world are seeing substantial

increases in identity theft and cyberattacks, as well as illicit transactions, including drug trafficking and human trafficking, being made through the dark web internet. Sex offenders and murderers explore unconventional methods of finding and contacting their victims through Facebook, Instagram, popular dating sites, etc., while pedophiles rely on these channels to obtain information and photographs of children, which are shared on hidden community sites. As criminals continue to harness technological advancements that are outpacing legal and ethical standards, law enforcement and government officials are faced with the challenge of devising new and alternative strategies to identify and apprehend criminals to preserve the safety of

society. The Encyclopedia of Criminal Activities and the Deep Web is a three-volume set that includes comprehensive articles covering multidisciplinary research and expert insights provided by hundreds of leading researchers from 30 countries including the United States, the United Kingdom, Australia, New Zealand, Germany, Finland, South Korea, Malaysia, and more. This comprehensive encyclopedia provides the most diverse findings and new methodologies for monitoring and regulating the use of online tools as well as hidden areas of the internet, including the deep and dark web. Highlighting a wide range of topics such as cyberbullying, online hate speech, and hacktivism, this book will offer strategies for the prediction and prevention of online

criminal activity and examine methods for safeguarding internet users and their data from being tracked or stalked. Due to the techniques and extensive knowledge discussed in this publication it is an invaluable addition for academic and corporate libraries as well as a critical resource for policy makers, law enforcement officials, forensic scientists, criminologists, sociologists, victim advocates, cybersecurity analysts, lawmakers, government officials, industry professionals, academicians, researchers, and students within this field of study.

*ICT Influences on Human Development, Interaction, and Collaboration* -

Chhabra, Susheel 2012-08-31

"This book creates awareness on how ICTs contribute to human development

in multiple areas, including the link between ICTs and economic, social, and political aspects of human development"--Provided by publisher.

*Computer and Information Sciences III*

- Erol Gelenbe 2012-10-29

A collection of papers from ISCIS 27th Annual Symposium. Based on a rigorous selection of worldwide submissions of advanced research papers, this volume includes some of the most recent ideas and technical results in computer systems, computer science, and computer-communication networks. This book provides the reader with a timely access to the work of vibrant research groups in many different areas of the world where the new frontiers of computing and communications are being created.

Current Index to Journals in Education - 1998-04

## Using Computational Intelligence for the Dark Web and Illicit Behavior

Detection - Rawat, Romil 2022-05-06

The Dark Web is a known hub that hosts myriad illegal activities behind the veil of anonymity for its users. For years now, law enforcement has been struggling to track these illicit activities and put them to an end. However, the depth and anonymity of the Dark Web has made these efforts difficult, and as cyber criminals have more advanced technologies available to them, the struggle appears to only have the potential to worsen. Law enforcement and government organizations also have emerging technologies on their side, however. It is essential for these organizations to stay up to date on these emerging technologies, such as computational intelligence,

in order to put a stop to the illicit activities and behaviors presented in the Dark Web. Using Computational Intelligence for the Dark Web and Illicit Behavior Detection presents the emerging technologies and applications of computational intelligence for the law enforcement of the Dark Web. It features analysis into cybercrime data, examples of the application of computational intelligence in the Dark Web, and provides future opportunities for growth in this field. Covering topics such as cyber threat detection, crime prediction, and keyword extraction, this premier reference source is an essential resource for government organizations, law enforcement agencies, non-profit organizations, politicians, computer scientists, researchers, students, and

academicians.

*Network Security Through Data Analysis* - Michael Collins 2014-02-10

In this practical guide, security researcher Michael Collins shows you several techniques and tools for collecting and analyzing network traffic datasets. You'll understand how your network is used, and what actions are necessary to protect and improve it. Divided into three sections, this book examines the process of collecting and organizing data, various tools for analysis, and several different analytic scenarios and techniques.

Communication for Behavior Change - Esta de Fossard 2016-04-01

A valuable resource for creating distance education programs for radio, TV, Internet and mobile, using an entertainment–education approach.

This volume explains how to integrate education into an entertaining format so that it holds people's attention while enabling them to improve their health and lifestyle. It also demonstrates how mobile technologies can be used for educating frontline health workers as well as the communities they serve. This book offers many examples of such programs and provides essential guidance on:

- Understanding precisely what knowledge must be imparted to the communities and how.
- Why mobile technology is the new and most important platform for delivering educational content.
- How educational programs can be designed for developing countries in a sustainable way.

*Communication Technologies, Information Security and Sustainable*



*Development* - Faisal Karim Shaikh  
2014-09-10

This book constitutes the thoroughly refereed proceedings of the Third International Multi-topic Conference on Communications, Technologies, Information Security and Sustainable Development, IMTIC 2013, held in Jamshoro, Pakistan, in December 2013. The 27 revised papers presented in this volume were carefully reviewed and selected from 140 submissions. The topics presented had a reasonable balance between theory and practice in multi-disciplined topics including wireless sensor networks, cloud computing, wireless communication, antenna design, signal processing, software engineering, image processing, bioinformatics and telemedicine, neural networks, automation and control, and green

renewable energy.

**Interdependencies in the Discovery and Adoption of Facebook Applications**

- Philip Mayrhofer 2012-12-13

Facebook Platform is a prominent example of an internet-based service for which third-party developers can offer small add-on programs, which are often referred to as applications. The distribution of success of applications is highly skewed. This market structure with few blockbuster applications and a long-tail of unpopular ones is common in the media, entertainment and software industries. Philip Mayrhofer compiled individual-level survey data as well as an original panel data set of Facebook applications in order to examine the market for Facebook applications in detail. Specifically he identifies interdependencies such

as bandwagon effects between users and spillovers between applications and analyzes whether they contribute to the concentrated market structure.

□

Advances in Digital Forensics XIII -

Gilbert Peterson 2017-08-31

Digital forensics deals with the acquisition, preservation, examination, analysis and presentation of electronic evidence. Networked computing, wireless communications and portable electronic devices have expanded the role of digital forensics beyond traditional computer crime investigations. Practically every crime now involves some aspect of digital evidence; digital forensics provides the techniques and tools to articulate this evidence. Digital forensics also has myriad

intelligence applications.

Furthermore, it has a vital role in information assurance -- investigations of security breaches yield valuable information that can be used to design more secure systems. Advances in Digital Forensics XIII describes original research results and innovative applications in the discipline of digital forensics. In addition, it highlights some of the major technical and legal issues related to digital evidence and electronic crime investigations. The areas of coverage include: Themes and Issues; Mobile and Embedded Device Forensics; Network and Cloud Forensics; Threat Detection and Mitigation; Malware Forensics; Image Forensics; and Forensic Techniques. This book is the thirteenth volume in the annual

series produced by the International Federation for Information Processing (IFIP) Working Group 11.9 on Digital Forensics, an international community of scientists, engineers and practitioners dedicated to advancing the state of the art of research and practice in digital forensics. The book contains a selection of sixteen edited papers from the Thirteenth Annual IFIP WG 11.9 International Conference on Digital Forensics, held in Orlando, Florida, USA in the winter of 2017. *Advances in Digital Forensics XIII* is an important resource for researchers, faculty members and graduate students, as well as for practitioners and individuals engaged in research and development efforts for the law enforcement and intelligence communities. Gilbert Peterson, Chair,

IFIP WG 11.9 on Digital Forensics, is a Professor of Computer Engineering at the Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, USA. Sujeet Shenoj is the F.P. Walter Professor of Computer Science and a Professor of Chemical Engineering at the University of Tulsa, Tulsa, Oklahoma, USA.

**Scientific and Technical Aerospace Reports** - 1994

International Conference on Computer Applications 2012 :: Volume 05 -  
Kokula Krishna Hari K

*Social Media, Artificial Intelligence and Carbon Neutrality* - Rita Yi Man Li 2023-02-13

Motivation, Language Identity and the

L2 Self - Zoltán Dörnyei 2009-01-12  
Due to its theoretical and educational significance within the language learning process, the study of L2 motivation has been an important area of second language acquisition research for several decades. Over the last few years L2 motivation research has taken an exciting new turn by focusing increasingly on the language learner's situated identity and various self-perceptions. As a result, the concept of L2 motivation is currently in the process of being radically reconceptualised and re-theorised in the context of contemporary notions of self and identity. With contributions by leading European, North American and Asian scholars, this volume brings together the first comprehensive

anthology of key conceptual and empirical papers that mark this important paradigmatic shift.

**Library & Information Science Abstracts** - 2008

**Criminal Investigations of Sexual Offenses** - Nadine Deslauriers-Varin 2021-11-22

It is startling to read how few sexual offenses are reported in a year; even more shocking to see how few reports lead to an arrest. Research on effective responses to sexual offenses is much needed and ongoing. This edited book presents the current state of research on investigative techniques used in sex crime investigations, and the operational challenges and issues that arise in these investigations. Including an international cohort of

scholars from various academic backgrounds, it reviews current efforts in key areas, synthesizing the research to make recommendations for the improvement of investigative practices and the criminal justice system's response to sexual offenses. Each chapter includes a brief introduction to the topic, a review of the research, and a discussion of the key points relevant to investigators and researchers as they move forward. The book is broken down into four themes: suspect prioritization and identification techniques, investigative interviewing, operational challenges and issues, and the investigation of specific types of sex crimes. With its broad coverage and international scope, this volume fills the gap in knowledge for investigators working

on sex crimes cases. With its easy-to-read style and clear analysis of the research, this is the perfect volume for investigators and law enforcement officers, policy makers and researchers, and students in criminology and criminal justice, forensic psychology, and victim services.

*Advances in Data and Web Management* -  
Guozhu Dong 2007-06-26

This book constitutes the refereed proceedings of the joint 9th Asia-Pacific Web Conference, APWeb 2007, and the 8th International Conference on Web-Age Information Management, WAIM 2007, held in Huang Shan, China, June 2007. Coverage includes data mining and knowledge discovery, P2P systems, sensor networks, spatial and temporal databases, Web mining, XML and semi-structured data, privacy and

security, as well as data mining and data streams.

**Performance Evaluation of Computer and Communication Systems. Milestones and Future Challenges** - Karin Anna Hummel 2011-12-13

Hummel 2011-12-13

This Festschrift volume is published in honor of Günter Haring on the occasion of his emerital celebration and contains invited papers by key researchers in the field of performance evaluation presented at the workshop Performance Evaluation of Computer and Communication Systems - Milestones and Future Challenges, PERFORM 2010, held in Vienna, Austria, in October 2010. Günter Haring has dedicated most of his scientific professional life to performance evaluation and the design of distributed systems, contributing in particular to the field of

workload characterization. In addition to his own contributions and leadership in international research projects, he is and has been an excellent mentor of young researchers demonstrated by their own brilliant scientific careers. The 20 thoroughly refereed papers range from visionary to in-depth research papers and are organized in the following topical sections: milestones and evolutions; trends: green ICT and virtual machines; modeling; mobility and mobile networks; communication and computer networks; and load balancing, analysis, and management.

**Machine Intelligence and Signal Analysis** - M. Tanveer 2018-08-07

The book covers the most recent developments in machine learning, signal analysis, and their applications. It covers the topics of

machine intelligence such as: deep learning, soft computing approaches, support vector machines (SVMs), least square SVMs (LSSVMs) and their variants; and covers the topics of signal analysis such as: biomedical signals including electroencephalogram (EEG), magnetoencephalography (MEG), electrocardiogram (ECG) and electromyogram (EMG) as well as other signals such as speech signals, communication signals, vibration signals, image, and video. Further, it analyzes normal and abnormal categories of real-world signals, for example normal and epileptic EEG signals using numerous classification techniques. The book is envisioned for researchers and graduate students in Computer Science and Engineering, Electrical Engineering, Applied

Mathematics, and Biomedical Signal Processing.

**Unsafe human behavior at construction sites** - Qingfeng Meng 2023-01-04

### **Artificial Intelligence and Security**

- Xingming Sun 2020-09-12

The 3-volume set CCIS 1252 until CCIS 1254 constitutes the refereed proceedings of the 6th International Conference on Artificial Intelligence and Security, ICAIS 2020, which was held in Hohhot, China, in July 2020. The conference was formerly called "International Conference on Cloud Computing and Security" with the acronym ICCCS. The total of 178 full papers and 8 short papers presented in this 3-volume proceedings was carefully reviewed and selected from 1064 submissions. The papers were organized in topical sections as

follows: Part I: artificial intelligence; Part II: artificial intelligence; Internet of things; information security; Part III: information security; big data and cloud computing; information processing.

**Social Media and the Transformation of Interaction in Society** - Sahlin, John P. 2015-09-21

The availability of various technological platforms enables individuals to feel a deeper sense of connectivity and contribution to their social circles and the world around them. This growing dependence on social networking platforms has altered the ways in which society functions and communicates. *Social Media and the Transformation of Interaction in Society* is a definitive reference source for

timely scholarly research evaluating the impact of social networking platforms on a variety of relationships, including those between individuals, governments, citizens, businesses, and consumers. Featuring expansive coverage on a range of topics relating to social media applications and uses across industries, this publication is a critical reference source for professionals, educators, students, and academicians seeking current research on the role and impact of new media on modern society. This publication features authoritative, research-based chapters across a range of relevant topics including, but not limited to, computer-mediated communication, nonprofit projects, disaster response management, education, cyberbullying,



microblogging, digital paranoia, user interaction augmentation, and viral messaging.

**Knowledge-Based Intelligent Information and Engineering Systems -**

Bruno Apolloni 2007-09-12

This book is part of a three-volume set that constitutes the refereed proceedings of the 11th International Conference on Knowledge-Based Intelligent Information and Engineering Systems, KES 2007. Coverage in this first volume includes artificial neural networks and connectionists systems, fuzzy and neuro-fuzzy systems, evolutionary computation, machine learning and classical AI, agent systems, and information engineering and applications in ubiquitous computing environments.

*Big Data Management, Technologies,*

*and Applications* - Hu, Wen-Chen  
2013-10-31

"This book discusses the exponential growth of information size and the innovative methods for data capture, storage, sharing, and analysis for big data"--Provided by publisher.

*Intelligent Technologies for Information Analysis* - Ning Zhong  
2013-03-14

Intelligent Information Technology (iiT) encompasses the theories and applications of artificial intelligence, statistical pattern recognition, learning theory, data warehousing, data mining and knowledge discovery, Grid computing, and autonomous agents and multi-agent systems in the context of today's as well as future IT, such as Electronic Commerce (EC), Business Intelligence (BI), Social Intelligence (SI), Web

Intelligence (WI), Knowledge Grid (KG), and Knowledge Community (KC), among others. The multi-author monograph presents the current state of the research and development in intelligent technologies for information analysis, in particular, advances in agents, data mining, and learning theory, from both the theoretical and application aspects. It investigates the future of information technology (IT) from a new intelligent IT (iiT) perspective, and highlights major iiT-related topics by structuring an introductory chapter and 22 survey/research chapters into 5 parts: (1) emerging data mining technology, (2) data mining for Web intelligence, (3) emerging agent technology, (4) emerging soft computing technology, and (5) statistical learning theory.

Each chapter includes the original work of the author(s) as well as a comprehensive survey related to the chapter's topic. This book will become a valuable source of reference for R&D professionals active in advanced intelligent information technologies. Students as well as IT professionals and ambitious practitioners concerned with advanced intelligent information technologies will appreciate the book as a useful text enhanced by numerous illustrations and examples.

Advances in Computer Science and Engineering - Hamid Sarbazi-Azad  
2008-11-23

It is our pleasure to welcome you to the proceedings of the 13th International Computer Society of Iran Computer Conference (CSICC-2008). The conference has been

held annually since 1995, except for 1998, when it transitioned from a year-end to first-quarter schedule. It has been moving in the direction of greater selectivity (see Fig.1) and broader international participation. Holding it in Kish Island this year represents an effort to further facilitate and encourage international contributions. We feel privileged to participate in further advancing this strong technical tradition.

|       |                  |                  |             |       |       |       |     |
|-------|------------------|------------------|-------------|-------|-------|-------|-----|
| 60    | 50               | 40               | 30          | 20    | 10    | 0     | Dec |
| 23-26 | Dec              | 23-25            | Dec         | 23-25 | Jan   | 26-28 |     |
| Mar   | 8-10             | Feb              | 21-23       | Feb   | 28-30 | Feb   |     |
| 23-26 | Feb              | 16-19            | Feb         | 15-18 | Jan   | 24-26 |     |
| Feb   | 20-22            | Mar              | 9-11        | 1995  | 1996  | 1997  |     |
| Iran  | 1999             | 2000             | 2001        | U of  | 2002  | Iran  |     |
| 2003  | 2004             | 2005             | Iran        | 2006  | IPM,  | 2007  |     |
| 2008  | Sharif U         | Amirkabir U      | of Sharif U |       |       |       |     |
|       | Shahid Isfahan,  | Telecom Ferdowsi |             |       |       |       |     |
|       | Sharif U Telecom | Tehran Shahid    | Sharif      |       |       |       |     |

U of Tech, U of Tech, Sci/Tech, of Tech, Beheshti Isfahan Res. U, of Tech, Res. Beheshti of Tech, Tehran Tehran Tehran U, Tehran Center Mashhad Tehran Center U, Tehran Kish Island Dates, Year, Venue

*CMS Security Handbook* - Tom Canavan  
2011-03-31

Learn to secure Web sites built on open source CMSs Web sites built on Joomla!, WordPress, Drupal, or Plone face some unique security threats. If you're responsible for one of them, this comprehensive security guide, the first of its kind, offers detailed guidance to help you prevent attacks, develop secure CMS-site operations, and restore your site if an attack does occur. You'll learn a strong, foundational approach to CMS operations and security from an expert in the field. More and more

Web sites are being built on open source CMSs, making them a popular target, thus making you vulnerable to new forms of attack This is the first comprehensive guide focused on securing the most common CMS platforms: Joomla!, WordPress, Drupal, and Plone Provides the tools for integrating the Web site into business operations, building a security protocol, and developing a disaster recovery plan Covers hosting, installation security issues, hardening servers against attack, establishing a contingency plan, patching processes, log review, hack recovery, wireless considerations, and infosec policy CMS Security Handbook is an essential reference for anyone responsible for a Web site built on an open source CMS.

*Engineering Software for Modern Challenges* - Dayang Norhayati A. Jawawi 2022-12-17

This volume constitutes selected papers presented at the First International Conference on Engineering Software for Modern Challenges, ESMoC 2021, held in Johor, Malaysia, in October 20-21, 2021. The 17 papers presented were thoroughly reviewed and selected from the 167 submissions. They are organized in the topical sections on □software engineering; intelligent systems; software quality.

**Proceedings of the ... ACM SIGCOMM Internet Measurement Conference - 2006**

Cyberspace Safety and Security - Guojun Wang 2013-11-08

This book constitutes the proceedings

of the 5th International Symposium on Cyberspace Safety and Security, CSS 2013, held in Zhangjiajie, China, in November 2013. The 30 full papers presented in this volume were carefully reviewed and selected from 105 submissions. In addition the book contains 6 workshop papers. The papers are organized in topical sections named: data and applications security; network and communications security; software and systems security; and cloud security and cyberspace safety.

Web Caching and Content Delivery - A. Bestavros 2001-12-14

The International Web Content Caching and Distribution Workshop (WCW) is a premiere technical meeting for researchers and practitioners interested in all aspects of content caching, distribution and delivery on

the Internet. The 2001 WCW meeting was held on the Boston University Campus. Building on the successes of the five previous WCW meetings, WCW01 featured a strong technical program and record participation from leading researchers and practitioners in the field. This book consists of all the technical papers presented at WCW'01. It includes 20 full papers and four R&D synopses that were presented at the workshop. The collection reflects the latest research in this important area, including such topics as Content Delivery Networks (CDNs), tools and methodology of performance measurements, Web characterization as relates to caching and content distribution, scalable web server architectures, cache prefetching, emerging new edge services, and delivery of streaming content.

## **Gaining Insight Into User and Search Engine Behaviour by Analyzing Web Logs - Jeeva Jose 2016-11**

Web Usage Mining, also known as Web Log Mining, is the result of user interaction with a Web server including Web logs, click streams and database transaction or the visits of search engine crawlers at a Website. Log files provide an immense source of information about the behavior of users as well as search engine crawlers. Web Usage Mining concerns the usage of common browsing patterns, i.e. pages requested in sequence from Web logs. These patterns can be utilized to enhance the design and modification of a Website. Analyzing and discovering user behavior is helpful for understanding what online information users inquire and how they behave.

The analyzed result can be used in intelligent online applications, refining Websites, improving search accuracy when seeking information and lead decision makers towards better decisions in changing markets, for instance by putting advertisements in ideal places. Similarly, the crawlers or spiders are accessing the Websites to index new and updated pages. These traces help to analyze the behavior of search engine crawlers. The log files are unstructured files and of huge size. These files need to be extracted and pre-processed before any data mining functionality to follow. Pre-processing is done in unique ways for each application. Two pre-processing algorithms are proposed based on indiscernibility relations in rough set theory which generates Equivalence Classes. The

first algorithm generates a pre-processed file with successful user requests while the second one generates a pre-processed file for pre-fetching and caching purposes. Two algorithms are proposed to extract usage analytics. The first algorithm identifies the origin of visits, the top referring sites and the most popular keywords used by the visitor to arrive at a Website. The second algorithm extracts user agents like browsers and operating systems used by a visitor to access a Website. In this study, clustering of users based on Entry Pages to a Website is done to analyze the deep linked traffic at a Website. The Top Ten Entry Pages, the traffic and the temporal information of the Top Ten Entry Pages are also studied.

**Big Data: Concepts, Methodologies,**

**Tools, and Applications** - Management Association, Information Resources 2016-04-20

The digital age has presented an exponential growth in the amount of data available to individuals looking to draw conclusions based on given or collected information across industries. Challenges associated with the analysis, security, sharing, storage, and visualization of large and complex data sets continue to plague data scientists and analysts alike as traditional data processing applications struggle to adequately manage big data. Big Data: Concepts, Methodologies, Tools, and Applications is a multi-volume compendium of research-based perspectives and solutions within the realm of large-scale and complex data sets. Taking a multidisciplinary

approach, this publication presents exhaustive coverage of crucial topics in the field of big data including diverse applications, storage solutions, analysis techniques, and methods for searching and transferring large data sets, in addition to security issues.

Emphasizing essential research in the field of data science, this publication is an ideal reference source for data analysts, IT professionals, researchers, and academics.

**Linguistics and Language Behavior Abstracts - 2008-12**