**C1. Introduction to webdatametrics, online research and its legal and ethical aspects**

|  |
| --- |
| **Subject Information:** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Code | C1 | | Plan | 2014 | ECTS | 5 |
| Type of Subject | Core | | Year | 2014 | Semester | 1 |
| Knowledge area: | Webdata management and analysis | | | | | |
| Department: | Applied Economics | | | | | |
| Virtual Platform | Platform: | Studium.usal.es | | | | |
| Access URL: | Studium.usal.es | | | | |

|  |
| --- |
| **Data about the instructors-teachers** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Teachers | Pablo de Pedraza and Francis Serrano | | | Group / s |  |
| Department | Applied Economics | | | | |
| Knowledge area | Webdatametrics | | | | |
| Academic Centre and University | University of Salamanca & University Complutense of Madrid | | | | |
| Office Address: | Campus Unamuno, Office 118, Facultad de Derecho,  37007,Salamanca, Spain | | | | |
| Personal URL: | [www.webdatanet.eu](http://www.webdatanet.eu) | | | | |
| E-mail | [pablodepedraza@usal.es](mailto:pablodepedraza@usal.es)  [fj.serrano@estumail.ucm.es](mailto:fj.serrano@estumail.ucm.es) | Telephone /Skype User: | +34923294500  (ext.) 1626 | | |

|  |
| --- |
| **Remarks (previous requirements, coordination, other, if any)** |

|  |
| --- |
| This is an introductory course, thus no specific formal requirements from students are needed. However it is expected that students are familiar with basic concepts of social science research and its methodology. |

|  |
| --- |
| **Objectives and competences of the subject (basics, general, transversal, specifics)** |

|  |
| --- |
| **Objectives:**  This course is to serve as an introductory and illustrative source of knowledge about web-based research in social sciences.  It aims to pave the way of students during the Master and their professional prospects after the Master. The program is presented in a way that helps students to select their itinerary in selective courses, the most appropriate associate partner for internship, the Master dissertation and to take these decisions on the basis of a carrier plan in line with student´s background, interests and skills. Specific goals are:  -Presenting and discuss the methodological and epistemological considerations of Internet as a research object as well as a data collection medium.  -Introducing students with a conceptualization of main online research methods and techniques -qualitative/quantitative, reactive/non-reactive, data collection/data analysis- and main fields of application in social sciences;  -Presenting different types of existing web data (I.e. Big Data, survey data) and analysis techniques options;  -Pointing out the legal and ethical issues involved in Internet research;  -Making a review of relevant practices in “real life” involving Internet research by both public and private institutions and stakeholders.  **Competences:**  The course is intended to provide students with, at least, the following competences:  -The capacity to identify and to choose appropriate Internet research methods of web data collection and analysis, in relation to specific goals and constraints of a given research project.  -The ability to critically approach a web-based research project and to be reflective as a subject of research mainly regarding ethical and legal issues.  -The capacity to write a tentative draft for the Master dissertation project in terms of reviewing literature/state of the art, defining research goals, choosing methodology, etc.  -The capacity to write a carrier plan that identifes relevant stakeholders and potential employers |

|  |
| --- |
| **Programme (Brief Description of modules) and expected learning outcomes** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | | 1. **Introduction. Doing research through the Internet.**  -Methodological and epistemological considerations on web-based research  -Offline vs. Online research.  -Data quality in Internet research | | 2. **Legal and ethical issues of online research**  -Data privacy and data security.  -Ethics in Internet research.  -Best practices in online research: Codes and standards | | 3. **Introduction to online research data collection methods**  -Web surveys  -Non-reactive methods  -Web based Tests and Experiments | | **4. Internet data analysis and tools**  **-**Online research tools: an overview  **-**Introduction to Big Data analysis tools  **-**Social network analysis, data & text mining and semantic web technologies  **-**Software for web surveys  **-**Internet data archives and sources | | **5. Applied Internet research: fields of application**  -Web data and applied economics  -Media research and web data  -Sociological and political research using the Internet  -Psychological research and web data  -Application of web data in other fields | |

|  |
| --- |
| **Methodology** |

|  |
| --- |
| The course will consist of a mixture of face-to-face and online lectures, participation, self-study & assignments during the course. |

|  |
| --- |
| **- Resources:** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| |  | | --- | | Bibliography: | | Artola, C. and E. Galan (2012). Tracking the future of the web: constructing of leading indicators using internet searches. *Banco de España, Documentos Ocasionales Nº1203.*  <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/12/Fich/do1203e.pdf>  Askitas, N. and K. F. Zimmermann (2009). Google Econometrics and Unemployment Forecasting. *IZA Discussion paper No.4201, June 2009.*  Askitas, N. and K. F. Zimmermann (2011a). Health and Well-Being in the Crisis. *IZA Discussion paper No.5601, March 2011.*  Askitas, N. and K. F. Zimmermann (2011b). Detecting Mortgage Delinquencies. *IZA Discussion paper No.5895, July 2011.*  Askitas, N. and K. F. Zimmermann (2011c). Nowcasting Business Cycles Using Toll Data. *IZA Discussion paper No.5522, Ferbruary 2011.*  Arthur, W.B. (2011) The second Economy Mackinsey Quarterly, October 2011 <http://www.mckinsey.com/insights/strategy/the_second_economy>  Biffignandi S., Bethlehem J., Handbook of Web Surveys, Wiley, Hoboken, United States, 2012, ISBN 978-0-470-60356-7  Birnbaum, M. H. (Ed.)(2000).[*Psychological experiments on the Internet*](http://psych.fullerton.edu/mbirnbaum/web/IntroWeb.htm)*.* San Diego, CA: Academic Press.  Bollena, J., Maoa,H., Zengb, X. (2011) Twitter mood predicts the stock market. Preprint submitted to Journal of Computational Science February 24, 2011  Brown, D. K. (2001). Labor Standards: Where Do They Belong on the International Trade Agenda. *Journal of Economic Perspectives, Volume 15, Number 3, Summer 2001, Pages 89-112.*  Carson-Berndsen, J., H. Somers, C. Vogel and A. Way (2010). Integrated Language Technology as part of Next Generation Localisation. Localisation Focus 8(1):53-66. Choi and Variant (2011). Predicting the Present with Google Trends. *The Economic Record, vol. 88, Special Issue, June, 2012, 2-9.*  Carrière-Swallow, Yan and Labbé, Felipe (2010). Nowcasting with Google Trends in an Emerging Market. Banco Central de Chile Documentos de Trabajo, Central Bank of Chile Working Papers N° 588, Julio 2010  Dillman, D. A. and Bowker, D. (2001). The Web Questionnaire Challenge to Survey Methodologists. In U.-D. Reips and M. Bosnjak (eds.), *Dimensions of Internet Science* (pp. 159-178). Lengerich, Germany: Pabst Science.  Eliott, A. K. and R. Freeman (2003) Can labor standards improve under globalization  Fielding, N. G., Lee, R. M. and Blank, G. (2008) The SAGE Handbook of Online Research Methods. SAGE.  Freeman, R. (1995)  Fuchs, M. and B. Busse (2009). The Coverage bias of Mobile Web Surveys Across European Countries. *International Journal of Internet Science, 2009, 4 (1), 21-33.*  Gerow, A. and M. Keane (2010). The Voice of the Herd: Power-law Regularities in Newspaper Articles Predict Market Bubbles. http://econpapers.repec.org/paper/arxpapers/1212.2676.htm  Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomicdata. Journal of Monetary Economics, 55 (4), 665–676.  Hewson, C. and D. Laurent and C. Vogel, (1996). Proper Methodologies for Psychological and Sociological Experiments Administered via Internet, Behavior Research Methods, Instruments & Computers, 28, (2), p186 - 191.Hewson, C., P. Yule, D. Laurent and C. Vogel, (2003). Internet Research Methods: A Practical Guide for the Social and Behavioural Sciences, 1st, London, SAGE Publications, ix+155pp.  Hitzler, P. and Krysztof, J (?) Linked Data, Big Data, and the 4th Paradigm, *Semantic Web 0 (0) 1, editorial,* IOs Press  Jain P., Hitzler P., Janowicz K., Venkatramani C. (2013) There’s no money in Linked Data. Available at http://knoesis.wright.edu/pascal/pub/nomoneylod.pdf, last access October 2013.  Janetzko D.(2008) Non-reactive Data Collection on the Internet. The SAGE Handbook of Online Research Methods, Nigel Fielding & Raymond M. Lee & Grant Blank.  Krantz, J. H. and Dalal, R. (2000). Validity of Web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35–60). San Diego, CA: Academic Press.  Kosinskia, M., Stillwella, D. and Graepelb, T. (2013) Private traits and attributes are predictable from digital records of human behaviour. *Proceedings of the National academy of Sciences of the United States of America.* <http://www.pnas.org/content/early/2013/03/06/1218772110.full.pdf+html>  http://www.pnas.org/content/early/2013/03/06/1218772110.abstract  Masing (2012), Herausforderungen des Datenschutzes, NJW 2012, pp. 2305-2312.  McKinsey Global Institute (2011). Big data the next frontier for innovation. <http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_forinnovation>  McLaren, Nick and Shanbhogue, Rachana (2011) Using internet search data as economic indicators. Bank of England Quarterly Bulletin 2011 Q2, pp134-140 . <http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/qb110206.pdf>  Musch, J. and Reips, U.-D. (2000). A brief history of Web experimenting. In M. H. Birnbaum (Ed.), [*Psychological experiments on the Internet*](http://psych.fullerton.edu/mbirnbaum/web/IntroWeb.htm) (pp. 61-88). San Diego, CA: Academic Press. doi 10.1016/B978-012099980-4/50004-6  Layard, R. Rethinking Public Economics: The Implications of Rivalry and Habit. In *Economics and Happiness: Framing the Analysis CN  - 1059909*; Oxford and New York: Oxford University Press ST  - Rethinking Public Economics, 2005; pp. 147–169.  [Musgrave](http://en.wikipedia.org/wiki/Richard_A._Musgrave), Richard A. (1957). "A Multiple Theory of Budget Determination," [*FinanzArchiv*](http://en.wikipedia.org/wiki/FinanzArchiv), New Series 25(1), pp. 33–43.  OCDE (2013), Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value, OCDE Digital Economy Papers, No 220, OCDE Publishing. <http://www.keepeek.com/Digital-Asset-Management/oecd/science-and-technology/exploring-the-economics-of-personal-data_5k486qtxldmq-en>  Pedraza, P. de; Tijdens, K.; Bustillo, R. M. de; Steinmetz, S. A Spanish Continuous Volunteer Web Survey: Sample Bias, Weighting and Efficiency. *Revista Española de Investigaciones Sociológicas* 2010, *131*, 109–130.  Pedraza, P. de and R. Muñoz de Bustillo (2010). Determinants of Job Insecurity in Five European Countries. *European Journal of Industrial Relations, vol. 16, number 1, March 2010, 5-20.*  Pedraza, P. de and M. Guzi (2013) "A Web Survey analysis of Subjective Well-Being of Spanish workers". IZA Discussion paper No. 7618. <http://ftp.iza.org/dp7618.pdf>  Poynter, Ray (2010): Handbook of Online and Social Media Research: Tools and Techniques for Market Researchers. Esomar.  Reips, U.-D. (1996). Experimenting in the World Wide Web. Paper presented at the 1996 Society for Computers in Psychology Conference. Chicago, IL.  Reips, U.-D. (2001). [The Web Experimental Psychology Lab: Five years of data collection on the Internet](http://personalwebpages.deusto.es/papers/Reips2001.pdf). *Behavior Research Methods, Instruments, & Computers, 33*, 201-211.  Reips, U.-D. (2002). Standards for Internet-based experimenting. [*Experimental Psychology, 49*, 243-256](http://www.hogrefe.com/periodicals/experimental-psychology/).  Reips, U.-D. (2006). Web-based methods. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 73-85). Washington, DC: American Psychological Association. doi:10.1037/11383-006.  Reips, U.-D. (2011). Privacy and the disclosure of information on the Internet: Issues and measurement. In A. Blachnio, A. Przepiórka & T. Rowinski (eds.), *Internet in Psychological Research* (pp. 71-104). Warsaw: UKSW Publishing House.  Reips, U. D. (2013). Big Data: Using Google services for research, forthcoming).  Reips, U.-D., & Bosnjak, M. (Eds.) (2001). Dimensions of Internet Science. Lengerich: Pabst.  Reips, U.-D. and Garaizar, P. (2011). Mining Twitter: Microblogging as a source for psychological wisdom of the crowds. *Behavior Research Methods, 43*, 635-642. [doi](http://dx.doi.org/10.3758/s13428-011-0116-6):10.3758/s13428-011-0116-6  Snijders, C., Matzat, U. and Reips, U.-D. (2012). ‘Big Data’: Big gaps of knowledge in the field of Internet science. *International Journal of Internet Science, 7*, 1-5.  Soros, G. (2008) *The New Paradigm for Financial Markets.* Public Affairs: New York. Xxiv+162pp.  Steinmetz, S., Tijdens, K. Raess D., Pedraza, P. de (2013) "Measuring wages worldwide – exploring the potentials and constraints of volunteer web surveys". Forthcoming in *Advancing Social and Business Research Methods with New Media Technologies*. Sappleton, N. (ed.) Hershey, PA: IGI Global.<http://www.igi-global.com/book/advancing-research-methods-new-technologies/72357>  Stieger, S. and Reips, U.-D. (2010). What are participants doing while filling in an online questionnaire: A paradata collection tool and an empirical study. *Computers in Human Behavior, 26(6)*, 1488–1495. doi:10.1016/j.chb.2010.05.013  Suhoy, Tanya (2009). Query Indices and a 2008 Downturn: Israeli Data. Research Department, Bank of Israel. http://www.boi.org.il Discussion Paper No. 2009.06, July 2009.  Tijdens, K. and M. van Klaveren (2013) "The challenge of measuring wages globally, comparing individual and country-level data: approaches and results from the WageIndicator data collection on work and wages in 75 countries". 3rd Conference of the Regulating for Decent Work Network Regulating Decent Work, Organized by International Labour Office (ILO), Geneva, 3‐5 July 2013.  Variant, Hal R. (2013). Big Data: New Tricks for Econometrics, June 2013, Revised: November 12, 2013  Vogel, C. and J. Janssen (2009). Emoticonsciousness. In Esposito, A., A. Hussain, M. Marinaro and R. Martone, Multimodal Signals: Cognitive and Algorithmic Issues, Lecture Notes in Computer Science: Volume 5398/2009, Springer, 2009, pp271 - 287.  Weltevreden, J. W. J., & Boschma, R. A. (2008). The influence of firm owner characteristics on Internet adoption by independent retailers: A business survey. *International Journal of Internet Science, 3*(1), 34-54.  \*\*\*Specific literature for each module will be provided with the final program. | |  |  |  | | --- | | Online resources: | | -[www.webdatanet.eu](file:///C:\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\STTC9JVN\www.webdatanet.eu)  -[www.websm.org](file:///C:\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\STTC9JVN\www.websm.org)  -[www.iscience.eu](file:///C:\AppData\Local\Microsoft\Windows\INetCache\Content.Outlook\STTC9JVN\www.iscience.eu)  -<http://gor14.gor.de/>  -<http://www.ijis.net/>  -<http://english.1ka.si/> | |

|  |
| --- |
| **Evaluation System:** |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | General Considerations: | | Students will receive material and literature one month before the course.  Face to Face lectures will take place during 15 hours in September and it is assume the students have been doing individual work using the material. There will be individual meeting before the beginning of the course.  There will be two on-line individual meetings where teachers monitor the work to be done. One will take place at the end of September and the other by mid-October depending on the number of students.  There will be two group meetings where students will share the state of their work by mid-November.  Evaluation is based on assignments and participation carried out during the course. Students must submit a literature review on a specific research line, a 3000 hundred words Master Dissertation proposal and a carrier development plan. The Master dissertation proposal will preferably –not compulsory- serve as the basis for the design of the Master Thesis Project. |  |  | | --- | | Evaluation Criteria: | | Students will be evaluated according to the following criteria:  - Participation in lectures and online follow up: 20 %  - Online presentation: 15 %  - Literature review: 25%  - Master/Dissertation research project proposal: 25%  - Master itinerary and Carrier development plan: 15%  Minimum 60% | |  |  |  | | --- | | Recommendation for second and following evaluations: | | Students scoring less than 60% will be requested to repeat the Online presentation, the Literature review, the Master/Dissertation Research project proposal, and/or Master itinerary and Carrier development plan. | |  | |

|  |
| --- |
| **Employment Opportunities (optional)** |

|  |
| --- |
| -- |