

C – Charakteristika studijního předmětu nebo tematického bloku	
Název studijního předmětu	Získávání znalostí z webu
Způsob zakončení	Zkouška
Další požadavky na studenta	Zpracování seminární práce
Přednášející	Doc. Ing. Ladislav Beránek, MBA, CSc.
Stručná anotace předmětu	Předmět je zaměřen na technologie získávání informací a znalostí z velkých objemů dat na WWW. Tyto data na webu mohou mít jak textovou podobu tak i multimediální. Důležitá je přitom i získání poznatků o struktuře dat, o odkazech a jejich vzájemném propojení. Cílem je získání celkového rámcového přehledu o oboru a znalostí jak např. vytipovat v kontextu řešení určitého praktického nebo vědeckého problému vhodnou metodu automatického zpracování webových dat a zorientovat se v možném postupu její aplikace.
Odborná literatura	<p>Základní: MUNZERT, S. <i>Automated data collection with R: a practical guide to Web scraping and text mining</i>. New York: Wiley, 2014. LIU, B. <i>Web data mining: exploring hyperlinks, contents, and usage data</i>. 2. vyd. Berlin: Springer, 2009, xix.</p> <p>Doporučená: LIU, B. <i>Web data mining: exploring hyperlinks, contents, and usage data</i>. Berlin: Springer, c2007, xix. DE VIRGILIO, R., GUERRA, F., VELEGRAKIS, Y. <i>Semantic search over the Web</i>. New York: Springer, c2012, xii. Data-centric systems and applications.</p>