**Программа учебной дисциплины Цифровые технологии в гуманитарных науках**

Утверждена

Академическим советом ООП

Протокол № от «\_\_»\_\_\_\_\_2017 г.

|  |  |
| --- | --- |
| Автор | Черняк Е.Л., Ильвовский Д.А. |
| Число кредитов |  |
| Контактная работа (час.) | 46 |
| Самостоятельная работа (час.) | 106 |
| Курс | 3 курс бакалавриата |
| Формат изучения дисциплины | Без использования онлайн курса |

1. **ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ**

Основная задача курса – изучение различных примеров использования цифровых технологий за пределами компьютерных наук и непосредственно индустрии ИТ.

В результате изучения дисциплины студенты должны:

* знать основные приложения цифровых технологий в гуманитарной сфере для обработки, анализа и хранения неструктурированных данных
* понимать основные принципы анализа неструктурированных данных
* уметь использовать некоторые библиотеки для анализа неструктурированных данных

Настоящая учебная дисциплина относится к профессиональному циклу и является дисциплиной по выбору в учебной программе подготовки бакалавра направления 01.03.02

«Прикладная математика и информатика» – на 3 и 4 году обучения. Она продолжает цикл дисциплин, связанных с основами анализа данных, информационных технологий и программирования. Ее изучение требует предварительных знаний по дисциплинам

«Информатика и программирование», «Дискретная математика», «Алгоритмы и структуры данных».

Основные положения рассматриваемой дисциплины могут быть использованы при выполнении курсовых и выпускных квалификационных работ.

# Содержание УЧЕБНОЙ ДИСЦИПЛИНЫ

**Тема 1. Диахронические корпуса**

Различные подходы к формализации определения изменения смысла слова со временем. Введение в дистрибутивную семантику. Темпоральные вектора слов (temporal word emeddings). Базовые понятия из теории временных рядов. Кластеризация временных рядов. Базовые понятия о Байесовских моделях. Динамические Байесовские модели.

Основные источники:

1. Hamilton, William L., Jure Leskovec, and Dan Jurafsky. "Diachronic word embeddings reveal statistical laws of semantic change." *arXiv preprint arXiv:1605.09096* (2016).
2. Frermann, Lea, and Mirella Lapata. "A Bayesian Model of Diachronic Meaning Change." *TACL* 4 (2016): 31-45.
3. Stewart, Ian, Dustin Arendt, Eric Bell, and Svitlana Volkova. "Measuring, Predicting and Visualizing Short-Term Change in Word Representation and Usage in VKontakte Social Network." *arXiv preprint arXiv:1703.07012* (2017).

**Тема 2. Наукометрика**

Индексы. Модели оценки деятельности ученых. Пространство публикаций. Граф взаимного цитирования. Модели научных сообществ. Оценка исследовательских траекторий.

Основные источники:

1. Klavans, Richard, and Kevin W. Boyack. "Toward a consensus map of science." *Journal of the American Society for information science and technology* 60, no. 3 (2009): 455-476.
2. Hall, David, Daniel Jurafsky, and Christopher D. Manning. "Studying the history of ideas using topic models." In *Proceedings of the conference on empirical methods in natural language processing*, pp. 363-371. Association for Computational Linguistics, 2008.
3. Garfield, Eugene. "Citation indexes for science. A new dimension in documentation through association of ideas." *International journal of epidemiology* 35, no. 5 (2006): 1123-1127.
4. Garfield, Eugene. "Citation analysis as a tool in journal evaluation." American Association for the Advancement of Science, 1972.
5. Guevara, Miguel R., Dominik Hartmann, Manuel Aristarán, Marcelo Mendoza, and César A. Hidalgo. "The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations." *Scientometrics* 109, no. 3 (2016): 1695-1709.

**Тема 3. Музейное наследие**

Виртуальные музеи. Трехмерное сканирование и моделирование библиотечных объектов. Моделирование архитектурных объектов. Создание и хранение текстур.

Основные источники:

1. Levoy, Marc, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton et al. "The digital Michelangelo project: 3D scanning of large statues." In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 131-144. ACM Press/Addison-Wesley Publishing Co., 2000.
2. Debevec, Paul E., Camillo J. Taylor, and Jitendra Malik. "Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach." In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 11-20.ACM, 1996.
3. Styliani, Sylaiou, Liarokapis Fotis, Kotsakis Kostas, and Patias Petros. "Virtual museums, a survey and some issues for consideration." *Journal of cultural Heritage* 10, no. 4 (2009): 520-528.

**Тема 4. Библиотечное дело**

Науки о библиотеках и информации (library and information studies). Модели организации знаний и информации. Информационный поиск. Информационные системы.

Информационная культура.

Онтология. Таксономия. Индексирование. Прямой и обратный индекс.

Основные источники:

1. Miksa, Francis L. "Library and information science: two paradigms." Conceptions of library and information science: Historical, empirical and theoretical perspectives (1992): 229-52.
2. Hjørland, Birger. "Library and information science: practice, theory, and philosophical basis."

*Information processing & management* 36, no. 3 (2000): 501-531.

1. Warner, Julian. *Human information retrieval*. Cambridge, MA: MIT Press, 2010.

**Тема 5. Психо- и социометрия**

Методы оценки валидности и надежности психодиагностических методик.

Методы анализа социальных сетей в социометрических задачах. Выделение групп. Меры центральности. Визуализация социальных сетей. Основные источники:

1. Bryant, Fred B., and Paul R. Yarnold. "Principal-components analysis and exploratory and confirmatory factor analysis." (1995).
2. Wasserman, Stanley, and Katherine Faust. Social network analysis: Methods and applications. Vol.

8. Cambridge university press, 1994.

1. Lazer, David, Alex Sandy Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis et al. "Life in the network: the coming age of computational social science." *Science (New York, NY)* 323, no. 5915 (2009): 721.

**Тема 6. Музыкальные архивы**

Обработка сигналов. Анализ музыкальных записей. Кластеризация музыкальных записей. Поиск в музыкальных архивах. Рекомендация музыкальных записей.

Основные источники:

1. Muller, Meinard, Daniel PW Ellis,Anssi Klapuri, and Gaël Richard. "Signal processing for music analysis." *IEEE Journal of Selected Topics in Signal Processing* 5, no. 6 (2011): 1088-1110.
2. Févotte, Cédric, Nancy Bertin, and Jean-Louis Durrieu. "Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis." *Neural computation* 21, no. 3 (2009): 793-830.
3. Wang, Avery. "The Shazam music recognition service." *Communications of the ACM* 49, no. 8 (2006): 44-48.
4. Levy, Mark, and Mark Sandler. "Music information retrieval using social tags and audio." *IEEE Transactions on Multimedia* 11, no. 3 (2009): 383-395.
5. Kaminskas, Marius, and Francesco Ricci. "Contextual music information retrieval and recommendation: State of the art and challenges." *Computer Science Review* 6, no. 2 (2012): 89-119.

**Тема 7. Цифровая урбанистика**

Моделирование городских объектов. Транспортные потоки. Городское планирование. Основные источники:

1. Bazzan, Ana Lécia C., Joachim Wahle, and Franziska Klügl. "Agents in traffic modelling—from reactive to social behaviour." In *Annual Conference on Artificial Intelligence*, pp. 303-306. Springer Berlin Heidelberg, 1999.
2. Ishii, Hiroshi, Eran Ben-Joseph, John Underkoffler, Luke Yeung, Dan Chak, Zahra Kanji, and Ben Piper. "Augmented urban planning workbench: overlaying drawings, physical models and digital simulation." In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, p. 203. IEEE Computer Society, 2002.
3. Phithakkitnukoon, Santi, Teerayut Horanont, Giusy Di Lorenzo, Ryosuke Shibasaki, and Carlo Ratti. "Activity-aware map: Identifying human daily activity pattern using mobile phone data." In *International Workshop on Human Behavior Understanding*, pp. 14-25. Springer Berlin Heidelberg, 2010.

**Тема 8. Новые медиа**

Степенной закон. Модели распространения влияния. Каскады цитирований. Алгоритмы распространения меток. Алгоритмы распространения вирусов.

Основные источники:

1. Adamic, Lada A., and Bernardo A. Huberman. "Power-law distribution of the world wide web." S*cience* 287, no. 5461 (2000): 2115-2115.
2. Cheng, Justin, Lada Adamic, P. Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. "Can cascades be predicted?." In *Proceedings of the 23rd international conference on World wide web*, pp. 925-936. ACM, 2014.
3. Easley, David, and Jon Kleinberg. Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press, 2010.

# ОЦЕНИВАНИЕ

Основной инструмент оценивания - презентация и обзор по выбранной теме. Студенты могут выбрать любую статью из списка основных источников или предложить статью самостоятельно для согласования с преподавателем.

# ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ

* 1. Поиск слов, изменивших смысл в НКРЯ
  2. Построение графа цитирования сотрудников одного департамента ФКН
  3. Виртуальный тур по одному залу Третьяковской галереи
  4. Электронная библиотека студентка ФКН
  5. Визуализация динамической социальной сети
  6. Поиск похожей песни по аудиозаписям в vk.com
  7. Модель автомобильной пробки
  8. Анализ популярности хэштегов

1. **РЕСУРСЫ**
   1. **Основная литература**
2. Hamilton, William L., Jure Leskovec, and Dan Jurafsky. "Diachronic word embeddings reveal statistical laws of semantic change." arXiv preprint arXiv:1605.09096 (2016).
3. Frermann, Lea, and Mirella Lapata. "A Bayesian Model of Diachronic Meaning Change." TACL 4 (2016): 31-45.
4. Stewart, Ian, Dustin Arendt, Eric Bell, and Svitlana Volkova. "Measuring, Predicting and Visualizing Short- Term Change in Word Representation and Usage in VKontakte Social Network." arXiv preprint arXiv:1703.07012 (2017).
5. Klavans, Richard, and Kevin W. Boyack. "Toward a consensus map of science." Journal of the American Society for information science and technology 60, no. 3 (2009): 455-476.
6. Hall, David, Daniel Jurafsky, and Christopher D. Manning. "Studying the history of ideas using topic models." In Proceedings of the conference on empirical methods in natural language processing, pp. 363-371. Association for Computational Linguistics, 2008.
7. Garfield, Eugene. "Citation indexes for science. A new dimension in documentation through association of ideas." International journal of epidemiology 35, no. 5 (2006): 1123-1127.
8. Garfield, Eugene. "Citation analysis as a tool in journal evaluation." American Association for the Advancement of Science, 1972.
9. Guevara, Miguel R., Dominik Hartmann, Manuel Aristarán, Marcelo Mendoza, and César A. Hidalgo. "The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations." Scientometrics 109, no. 3 (2016): 1695-1709.
10. Levoy, Marc, Kari Pulli, Brian Curless, Szymon Rusinkiewicz, David Koller, Lucas Pereira, Matt Ginzton et al. "The digital Michelangelo project: 3D scanning of large statues." In Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pp. 131-144. ACM Press/ Addison-Wesley Publishing Co., 2000.
11. Debevec, Paul E., Camillo J. Taylor, and Jitendra Malik. "Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach." In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, pp. 11-20. ACM, 1996.
12. Styliani, Sylaiou, Liarokapis Fotis, Kotsakis Kostas, and Patias Petros. "Virtual museums, a survey and some issues for consideration." Journal of cultural Heritage 10, no. 4 (2009): 520-528.
13. Miksa, Francis L. "Library and information science: two paradigms." Conceptions of library and information science: Historical, empirical and theoretical perspectives (1992): 229-52.
14. Hjørland, Birger. "Library and information science: practice, theory, and philosophical basis."
15. Information processing & management 36, no. 3 (2000): 501-531.
16. Warner, Julian. Human information retrieval. Cambridge, MA: MIT Press, 2010.
17. Bryant, Fred B., and Paul R. Yarnold. "Principal-components analysis and exploratory and confirmatory factor analysis." (1995).
18. Wasserman, Stanley, and Katherine Faust. Social network analysis: Methods and applications. Vol. 8. Cambridge university press, 1994.
19. Lazer, David, Alex Sandy Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis et al. "Life in the network: the coming age of computational social science." Science (New York, NY) 323, no. 5915 (2009): 721.
20. Muller, Meinard, Daniel PW Ellis, Anssi Klapuri, and Gaël Richard. "Signal processing for music analysis."IEEE Journal of Selected Topics in Signal Processing 5, no. 6 (2011): 1088-1110.
21. Févotte, Cédric, Nancy Bertin, and Jean-Louis Durrieu. "Nonnegative matrix factorization with the Itakura- Saito divergence: With application to music analysis." Neural computation 21, no. 3 (2009): 793-830.
22. Wang, Avery. "The Shazam music recognition service." Communications of the ACM 49, no. 8 (2006): 44- 48.
23. Levy, Mark, and Mark Sandler. "Music information retrieval using social tags and audio." IEEE Transactions on Multimedia 11, no. 3 (2009): 383-395.
24. Kaminskas, Marius, and Francesco Ricci. "Contextual music information retrieval and recommendation: State of the art and challenges." Computer Science Review 6, no. 2 (2012): 89-119.
25. Bazzan, Ana Lécia C., Joachim Wahle, and Franziska Klügl. "Agents in traffic modelling—fromreactive to social behaviour." In Annual Conference on Artificial Intelligence, pp. 303-306. Springer Berlin Heidelberg, 1999.
26. Ishii, Hiroshi, Eran Ben-Joseph, John Underkoffler, Luke Yeung, Dan Chak, Zahra Kanji, and Ben Piper. "Augmented urban planning workbench: overlaying drawings, physical models and digital simulation." In Proceedings of the 1st International Symposium on Mixed and Augmented Reality, p. 203. IEEE Computer Society, 2002.
27. Phithakkitnukoon, Santi, Teerayut Horanont, Giusy Di Lorenzo, Ryosuke Shibasaki, and Carlo Ratti. "Activity-aware map: Identifying human daily activity pattern using mobile phone data." In International Workshop on Human Behavior Understanding, pp. 14-25. Springer Berlin Heidelberg, 2010.
28. Adamic, Lada A., and Bernardo A. Huberman. "Power-law distribution of the world wide web." Science 287, no. 5461 (2000): 2115-2115.
29. Cheng, Justin, Lada Adamic, P. Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. "Can cascades be predicted?." In Proceedings of the 23rd international conference on World wide web, pp. 925-936. ACM, 2014.
30. Easley, David, and Jon Kleinberg. Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press, 2010.
    1. **Дополнительная литература**
31. Klavans, Richard, and Kevin W. Boyack. "Toward a consensus map of science." Journal of the American Society for information science and technology 60, no. 3 (2009): 455-476.
32. Hall, David, Daniel Jurafsky, and Christopher D. Manning. "Studying the history of ideas using topic models." In Proceedings of the conference on empirical methods in natural language processing, pp. 363-371. Association for Computational Linguistics, 2008.
33. Garfield, Eugene. "Citation indexes for science. A new dimension in documentation through association of ideas." International journal of epidemiology 35, no. 5 (2006): 1123-1127.
34. Garfield, Eugene. "Citation analysis as a tool in journal evaluation." American Association for the Advancement of Science, 1972.
35. Guevara, Miguel R., Dominik Hartmann, Manuel Aristarán, Marcelo Mendoza, and César A. Hidalgo. "The research space: using career paths to predict the evolution of the research output of individuals, institutions, and nations." Scientometrics 109, no. 3 (2016): 1695-1709.
    1. **Программное обеспечение**

|  |  |  |
| --- | --- | --- |
| **№ п/п** | **Наименование** | **Условия доступа** |
| 1. | Microsoft Windows 7 Professional RUS  Microsoft Windows 10  Microsoft Windows 8.1 Professional RUS | *Из внутренней сети университета (договор)* |
| 2. | Microsoft Office Professional Plus 2010 | *Из внутренней сети университета (договор)* |

* 1. **Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)**

|  |  |  |
| --- | --- | --- |
| **№ п/п** | **Наименование** | **Условия доступа** |
|  | ***Профессиональные базы данных, информационно-справочные системы*** | |
| 1. | Консультант Плюс | *Из внутренней сети университета (договор)* |
| 2. | Электронно-библиотечная система Юрайт | URL: https://biblio-online.ru/ |
|  | ***Интернет-ресурсы (электронные образовательные ресурсы)*** | |
| 1. | Открытое образование | URL: https://openedu.ru/ |

* 1. **Материально-техническое обеспечение дисциплины**

Учебные аудитории для лекционных занятий по дисциплине обеспечивают использование и демонстрацию тематических иллюстраций, соответствующих программе дисциплины в составе:

ПЭВМ с доступом в Интернет (операционная система, офисные программы, антивирусные программы);

мультимедийный проектор с дистанционным управлением.

Учебные аудитории для лабораторных и самостоятельных занятий по дисциплине оснащены ­­­­­­­­­­­­­­­­­­­­­­­­ \_\_\_\_\_\_\_\_\_\_\_\_\_\_, с возможностью подключения к сети Интернет и доступом к электронной информационно-образовательной среде НИУ ВШЭ.