

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

УТВЕРЖДАЮ

Проректор по учебной работе
и образовательным инновациям

О.И. Чуприс

«12» августа 2019 г.

Регистрационный № УД-7692/уч.

**ДИЗАЙН СОЦИАЛЬНОГО ИССЛЕДОВАНИЯ И СТАТИСТИЧЕСКИЙ
АНАЛИЗ ТЕКСТА**

**Учебная программа учреждения высшего образования
по учебной дисциплине для специальности:**

1-23 80 11 Коммуникации

Профилизация: Исследование социальной коммуникации

2019 г.

Учебная программа составлена на основе: ОСВО 1-23 80 11-2019 и учебного плана Е23а-085/уч. от 11.04.2019

СОСТАВИТЕЛЬ:

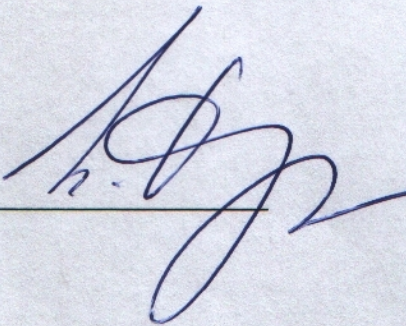
М. С. Фабрикант, доцент кафедры психологии факультета философии и социальных наук Белорусского государственного университета, кандидат психологических наук

РЕКОМЕНДОВАНА К УТВЕРЖДЕНИЮ:

Кафедрой психологии
(протокол № 12 от 27.06.2019);

Научно-методическим Советом БГУ
(протокол № 5 от 28.06.2019)

Заведующий кафедрой
психологии


И. А. Фурманов

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Дизайн социального исследования и статистический анализ текста – важнейшие составляющие профессиональной деятельности современного исследователя социальной коммуникации. Для получения релястичной,3 разносторонней и эмпирически обоснованной картины в изучении коммуникации часто требуется как генерировать данные посредством специально разработанного опросного инструментария, так и анализировать большие массивы существующего в естественной среде неструктурированного текста, в том числе с использованием современных методов статистического анализа данных и компьютерных ресурсов обработки данных.

«Дизайн социального исследования и статистический анализ текста» – практико-ориентированная дисциплина, направленная на использование современных методов сбора и обработки структурированных и неструктурированных данных. Использование опросных методов социальных исследований и статистического анализа текста повышает надежность и валидность инструментария исследований социальной коммуникации и обоснованность его результатов.

Цели и задачи учебной дисциплины

Цель учебной дисциплины – подготовка специалиста в области исследования социальной коммуникации, владеющего современными техниками дизайна социального исследования и статистического анализа текста.

Задачи учебной дисциплины:

1. изучить основные принципы дизайна социального исследования;
2. изучить принципы, техники и инструменты разработки инструментария для проведения социальных опросов;
3. изучить принципы и методы статистического анализа текста;
4. освоить компьютерный инструментарий статистического анализа текста.

Место учебной дисциплины в системе подготовки магистра.

Учебная дисциплина «Дизайн социального исследования и статистический анализ текста» относится к компоненту учреждений высшего образования и входит в модуль «Количественные методы в социальных науках».

Учебная дисциплина «Дизайн социального исследования и статистический анализ текста» читается в первом семестре для магистрантов, обучающихся на английском языке, и тесно связана с другими дисциплинами учебного плана, такими как «Дискурс-анализ», и «Теория современной коммуникативистики».

Требования к компетенциям

Освоение учебной дисциплины «Дизайн социального исследования и статистический анализ текста» должно обеспечить формирование **специализированной компетенции:**

СК-3. Уметь визуализировать полученную в ходе исследования информацию с помощью инфографики.

В результате освоения учебной дисциплины студент должен:

знать:

- основные принципы дизайна социального исследования;
- особенности поведения респондентов в зависимости от выбранного дизайна социального исследования;
- основные методы статистического анализа текста;

уметь:

- разрабатывать опросный инструментальный социального исследования;
- анализировать неструктурированный текст посредством статистических методов;
- представлять результаты статистического анализа текста при помощи инфографики;

владеть:

- навыками валидации инструментария социального исследования;
- компьютерными инструментами статистического анализа данных;
- навыками интерпретации результатов статистического анализа текста.

Структура учебной дисциплины

Дисциплина изучается в 1 семестре. Всего на изучение учебной дисциплины «Дизайн социального исследования и статистический анализ текста» отведено 90 часов, в том числе 36 аудиторных часов, из них: лекции – 18 часов, семинарские занятия – 18 часов. Трудоемкость учебной дисциплины составляет 3 зачетные единицы.

Форма текущей аттестации – экзамен.

СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

Theme 1. Designing questionnaire structure

Methodological principles of social survey design. Item response theory and its practical implications. Translating research question into survey items. Types of scales and converting parameters into variables. Predicting respondents' behavior. Satisfying behavior and the ways of its prevention. Social desirability effects and the ways of its correction. Survey layout and its impact on respondents' perceptions. Offline and online surveys.

Theme 2. Measuring sociodemographic and behavioral parameters

Sociodemographic module of surveys, its place in the survey structure and role in data analysis. Respondents' perception of sociodemographic data. Ways of measuring age and gender, their opportunities and drawbacks. Contrasting ways of measuring individual and households' income. Missing data in socioeconomic items responses and the ways of increasing the response rate.

Behavioral parameters appropriate and inappropriate for surveying. Studying behavior via self-reports and its limitations. Designing survey items for unique and repetitive kinds of behavior. Special techniques for asking sensitive questions about socially disapproved kind of behavior.

Theme 3. Measuring attitudes

The notion of attitude, its definition and structure. Disentangling cognitive, emotional, and behavioral components of attitudes. Interaction between attitudes and social behavior. Measuring attitudes using Likert scales. Measuring attitudes using open and semi-open questions in surveys. Method factor in measuring series of attitudes and the ways of its minimization. Measuring latitudes, values, and normative beliefs. Techniques for measuring attitudes on socially significant issues.

Theme 4. Questionnaire validization

Notions of reliability and validity. Test-retest method of testing reliability. Kinds of validity. Testing internal validity. Testing ecological validity. Testing construct validity. The use of cognitive interviews for questionnaire validization. Validization of online survey questionnaires. Validity of questionnaires used in longitudinal and panel surveys. Cross-cultural questionnaire validization.

Theme 5. Quantitative analysis of unstructured text

Purposes and expected outcomes of statistical text analysis. Regularity and irregularity of textual patterns. Implicit text structures. Contemporary content analysis. Principles, techniques, and approaches to categorization of textual components. R basics: data importing and processing, syntax, writing functions. Text mining R packages.

Theme 6. Text preprocessing

Sampling principles and techniques for statistical text analysis. Selection of units of text analysis. Encoding of textual units. Preprocessing of textual big data. Notion of untidy text. Text tidying techniques. Special cases of text parsing. Sentiment analysis and rhetorical analysis. Text tidying in R. Web scraping.

Tema 7. Unsupervised machine learning algorithms for text analysis

Parameter reduction and explication of hidden text structures. Clustering Latent dirichlet allocation (LDA) and latent class analysis (LCA). Principle components analysis (PCA) and factor analysis of textual data in R. K-means, K-mediod and hierarchical clustering, LDA and LCA in R.

Theme 8. Supervized machine learning algorithms for text analysis

Estimation of predictive models. Assessment and respecification of predictive models. Kind of regression analysis. Linear regression. Binary, multinomial and ordinal logistic regression. Loglinear regression. Poisson regression. Multilevel regression. Multivariate regression analysis of textual data in R. Applying predictive models for specified texts construction with AI.

Theme 9. Presenting text analysis results

Data visualization in text analysis. Word clouds as data visualization techniques. Visualization of sentiment analysis results. Visualization of frequencies and interrelation in textual data. Visualization of text structure. Data storytelling in text analysis. Visual representations of narrative structures. Temporal layouts of data visualization for text analysis results.

УЧЕБНО-МЕТОДИЧЕСКАЯ КАРТА УЧЕБНОЙ ДИСЦИПЛИНЫ

Дневная форма получения образования

Номер раздела,	Название раздела, темы	Количество аудиторных часов					Количество часов УСР	Форма контроля знаний
		Лекции	Практические занятия	Семинарские занятия	Лабораторные занятия	Иное		
1	2	3	4	5	6	7	8	9
1	Designing questionnaire structure	2		2				открытое эвристическое задание
2	Measuring sociodemographic and behavioral parameters	2		2				разрешение казусов
3	Measuring attitudes	2		2				разрешение казусов
4	Questionnaire validization	2		2				решение задач
5	Quantitative analysis of unstructured text	2		2				коллоквиум
6	Text preprocessing	2		2				решение задач
7	Unsupervized machine learning algorithms for text analysis	2		2				проект
8	Supervized machine learning algorithms for text analysis	2		2				проект
9	Presenting text analysis results	2		2				презентация

ИНФОРМАЦИОННО-МЕТОДИЧЕСКАЯ ЧАСТЬ

Перечень основной литературы

1. Beatty, P. C., Collins, D., Kaye, L., Padilla, J. L., Willis, G. B., & Wilmot, A. (Eds.). (2019). *Advances in Questionnaire Design, Development, Evaluation and Testing*. John Wiley & Sons.
2. Brace, I. (2018). *Questionnaire design: How to plan, structure and write survey material for effective market research*. Kogan Page Publishers.
3. Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice and using software*. Sage. Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2013). *Handbook of latent semantic analysis*. Psychology Press.
4. Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2014). *Automated data collection with R: A practical guide to web scraping and text mining*. John Wiley & Sons.
5. Riff, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research*. Routledge.

Перечень дополнительной литературы

1. Kumar, A., & Paul, A. (2016). *Mastering Text Mining with R*. Packt Publishing Ltd.
2. Krosnick, J. A. (2018). Questionnaire design. In *The Palgrave handbook of survey research* (pp. 439-455). Palgrave Macmillan, Cham.
3. Kwartler, T. (2017). *Text mining in practice with R*. John Wiley & Sons.
4. Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. "O'Reilly Media, Inc."
5. Willis, G. B. (2015). *Analysis of the cognitive interview in questionnaire design*. Oxford University Press.

Перечень рекомендуемых средств диагностики и методика формирования итоговой оценки

Формой текущей аттестации по дисциплине «Дизайн социального исследования и статистический анализ текста» учебным планом предусмотрен экзамен.

При формировании итоговой оценки используется рейтинговая оценка знаний студента, дающая возможность проследить и оценить динамику процесса достижения целей обучения. Рейтинговая оценка предусматривает использование весовых коэффициентов для текущего контроля знаний и текущей аттестации студентов по дисциплине.

Примерные весовые коэффициенты, определяющие вклад текущего контроля знаний и текущей аттестации в рейтинговую оценку:

Формирование оценки за текущую успеваемость:

- открытое эвристическое задание – 20 %;
- решение задач – 20 %;
- коллоквиум – 20 %;
- проект – 30%;
- презентация – 10%.

Рейтинговая оценка по дисциплине рассчитывается на основе оценки текущей успеваемости и экзаменационной оценки с учетом их весовых коэффициентов. Вес оценки по текущей успеваемости составляет 50 %, экзаменационная оценка – 50 %.

Примерная тематика семинарских занятий

Семинар № 1. Разработка опросного инструментария.

Семинар № 2. Измерение социально-демографических и поведенческих параметров

Семинар № 3. Измерение социальных аттитюдов.

Семинар № 4. Валидизация опросного инструментария.

Семинар № 5. Количественный анализ неструктурированного текста

Семинар № 6. Предварительная обработка текста.

Семинар № 7. Алгоритмы машинного обучения без супервизии для анализа текста.

Семинар № 8. Алгоритмы машинного обучения с супервизией для анализа текста.

Семинар № 9. Презентация результатов анализа текста.

Описание инновационных подходов и методов к преподаванию учебной дисциплины (эвристический, проективный, практико-ориентированный)

При организации образовательного процесса используется **метод проектного обучения**, который предполагает:

- способ организации учебной деятельности студентов, развивающий актуальные для учебной и профессиональной деятельности навыки планирования, самоорганизации, сотрудничества и предполагающий создание собственного продукта;

- приобретение навыков для решения исследовательских, творческих, социальных, предпринимательских и коммуникационных задач.

Также при организации образовательного процесса используется **практико-ориентированный подход**, который предполагает:

- освоение содержания образования через решения практических задач;
- приобретение навыков эффективного выполнения разных видов профессиональной деятельности;
- ориентацию на генерирование идей, реализацию групповых студенческих проектов, развитие предпринимательской культуры;
- использованию процедур, способов оценивания, фиксирующих сформированность профессиональных компетенций.

Методические рекомендации по организации самостоятельной работы обучающихся

Для организации самостоятельной работы студентов по учебной дисциплине «Дизайн социального исследования и статистический анализ текста» в распоряжение учащихся предоставляются учебные и учебно-методические материалы:

- учебная программа;
- учебные издания для теоретического изучения дисциплины;
- методические указания к семинарским занятиям;
- методические рекомендации для выполнения коллоквиума;
- образцы инфографики;
- материалы с примерами использования инфографики в массмедиа;
- список основной литературы и дополнительной;
- перечень информационных ресурсов.

Примерный перечень вопросов к экзамену

1. Methodological principles of social survey design.
2. Item response theory and its practical implications.

3. Translating research question into survey items.
4. Types of scales and converting parameters into variables.
5. Predicting respondents' behavior.
6. Satisfying behavior and the ways of its prevention.
7. Social desirability effects and the ways of its correction.
8. Survey layout and its impact on respondents' perceptions.
9. Offline and online surveys.
10. Sociodemographic module of surveys, its place in the survey structure and role in data analysis.
11. Respondents' perception of sociodemographic data.
12. Ways of measuring age and gender, their opportunities and drawbacks.
13. Contrasting ways of measuring individual and households' income.
14. Missing data in socioeconomic items responses and the ways of increasing the response rate.
15. Behavioral parameters appropriate and inappropriate for surveying.
16. Studying behavior via self-reports and its limitations.
17. Designing survey items for unique and repetitive kinds of behavior.
18. Special techniques for asking sensitive questions about socially disapproved kind of behavior.
19. The notion of attitude, its definition and structure.
20. Disentangling cognitive, emotional, and behavioral components of attitudes.
21. Interaction between attitudes and social behavior.
22. Measuring attitudes using Likert scales.
23. Measuring attitudes using open and semi-open questions in surveys.
24. Method factor in measuring series of attitudes and the ways of its minimization. Measuring latitudes, values, and normative beliefs.
25. Techniques for measuring attitudes on socially significant issues.
26. Notions of reliability and validity.
27. Test-retest method of testing reliability. Kinds of validity.
28. Testing internal validity.
29. Testing ecological validity.
30. Testing construct validity.
31. The use of cognitive interviews for questionnaire validation.
32. Validation of online survey questionnaires.
33. Validity of questionnaires used in longitudinal and panel surveys.
34. Cross-cultural questionnaire validation.
35. Purposes and expected outcomes of statistical text analysis.
36. Regularity and irregularity of textual patterns.
37. Implicit text structures.
38. Contemporary content analysis.
39. Principles, techniques, and approaches to categorization of textual components.
40. R basics: data importing and processing, syntax, writing functions.
41. Text mining R packages.

42. Sampling principles and techniques for statistical text analysis.
43. Selection of units of text analysis.
44. Encoding of textual units. Preprocessing of textual big data.
45. Notion of untidy text. Text tidying techniques.
46. Special cases of text parsing.
47. Sentiment analysis and rhetorical analysis.
48. Text tidying in R. Web scraping.
49. Parameter reduction and explication of hidden text structures.
50. Clustering Latent dirichlet allocation (LDA) and latent class analysis (LCA).
51. Principle components analysis (PCA) and factor analysis of textual data in R.
52. K-means, K-mediod and hierarchical clustering in R.
53. LDA and LCA in R.
54. Estimation of predictive models.
55. Assessment and respecification of predictive models.
56. Kinds of regression analysis.
57. Linear regression.
58. Binary, multinomial and ordinal logistic regression.
59. Loglinear regression.
60. Poisson regression.
61. Multilevel regression.
62. Multivariate regression analysis of textual data in R.
63. Applying predictive models for specified texts construction with AI.
64. Data visualization in text analysis.
65. Word clouds as data visualization techniques.
66. Visualization of sentiment analysis results.
67. Visualization of frequencies and interrelation in textual data.
68. Visualization of text structure.
69. Data storytelling in text analysis.
70. Visual representations of narrative structures.
71. Temporal layouts of data visualization for text analysis results.

ПРОТОКОЛ СОГЛАСОВАНИЯ УЧЕБНОЙ ПРОГРАММЫ УВО

Название учебной дисциплины, с которой требуется согласование	Название кафедры	Предложения об изменениях в содержании учебной программы учреждения высшего образования по учебной дисциплине	Решение, принятое кафедрой, разработавшей учебную программу (с указанием даты и номера протокола)
1. Выборочный метод в социологическом исследовании	Кафедра социальной коммуникации	Нет	Изменений не требуется, протокол №12 от 27.06.2019
2. Сравнительные социальные исследования	Кафедра социальной коммуникации	Нет	Изменений не требуется, протокол №12 от 27.06.2019
3. Актуальные проблемы сравнительных исследований	Нет изменений	Нет	Изменений не требуется, протокол №12 от 27.06.2019

**ДОПОЛНЕНИЯ И ИЗМЕНЕНИЯ К УЧЕБНОЙ ПРОГРАММЕ ПО
ИЗУЧАЕМОЙ УЧЕБНОЙ ДИСЦИПЛИНЕ**

на ____ / ____ учебный год

№ п/п	Дополнения и изменения	Основание

Учебная программа пересмотрена и одобрена на заседании кафедры
психологии (протокол № ____ от _____ 201_ г.)

Заведующий кафедрой _____

И. А. Фурманов

УТВЕРЖДАЮ

Декан факультета _____

В. Ф. Гигин