

Кластеризация коротких текстовых сообщений на основе комбинированного вектора атрибутов

А.В. Сычев email: sav@sc.vsu.ru

Воронежский государственный факультет

***Аннотация.** В данной работе представлены предварительные результаты машинного эксперимента по кластеризации новостных текстовых сообщений.*

***Ключевые слова:** новостные сообщения, рейтинги, ключевые слова, лемматизация, кластеризация, корреляция, факторный анализ.*

Введение

Оргкомитет Международной конференции «Информатика: проблемы, методология, технология» рад приветствовать своих участников и потенциальных участников. Данный документ представляет собой шаблон оформления и описание требований оформления работ для участников данной конференции.

Данный документ является шаблоном для оформления работ. При использовании шаблона выделяются два важных пункта:

- Необходимые стили оформления;
- Полезные макросы.

1. Стили оформления

Оформление работы должны строго выполняться с использованием только стилей, которые представлены в данном шаблоне.

Прежде чем, вы продолжите читать данный документ далее, нажмите Alt + Ctrl + Shift + S или просто Alt + S. В правой стороне окна приложения (опционально) появляется список стилей, разработанных для данного документа.

На рис. 1 показан вид данного документа с открытым списком стилей в Microsoft Word 2003.

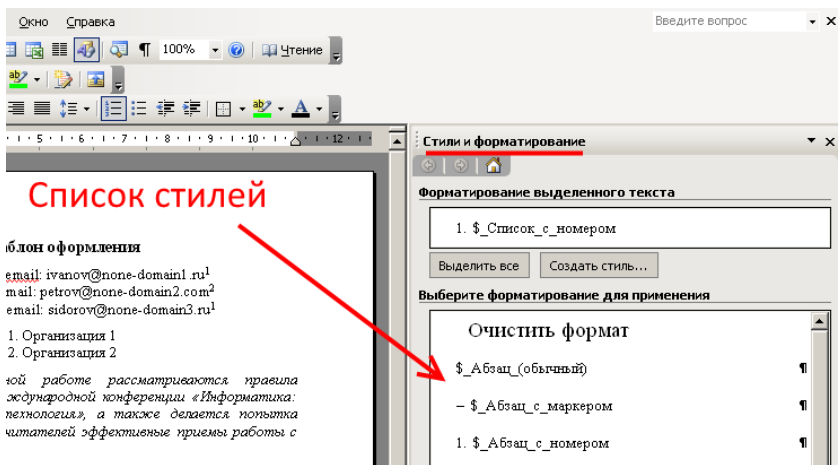


Рис. 1. Список стилей документа в MS Word 2003

В пакете Microsoft Word 2007 список стилей выглядит аналогично.

В Microsoft Word 2010 вид списка стилей немного изменился и выглядит, как показано на рис. 2.

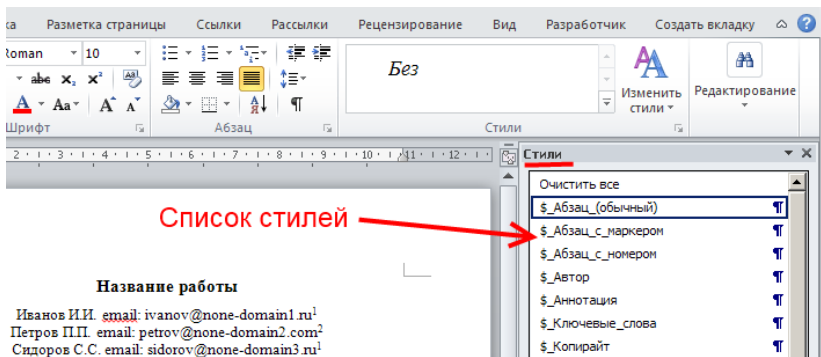


Рис. 2. Список стилей документа в MS Word 2010

В Microsoft Word 2016 список стилей появляется в виде плавающего окна, как на рис. 3.

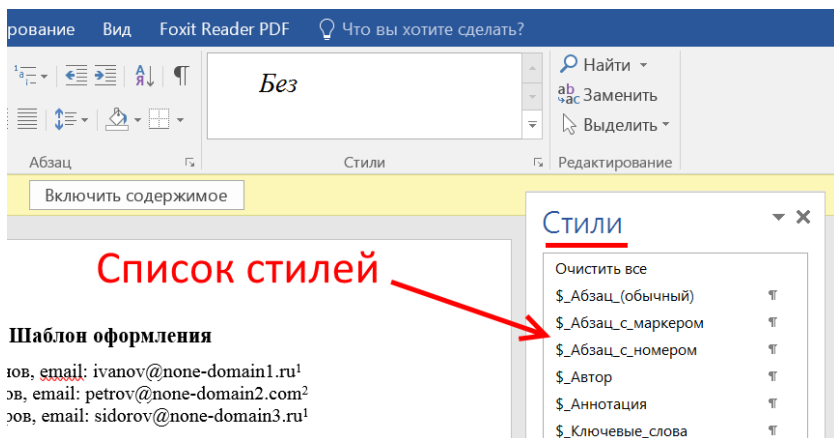


Рис. 3. Список стилей документа в MS Word 2016

Единственное, что допускается форматировать не через стили – это выделять отдельные слова (термины и т.п.) с помощью курсива, подчеркивания или «жирности».

Все стили в данном шаблоне начинаются с префикса \$_ – основные стили или @_ – служебные или вспомогательные стили (которые, в большинстве случаев вам не понадобятся).

При применении к абзацу какого-либо стиля, Microsoft Word может спрашивать что-то вроде «Переопределить»:

1. данный стиль, используя данный фрагмент как образец;
2. формат выделенного объекта, используя данный стиль».

Разработчики постарались дать интуитивно понятные названия стилям, однако определить назначения каждого стиля всегда можно путем выделения соответствующего абзаца (поставить курсор) в данном документе или в примерах правильно оформленных работ и посмотреть в списке стилей, какой именно стиль используется (используемый стиль для текущего абзаца в списке стилей выделяется рамочкой).

Заключение

Надеемся, что данный шаблон облегчит вам подготовку работ для Международной конференции «Информатика: проблемы, методология, технологии».

Также он должен существенно облегчить оргкомитету верстку сборника трудов данной конференции, поэтому мы просим авторов подойти к оформлению работ серьезно. Тем более, что как было сказано выше, материалы, оформленные без использования данного шаблона, с

несоблюдением указанных требований, автоматически исключаются из сборника трудов конференции!

Просьба к авторам, заметившим проблемы с данным шаблоном, например, неправильную нумерацию, ошибки при выполнении макросов и т.п., сообщить об этом организаторам (на e-mail: conference@sc.vsu.ru), предоставив описание проблемы и документ, на котором данная проблема наблюдается.

Список литературы

Литературные ссылки следует нумеровать в прямых скобках, нумерация сквозная. В случае нумерации диапазона источников необходимо весь диапазон включать в прямые скобки. Например: [1-4].

Библиографическое описание составляется по ГОСТ 7.1-2003 «Библиографическая запись. Библиографическое описание. Общие требования и правила составления» и ГОСТ 7.82-2001 «Библиографическая запись. Библиографическое описание электронных ресурсов».

Примеры наиболее используемых библиографических описаний приведены ниже.

Книги 1, 2, 3 автора:

1. Кириллов, В. И. Логика : учебник для юрид. вузов / В. И. Кириллов ; Моск. гос. юрид. академия. – 6-е изд., перераб. и доп. – М. : Проспект, 2009. – 233 с.
2. Грушевицкая, Т. Г. Культурология : учебник для вузов / Т. Г. Грушевицкая, А.П. Садохин. – 3-е изд., перераб. и доп. – М. : ЮНИТИ-ДАНА, 2008. – 687 с.

Книги 4 и более авторов:

3. Философия : учебник / Н. С. Савкин [и др.] ; отв. ред. Н. С. Савкин. – 3-е изд., испр. – Саранск : Морд. ун-т, 2004. – 355 с.

Журналы 1, 2, 3 автора:

4. Казаков, Н. А. Запоздалое признание : повесть / Н. Казаков // На боевом посту. – 2000. – № 9. – С. 64-76.
5. Баталов, А. Л. Сакральная топография средневекового города / А. Л. Баталов, Л. А. Беляев // Известия Института христианской культуры средневековья. – Москва, 1998. – Т. 1. – С. 13-22.
6. Боголюбов, А. Н. О вещественных резонансах в волноводе с неоднородным заполнением / А. Н. Боголюбов, А. Л. Делицын, М. Д. Малых // Вестник Московского университета. Сер. 3, Физика. Астрономия. – 2001. – № 5. – С. 23-25.

Журналы 4 и более авторов:

7. Мониторинг состояния оборудования систем связи в трубопроводном транспорте нефти / Л. И. Григорьев [и др.] // Автоматизация, телемеханизация и связь в нефтяной промышленности. – 2007. – № 5. – С. 3-8.

Материалы конференции:

8. Моисеева, М. В. Японская мультипликация и ее восприятие в России / М. В. Моисеева, Е. В. Кудряшева // Социокультурная миссия университета в современном обществе : сб. тр. участников I Всерос. научно-практической конф. "Человек, культура, образование" (Ульяновск, 2-4 июля 2005 г.). – Ульяновск, 2006. – С. 142-145.

Патентные документы:

9. Приемопередающее устройство [Текст] : пат. 2187888 Российская Федерация : МПК Н 04 В 1/38, Н 04 J 13/00 / Чугаева В. И. ; заявитель и патентообладатель Воронеж. науч.-ислед. ин-т связи. – № 2000131736/09 ; заявл. 18.12.00 ; опубл. 20.08.02, Бюл. № 23 (II ч.). – 3 с.

10. Пат. 2187888 Российская Федерация, МПК7 Н 04 В 1/38, Н 04 J 13/00. Приемопередающее устройство [Текст] / Чугаева В. И. ; заявитель и патентообладатель Воронеж. науч.-ислед. ин-т связи. – № 2000131736/09 ; заявл. 18.12.00 ; опубл. 20.08.02, Бюл. № 23 (II ч.). – 3 с.

ГОСТЫ:

11. Стекло безопасное для наземного транспорта. Общие технические условия [Текст] : ГОСТ 5727-88. – Изд. сент. 2001 с Изм. 1, 2, 3 (ИУС. 1992. № 2; ИУС. 1999. № 1; ИУС. 2001. № 11). – Взамен ГОСТ 5727-83 ; введ. 01.01.90.

12. ГОСТ Р 517721–2001. Аппаратура радиоэлектронная бытовая. Входные и выходные параметры и типы соединений. Технические требования [Текст]. – Введ. 2002-01-01. – М. : Изд-во стандартов, 2001. – IV, 27 с. : ил. ; 29 см.

Ресурсы локального доступа:

13. Сидыганов, В. У. Модель Москвы [Электронный ресурс] : электрон. карта Москвы и Подмосковья / В. У. Сидыганов, С. Ю. Толмачев, Ю. Э. Цыганков. – Версия 2.0. – Электрон, дан. и прогр. – М. : FORMOZA, 1998. – 1 электрон. опт. диск (CD-ROM).

14. Атлас-98 [Электронный ресурс] : 3D., 1998. – 1 электрон. опт. диск (CD-ROM).

Сетевые ресурсы:

15. Исследовано в России [Электронный ресурс] : многопредмет. науч. журн. / Моск. физ.-техн. ин-т. – Электрон. журн. – Режим доступа: <http://zhurnal.mipt.rssi.ru>

16. Электронный каталог ГПНТБ России [Электронный ресурс] : база данных. – Режим доступа : <http://www.gpntb.ru/win/search/help/el-cat.html>

При описании URL-адресов в конце URL-адреса точка не ставится!

В библиографических описаниях и в тексте материалов следует различать дефис («-»), который используется, как правило, внутри слов для их разделения, использования сокращения, указания диапазона страниц и т. д. (например: Ростов-на-Дону, д-р, 15-26), и тире (длинное тире «—» [«m-dash»]), которое используется для разделения библиографических описаний (например: ... Автоматизация, телемеханизация и связь в нефтяной промышленности. – 2007. – № 5).