

Извлечение терминов

Взят текст учебника по философии. Получившаяся таблица.

	Словосочетание	Частота	Словосочетание	МІ
1	человеческий природа	61	ядерный разоружение	14.054
2	первый очередь	56	эффектный заголовок	14.054
3	другой человек	51	шизоатлантический картография	14.054
4	другой сторона	46	церковно-нравственный дидактика	14.054
5	живой существо	44	цветной капуста	14.054
6	человеческий разум	42	ходячий трюизм	14.054
7	новый время	42	фонетический созвучие	14.054
8	научный знание	40	финикийский заселение	14.054
9	единичный вещь	36	финансово-податной округ	14.054
10	естественное состояние	31	ученый изыскатель	14.054
11	гегелевский философия	31	удивленный эфесец	14.054
12	иной слово	29	территориальный экспансия	14.054
13	человеческий познание	28	текстильный фабрикант	14.054
14	человеческий душа	28	сталелитейный магнат	14.054
15	большой часть	28	спусковой крючок	14.054
16	чистый разум	26	скользящий угорь	14.054
17	французский мыслитель	26	связный повествование	14.054
18	философский традиция	26	румянцевский музей	14.054
19	философский система	26	рачительный расходование	14.054
20	русский философия	26	равноапостольный кирилл	14.054

Терминами являются 'человеческий природа' и 'человеческий разум', т.к. это философские понятия. 'Живое существо' – термин (понятие из биологии). 'Новый время' также является термином, так как обозначает период истории и в тексте пишется с заглавной буквы. 'Научное знание', 'Человеческий познание', 'Человеческий душа', 'Чистый разум' также относятся к философским понятиям. 'гегелевский философия' и 'русский философия' – термины, так как обозначают течения в философии. Итого: 10 терминов.

Мера взаимной информации ставит на первые места словосочетания, которые редко встречались в тексте. В данном случае каждое из словосочетаний встречалось лишь однажды. Термины: 'ядерный разоружение' – исторический процесс, 'церковно-нравственный дидактика' – раздел науки, 'фонетический созвучие' – явление в фонетике, 'территориальный экспансия' – политическое явление, 'спусковой крючок' – часть оружия. Терминов 5.