

## Zagadnienia egzaminacyjne dla studiów Zarządzanie Big Data

### Bazy:

1. Charakterystyka relacyjnych i nierelacyjnych baz danych.
2. Dokumentowe bazy danych – charakterystyka i zastosowania.
3. Bazy klucz-wartość – charakterystyka i zastosowania.
4. Bazy rodziny kolumn -wartość – charakterystyka i zastosowania.
5. Bazy danych – replikacja, sharding, partycjonowanie.
6. Bazy grafowe – charakterystyka i zastosowania.
7. Normalizacja i denormalizacja w bazach danych – wady i zalety.
8. Problematyka rozproszonych systemów baz danych.
9. Główne cechy baz danych nazywanych bazami NoSQL
- 10.

### Big Data:

11. Rafinacja informacji.
12. Funkcje hashujące – rodzaje i zastosowanie
13. Metody deduplikacji informacji. Kolekcjonowanie danych ze źródeł internetowych – narzędzia, metody. Sposoby korzystania z zasobów Big Data.
14. Nieindeksowane zasoby sieciowe.
15. Zasoby publiczne. Open data. Dostęp do informacji publicznej. Prawo.
16. Metadane i ich rola w zarządzaniu zasobami cyfrowymi.
17. Istota rafinacji informacji
18. Kolekcjonowanie materiałów źródłowych
19. Machine learning
20. TSNE
21. Trendy i korelacje w statystyce
22. Wyrażenia regularne - charakterystyka, zastosowanie
23. TF-IDF - metoda, zastosowanie
24. Prawo Zipfa
25. Lematyzacja i stemming - wyjaśnić pojęcia, zastosowanie
26. Gramatyka wizualizacji danych
27. Przetwarzanie języka naturalnego
28. Komercyjne media elektroniczne na rynku.
- 29.

### IT:

30. Oprogramowanie open source i licencjonowane. Charakterystyka.
31. Systemy operacyjne – charakterystyka.
32. JSON,BSON,XML,CSV,YAML – cyfrowe formaty danych – charakterystyka i zastosowanie.
33. API (application programming interface) – charakterystyka, zastosowanie.
34. ACID i BASE – charakterystyka.
35. Przetwarzanie wsadowe, Mapreduce – charakterystyka i zastosowania.
36. Procedura budowy programu.
37. Charakterystyka i parametry elementów składowych PC. Benchmarki.

38. Oprogramowanie open source i licencjonowane. Charakterystyka.
39. Rodzaje sieci komputerowych. Charakterystyka. Zastosowanie.
40. Urządzenia peryferyjne komputera. Charakterystyka. Parametry.
41. Cyfrowa forma multimedialnych – istota. Cyfrowe nośniki informacji.
42. Parametry komputera
43. Grafika rastrowa i wektorowa. Charakterystyka. Formaty. Edytory.
44. Cyfrowy dźwięk i wideo. Charakterystyka. Formaty. Edytory.
45. Streaming audio i wideo. Charakterystyka. Formaty/kontenery.
46. Elementy procedury i konstrukcji niezbędne dla uruchomienia strony WWW.
47. Funkcjonalność serwisu WWW. Statystyki – narzędzia pomiaru.

#### Internet:

48. Ocena stron. Page Rank.
49. SEM
50. SEO
51. Domeny internetowe
52. Usługi w chmurze i hosting.
53. Usługi w internecie.
54. Wirtualizacja zasobów IT.
55. Istota wyszukiwarek
56. Komunikacja sieciowa
57. Zasoby informacyjne w sieci
58. Media internetowe

#### CSR

59. Znaczenie społecznej odpowiedzialności biznesu na rynku pracy
60. Zastosowanie CSR w rozwoju przedsiębiorstwa
61. Raportowanie niefinansowe - znaczenie i zakres informacji
62. Obszar zainteresowań badawczych CSR w mediach
63. Ekonomizacja CSR
64. Trendy w działaniach społecznie odpowiedzialnych

#### Systemy informatyczne zarządzania

65. Znaczenie skrótu i podstawowe cechy ZSI
66. Koncepcje (sposoby) pozyskania systemów informatycznych
67. Wady, niebezpieczeństwa i problemy wynikające z wdrożenia w firmie systemu dedykowanego – systemu opracowanego od podstaw „na miarę” dla przedsiębiorstwa
68. Zalety i korzyści dla firmy związane z zakupem i wdrożeniem systemu powielarnego (pudełkowego), tzn. systemu w pełni gotowego do pracy
69. Znaczenie skrótów i podstawowe cechy systemów klasy ERP i ERP II

#### Statystyka:

70. Omówić miary analizy struktury zbiorowości: położenia, rozproszenia, asymetrii, koncentracji.
71. Jakich testów użyjemy do analizy zmiennych jakościowych?
72. Jak weryfikujemy niezależność dwóch zmiennych ilościowych?
73. Omówić model regresji liniowej.
74. Scharakteryzować modele analizy regresji nieliniowej.
75. Zdefiniować przykładowe testy t-Studenta.
76. Opisać przykładowe nieparametryczne testy dla prób zależnych i niezależnych.
77. Omówić podstawowe funkcjonalności programu Excel/Statistica/R w zakresie analizy danych.

#### Ekonomia informacji - pytania:

78. Informacja jako: zasób ekonomiczny, produkt, towar, jako dobro konsumpcyjne - definicje i przykłady.
79. Proces informacyjny jako proces ekonomiczny.
80. System informacyjny jako system ekonomiczny.
81. Wartość i cena informacji.
82. Rola kadry kierowniczej w zarządzaniu informacją.
83. Elementy ekonomiki informacji w zarządzaniu obiektem gospodarczym.
84. Zastosowanie metod i mierników opracowanych przez ekonomikę informacji do oceny sytuacji ekonomicznej podmiotów gospodarczych.
85. Zastosowanie elementów ekonomiki informacji w zarządzaniu informacją.

#### Bezpieczeństwo IT

86. Kwalifikowany podpis elektroniczny
87. Szyfrowanie
88. Niezawodność IT
89. Wirusy, ransomware i inne podobne zagrożenia
90. Ochrona danych