

# BD01 Kompleksowe wprowadzenie do Big Data cz. 1

Harmonogram: łącznie 16 godz.

LP	Temat szkolenia	Data realizacji szkolenia	Godzina rozpoczęcia	Godzina zakończenia	Ilość godzin
<b>Dzień pierwszy</b>					
1	Czym jest Big Data		9:00	9:45	1
2	Cele i historia powstania, typowe zastosowania		9:45	10:30	1
	Przerwa kawowa		10:30	11:00	
3	Apache Hadoop - wprowadzenie		11:00	11:45	1
4	Hadoop: HDFS a. Architektura b. Korzystanie z CLI		11:45	12:30	1
	Przerwa obiadowa		12:30	13:30	
5	Hadoop - MapReduce a. Wprowadzenie do paradygmatu b. Tworzenie procesów MapReduce c. Uruchamianie i zarządzanie zadaniami MapReduce		13:30	15:00	2
	Przerwa kawowa		15:00	15:30	
6	Hadoop - YARN a. Czym jest YARN i powody powstania b. Konfiguracja i zarządzanie		15:30	17:00	2
<b>Dzień drugi</b>					
1	Apache Hive a. Wprowadzenie b. Architektura		9:00	10:30	2
	Przerwa kawowa		10:30	11:00	
2	Apache Hive a. Tabele zewnętrzne i wewnętrzne b. Przetwarzanie danych za pomocą języka HiveQL		11:00	12:30	2
	Przerwa obiadowa		12:30	13:30	

<b>3</b>	Import i eksport danych do baz SQL - sqoop		13:30	14:15	1
<b>4</b>	CAP Theorem		14:15	15:00	1
	Przerwa kawowa		15:00	15:30	
<b>5</b>	Apache Pig a. Wprowadzenie b. Architektura c. Typy danych d. Tryby pracy e. Przetwarzanie danych za pomocą języka PigLatin		15:30	17:00	2