

DATA+ Big Data & Business Analytics - program

1. dzień konferencji - czwartek, 26 listopada 2015

08.30-09.00	Rejestracja Uczestników. Poranny poczęstunek.
Plenarna sesja poranna.	
09.00-09.40	Data-Driven Storytelling. Ken Cherven , Senior Analyst and Data Visualization Specialist, General Motors
09.40-10.20	Next Step in Big Data: State-of-the-art Open Source Data Science Platform. • CASE STUDY: PETROL SECTOR Charles Cai , Head of Data Science Technology, BP (Trading Division)
10.20-10.40	Przerwa na kawę. Networking.
10.40-11.10	Why Data Modeling is essential in a Big Data environment. Mark Barringer , Product Line Manager, Enterprise Data Management, Tools, EMEA & LATAM Embarcadero
11.10-11.25	Big Data w Polsce i za granicą. Aleksandra Woźniak , Research Analyst w funduszu private equity
11.25-12.00	Współpraca europejskich badawczych centrów danych z biznesem. Maciej Remiszewski , Zastępca Dyrektora ICM ds. Rozwoju i Współpracy, ICM UW
12.00-13.00	Lunch

DATA+ Biznes

13.00-13.40

Data Scientist - wyzwania i rola w organizacji.

- Jak budować i jak zarządzać takimi zespołami?
- Kto w organizacji może zajmować się przetwarzaniem dużych wolumenów danych?

DATA+ Technologie

13.00-13.40

Data Science w środowisku korporacyjnym.

[Piotr Gawrysiak](#), Chief Data Scientist, mBank

- Jakich efektów działania zespołu BIG DATA możemy się spodziewać?

[Michał Janusz](#), Data Scientist, Grupa Wirtualna Polska

13.40-14.20

Zasady wizualizacji danych, które powinnyś znać Ty i Twój dostawca BI.

[Bartosz Czapiewski](#), Trener i konsultant, założyciel, SkuteczneRaporty.pl

13.40-14.20

Przetwarzanie danych - nowe metody i procesy

[Dominik Ryżko](#), Chief Data Scientist, Grupa Allegro

14.20-14.40

Przerwa na kawę. Networking.

14.40-15.20

CX w służbie poprawy konwersji, retencji użytkowników i budowaniu spójnego doświadczenia.

[Jarosław Trybuchowicz](#), Business Intelligence Analyst, Audioteka

15.20-16.00

Nowa era Social Business Intelligence - jakie tajemnice pozwoli Ci odkryć analiza społecznościowa.

- Jaka może być wartość informacji i kontaktów społecznościowych dla firmy.
- Co mógłbyś wiedzieć i zbierać o swoim kliencie - Analiza Focus on Social Customer
- Wykorzystaj Social Contact Center, aby uzyskać zaangażowanie klienta i polepszyć z nim kontakt.
- Monitorowanie konkurencji, aktywności, profilowanie, targetowanie, analizy sentymentu.
- Analiza zachowań - Konwersja z fanów / followers na kupujących.
- Czy docierasz do swoich klientów w miejsca gdzie oni przebywają i socjalizują się?

[Dawid Pawłowski](#), Chief Digital Officer, Aforti Holding S.A.

16.00-16.20

Zakończenie 1. dnia konferencji. Zebranie ankiet od Uczestników i losowanie upominków.

14.40-15.20

Dlaczego Spark? Co wnosi nowy silnik przetwarzania danych w Hadoop?

- Architektura Lambda. Analiza batch oraz streaming z zastosowaniem Spark.
- Hive, Hive on Spark, SparkSQL czy HiveContext w Spark - jak się w tym wszystkim odnaleźć?
- Hadoop i dalej. Co mogą zaoferować nam inne rozwiązania - Elasticsearch, Cassandra

[Krzysztof Adamski](#), Team lead w zespole Hadoop, ING Services Polska

15.20-16.00

Analityka predykcyjna - skuteczne narzędzie, czy wróżenie z fusów?

[Arkadiusz Olchawa](#), IT Director, ITAKA

2. dzień konferencji - INTERAKTYWNE SESJE WARSZTATOWE

08.30-09.00

Rejestracja Uczestników. Poranny poczęstunek i networking.

09.00-12.30

09.00-12.30

Warsztat I - Failing >> fast forward. How data drives lean innovation (UX, Agile, Tools)

Abstract:

"If we have data, let's look at data. If all we have are opinions, let's go with mine."

Jim Barksdale, former CEO of Netscape

As a manager or executive you are most probably aware of this imperative of change, many of the attempts to start new digital products or improve existing ones fail. In most cases because companies try to tackle new problems with the same old tools (organizational structures, methods, processes) that created the problems in the first place. How to avoid such and other traps and provide valuable experiences? In order to build a product that stands even a chance for survival or even success, quick steps, permanent analysis of relevant data and constant iterative feedback loops are needed.

During our more than a decade-long experience with digital projects in corporate and start-up contexts, we have identified a toolset that we consider essential for product innovation. It is based on the principles of Lean UX and Agile methodologies, data-driven, pragmatic and straight-forward. half-day-day workshop, entertaining, playful, hands-on, tangible results, learning by doing, less theory, more practice.

Programme:

- The importance of data in user experience
- Essential tools and methods for lean innovation
- Insights to how data drives the agile innovation process
- How to implement these in your work environment

[Wolfram Knelangen](#), Coach, Consultant

[Agnieszka M. Walorska](#), Creative Construction Heroes GmbH

12.30-13.30

Lunch

13.30-17.00

Warsztat II - Budowa modeli predykcyjnych w praktyce.

Modele predykcyjne wykorzystuje się w optymalizacji działań firm od lat. Najbardziej powszechne zastosowania to modele prognozujące zachowanie klientów (application scoring, behavioral scoring, churn). Jest wiele mniej standardowych zastosowań, choćby przewidywanie na podstawie scenariusza, ile zarobi nakręcony na jego podstawie film.

Niezależnie od przewidywanego zjawiska, sposób budowy modeli jest podobny. Próbuąc prognozować zjawiska z wykorzystaniem modeli predykcyjnych, spotykamy się jednak z wieloma trudnościami.

Warsztat III - Data Science - praktyczne warsztaty z analizy danych.

Profil warsztatów: biznesowy

Każdego dnia w firmie gromadzone są terabajty danych. Jak wyciągnąć z nich wiedzę i wykorzystać ją do podejmowania decyzji biznesowych?

Na warsztatach:

- Będziesz pracować na praktycznym przykładzie
- Poznasz warsztat pracy Data Scientist
- Dowiesz się jakich narzędzi i technik możesz użyć do analizy danych
- Zobaczysz, jak odkrywać wiedzę ukrytą w zbiorach danych

W programie warsztatów:

- Poznasz sposób na przejście od raportowania do analizy danych
- Dowiesz się, jak zebrać dobrej jakości dane do analizy
- Zobaczysz, jak przeprowadzać analizy danych w ogólnodostępnych narzędziach
- Nauczysz się przygotowywać dobre wizualizacje danych

Mile widziane:

- Posiadanie komputera z zainstalowaną przeglądarką internetową
- Podstawowa znajomość arkusza kalkulacyjnego

[Michał Bryś](#), Data Scientist, Grupa Allegro

13.30-17.00

Warsztat IV - Interaktywna eksploracja i wizualizacja danych przy użyciu Apache Spark.

Hadoop wciąż kojarzy się głównie z offline'owym przetwarzaniem olbrzymich zbiorów danych za pomocą Map-Reduce. Implementujemy naszą aplikację, zgłaszamy ją do wykonania i... idziemy na kawę albo lunch, gdyż obliczenia trwają zbyt długo. Na szczęście coraz większą popularność zdobywa narzędzie, które pozwala nam efektywniej wykorzystać cenny czas! Spark wprowadził istną rewolucję w modelu pracy z dużymi danymi. Odejdźcie od sztywnego paradygmatu Map-Reduce i wprowadzenie

Uczestnicy dowiedzą się, jakie to trudności, jak je pokonać oraz co jest ważne w modelowaniu predykcyjnym.

Przynajmniej od dekady darmowe oprogramowanie wykorzystywane jest coraz chętniej w biznesie. Często wypiera też drogie rozwiązania komercyjne. Użyjemy właśnie darmowego oprogramowania KNIME. Zdobyte wiedza i umiejętności będą wartościowe, niezależnie od wykorzystywanego w firmie oprogramowania.

Warsztaty będą przeprowadzone w sposób interaktywny, a uczestnicy dzięki pracy na własnych komputerach będą mogli wypróbować różne możliwości. Skorzystamy z ciekawych rzeczywistych danych.

Uczestnicy dowiedzą się:

- Co jest najważniejsze w modelowaniu predykcyjnym.
- Jak są typowe trudności w modelowaniu i sposoby ich omijania.
- Jak są etapy procesu budowy modelu predykcyjnego: od zebrania danych, poprzez wybór cech, ocenę jakości, aż do zastosowania modelu.
- Jak przygotowywać dane do budowy modeli predykcyjnych.
- Jak wybrać cechy, zbudować model i ocenić jego jakość.
- Poznają wybrane metody statystyczne.
- Jak zrobić to wszystko z wykorzystaniem darmowego oprogramowania KNIME.

W programie warsztatów

- Cykl życia modelu
- Proces budowy modelu predykcyjnego
- Praca z KNIME
- Przygotowanie danych
- Określenie targetu dla modelowania
- Metody wyboru cech
- Sposób budowy wybranych modeli klasyfikacyjnych
- Ocena jakości modeli

Wymagania

Jakiegokolwiek doświadczenie w pracy z danymi, w dowolnym narzędziu. Warsztaty będą interaktywne, dlatego zapraszamy do zabrania ze sobą laptopa z zainstalowaną najnowszą wersją KNIME (KNIME + all free extensions) - https://www.knime.org/downloads/overview?quicktabs_knimed=1#quicktabs-knimed (do ściągnięcia bezpłatnie).

dr inż. Artur Suchwałko, data scientist, założyciel i właściciel firmy QuantUp.pl

interaktywnego shell'a otworzyło zupełnie nowe możliwości przed każdym kto chce szybko i wydajnie pracować z danymi na klastrze.

Na potrzeby warsztatu przygotowaliśmy starannie dobry miks teorii, ćwiczeń, demonstracji, Q&A oraz zabawy. Po szybkim omówieniu architektury Sparka, przejdziemy do praktycznej części. Każdy uczestnik wcieli się w rolę analityka, którego zadaniem będzie eksploracja zbiorów danych przechowywanych na HDFS przy użyciu Sparka i wizualizacja wyników.

Uczestnicy

O ile szkolenie jest techniczne, wcześniejsze doświadczenie z Hadoop-em czy Sparkiem nie jest wymagane. Rekomendowana jest znajomość w stopniu przynajmniej podstawowym jednego z języków programowania: Java, Scala, Python. Każdy uczestnik powinien przyjść z własnym laptopem, z którego będzie używał podczas wykonywania ćwiczeń. Będziemy korzystać z wielowęzłowego klastra Hadoop zainstalowanego w publicznej chmurze. Laptop nie musi posiadać żadnego dedykowanego oprogramowania – wystarczy przeglądarka internetowa i terminal.

[Radosław Kita](#), Ekspert Data Scientist, GetinData

Organizatorzy dołożą wszelkich starań, aby konferencja odbyła się zgodnie z prezentowanym programem, jednak zastrzega się możliwość częściowych zmian.