|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  | | --- | --- | | POLITECHNIKA ŚLĄSKA  WYDZIAŁ INŻYNIERII MATERIAŁOWEJ I METALURGII | | | Kierunek: Informatyka Przemysłowa | | | Rodzaj studiów: stacjonarne | | | PROJEKT INŻYNIERSKI | | | Michał Ćwiękała | | | Integracja JAVA z Ms excel | | | Kierujący projektem: | Recenzent: | | Dr inż. Joanna Lisok | Dr inż. Marcin Blachnik | | Katowice, styczeń, 2016 r. | | |
| Spis treści  [Wstęp 3](#_Toc439618503)  [Microsoft Excel 3](#_Toc439618504)  [Java 4](#_Toc439618505)  [Statystyki 4](#_Toc439618506)  [Dlaczego programiści wybierają technologię Java? 5](#_Toc439618507)  [Biblioteka XLLoop 6](#_Toc439618508)  [Zastosowanie 6](#_Toc439618509)  [Posiada następujące właściwości: 6](#_Toc439618510)  [XLLoop składa się z dwóch głównych komponentów: 7](#_Toc439618511)  [RapidMiner 7](#_Toc439618512)  [Zarys 8](#_Toc439618513) |

1)      Wstęp – zawiera wprowadzenie + info skąd taki pomysł na projekt, i krótko zarysować w sposób ogólny skąd taki temat pracy. Tutaj można np. napisać że wiele aplikacji użytkowych w tym do przetwarzania dnaych napisanych jest w Javie i istnieje potrzeba ich integracji z Excelem, tak aby użytkownik mógł korzystać z dobrodziejstw obydwu rozwiązań

W dzisiejszym skomputeryzowanym świecie biznesowym praca z arkuszem kalkulacyjnym to podstawa w każdej firmie i instytucji, niezależnie od branży. Microsoft Excel to najpopularniejszy na rynku program służący do dokonywania obliczeń zestawionych w formie tabelarycznej. [1] Dostarcza on funkcje matematyczne, finansowe, bazodanowe oraz zapewnia półautomatyczne powtarzanie pisanych formuł. Excel jest na dla systemów Windows i MacOS. Program został napisany w 1987 przez Dan'a Bricklin'a, jednak nikt nie spodziewał się jego niesamowitego sukcesu. [2]

Excel został napisany początkowo w języku C, a następnie wprowadzono do niego dużo modyfikacji z C++. Wspiera developerów korzystających z C#, który został również stworzony przez Microsoft. Jednak najbardziej popularnym językiem programowania od wielu lat pozostaje Java.

Technologia Java została stworzona przez firmę Sun Microsystems, jego podstawowe koncepcje zostały przejęte z języka C++ - składnia i słowa kluczowe oraz Samlltalk - maszyna wirtualna i zarządzanie pamięcią. Java jest językiem wysokiego o wysokim poziomie abstrakcji, jest językiem kompilowanym do kodu bajtowego, który wykorzystywany jest przez JVM - wirtualną maszynę Java.

Java jest podstawą projektowania praktycznie każdego rodzaju aplikacji sieciowych jak również desktopowych. Jest ogólnoświatwym standardem z którego korzysta się do produkcji aplikacji mobilnych, wbudowanych, gier, zawartości i treści internetowych. Ponad 9 milionów developerów Java przyczyniło się do nowoczesnego tworzenia, implementacji oraz wykorzystywania tych aplikacji i usług.

Statystyki[3]

* 97% komputerów lokalnych, używanych w przedsiębiorstwach, korzysta z oprogramowania Java
* 87% komputerów lokalnych, używanych w Stanach Zjednoczonych, korzysta z oprogramowania Java
* 9 milionów programistów rozwija oprogramowanie oparte na technologii Java
* Nr 1 dla programistów
* Nr 1 jako platforma programistyczna
* 3 miliardy telefonów komórkowych korzysta z oprogramowania Java
* 100% odtwarzaczy dysków Blu-ray jest dostarczanych wraz z oprogramowaniem Java
* 5 miliardów kart, będących w użyciu, korzysta z oprogramowania Java
* 125 milionów urządzeń telewizyjnych korzysta z oprogramowania Java
* 5 z 5 czołowych producentów OEM dostarcza oprogramowania Java ME

Technologia Java została stworzona w ten sposób, by pozwolić na programowanie przenośnych oraz wysoce wydajnych aplikacji dla możliwie najszerszego spektrum platform przetwarzania cyfrowego. Dzięki aplikacjom dostępnym w heterogenicznych środowiskach, firmy mogą dostarczać więcej usług oraz zwiększać produktywność, komunikatywność i współpracę użytkowników końcowych, radykalnie zmniejszając koszty związane z prowadzeniem działalności i utrzymywaniem aplikacji konsumenckich.

1 - Computerworld: [Polski Word 2.0 i Excel 4.0 już w sprzedaży](http://www.computerworld.pl/artykuly/39139/Polski.Word.2.0.i.Excel.4.0.juz.w.sprzedazy.html). [dostęp 2009-01-31]

2 - http://naukaexcel.pl/podstawy/kurs-o-historii-excela-jak-to-sie-stalo-ze-excel-zostal-krolem-na-rynku-arkuszy-kalkulacyjnych/

3 - https://www.java.com/pl/about/ [dostęp 2015-12-27]

http://computeraidedfinance.com/2011/11/26/what-shall-i-choose-for-implementing-an-excel-plugin-and-udf/

2)      Cel i zakres pracy – gdzie opisuje pan krótko dokładnie zdefiniowany cel np. celem pracy jest integracja platformy MS Excel z aplikacjami napisanymi dla środowiska Java. Zakres pracy obejmuje:

-  Zbudowanie przykładowej aplikacji stworzonej dla systemu RapidMiner, tak aby możliwe było wywoływanie z poziomu excela funkcji stanowiących procesy napisane w RapidMinerze.

– wykorzystanie biblioteki XLLoop do integracji środowisk MS Excel i Java

- może coś jeszcze

                3) Możliwe narzędzia integracji MS Excel i Java

                               - opisać w pierwszym podrozdziale co to jest excel i jakie są możliwe technologie jego rozszerzenia w tym UDF i ta druga technologia

                               - opisać w podrozdziałach poszczególne biblioteki umożliwiające integrację Javy i Excela czyli to co najpierw chcieliśmy wykorzystać, krótko można to opisać i wymienić wady oraz zalety tego rozwiązania. Z Wad to oczywiście to że dla najnowszych systemów widows są problemy z jej wykorzystaniem

                               Potem opisać krótko drugie rozwiązanie tj. bibliotekę XLLoop

                4) Opisać krótko RapidMiner co to jest i do czego służy oraz dlaczego nam zależy na jego integracji z Excelem

                5) Opisać szczegóły implementacyjne projektu, tj. instalacja po stronie excela, instalacja po stronie klienta tj. jak pan podpiął własną funkcję pod XLLoopa itp.

                6) Przykład użycia – tutaj pokazać przykładowy proces RapidMinera i jak go użyto w Excelu

                7) Wnioski i podumowanie – tutaj krótko co było celem i jak się go udało osiągnąć – tak na ¾ strony razem z tytułem rozdziału. Całość dopbrze by zajęła około 25/30 stron łącznie ze spisem literatury itp.

# Wstęp

## Microsoft Excel

Microsoft Excel (pełna nazwa Microsoft Office Excel) – arkusz kalkulacyjny produkowany przez firmę Microsoft dla systemów Windows i MacOS. Pierwsza wersja programu przeznaczona dla Windows trafiła na rynek w roku 1987 i stała się przebojem. Postępujący sukces rynkowy programu sprawił, że w roku 1993 programy pakietu Microsoft Office zostały przeprojektowane tak, by przypominać wyglądem arkusz Excel. Od wersji 5 wydanej w 1993 program zawiera wbudowany język Visual Basic. Od wersji 4.0 dostępny w wersji polskiej[1].

Aplikacja jest powszechnie używana w firmach i instytucjach, a także przez użytkowników domowych. Jej główne zastosowanie to dokonywanie obliczeń (np. wydat ków) zestawionych w formie tabelarycznej. W tym użyciu mają zastosowanie liczne funkcje matematyczne, finansowe i bazodanowe dostępne w programie. Istotne znaczenie ma też półautomatyczne powielanie tworzonych formuł z zastosowaniem różnych wariantów adresowania (adresowanie względne, adresowanie bezwzględne, adresowanie mieszane). Microsoft Excel służy także do tworzenia wielu typów wykresów, przydatnych między innymi w fizyce, matematyce i ekonomii. Zawiera też system zestawiania raportów z użyciem tzw. tabel przestawnych, wykorzystywany przy wykonywaniu analiz biznesowych. [2]

|  |
| --- |
|  |
| Rys. 1.1 GUI aplikacji Microsoft Excel 2007 |

## Java[3]

Technologia Java to podstawa projektowania niemal każdego rodzaju aplikacji sieciowej — stanowi ogólnoświatowy standard wykorzystywany przy tworzeniu aplikacji dla urządzeń mobilnych, aplikacji wbudowywanych, gier, zawartości i treści internetowych oraz oprogramowania dla przedsiębiorstw. Społeczność ponad 9 milionów oddanych programistów przyczynia się do sprawnego tworzenia, implementacji oraz użytkowania aplikacji i usług.

|  |
| --- |
| Logo Java |
| Rys. 1.2 Logo platformy Java |

Java jest wszędzie — od laptopów po centra danych, od konsol do gier po superkomputery naukowe, od telefonów komórkowych po Internet.

### Statystyki

* 97% komputerów lokalnych, używanych w przedsiębiorstwach, korzysta z oprogramowania Java
* 87% komputerów lokalnych, używanych w Stanach Zjednoczonych, korzysta z oprogramowania Java
* 9 milionów programistów rozwija oprogramowanie oparte na technologii Java
* Nr 1 dla programistów
* Nr 1 jako platforma programistyczna
* 3 miliardy telefonów komórkowych korzysta z oprogramowania Java
* 100% odtwarzaczy dysków Blu-ray jest dostarczanych wraz z oprogramowaniem Java
* 5 miliardów kart, będących w użyciu, korzysta z oprogramowania Java
* 125 milionów urządzeń telewizyjnych korzysta z oprogramowania Java
* 5 z 5 czołowych producentów OEM dostarcza oprogramowania Java ME

### Dlaczego programiści wybierają technologię Java?

Technologia Java była testowana, udoskonalana, rozszerzana i ulepszana przez oddaną społeczność programistów, architektów i entuzjastów. Została zaprojektowana tak, aby umożliwić programowanie przenośnych, wysoce wydajnych aplikacji dla możliwie najszerszego spektrum platform przetwarzania cyfrowego. Dzięki aplikacjom dostępnym w heterogenicznych środowiskach, firmy mogą dostarczać więcej usług oraz zwiększać produktywność, komunikatywność i współpracę użytkowników końcowych, radykalnie zmniejszając koszty związane z prowadzeniem działalności i utrzymywaniem aplikacji konsumenckich. Technologia Java stała się nieocenionym narzędziem dla programistów, umożliwiając im:

* uruchamianie programów, pisanych na jednej platformie, na praktycznie dowolnej innej platformie;
* tworzenie programów, które mogą być uruchamiane w przeglądarkach internetowych i mogą uzyskiwać dostęp do usług internetowych;
* opracowywanie aplikacji serwerowych obsługujących fora internetowe, sklepy internetowe, ankiety, formularze HTML i wiele innych;
* łączenie aplikacji lub usług wykorzystujących język Java w celu stworzenia aplikacji i usług ściśle dostosowanych do konkretnych potrzeb;
* pisanie skutecznych i wydajnych aplikacji dla telefonów komórkowych, odległych urządzeń przetwarzających, mikrokontrolerów, modułów bezprzewodowych, sensorów, bramek, produktów konsumenckich oraz praktycznie wszelkich innych urządzeń elektronicznych.[3]

## Biblioteka XLLoop[4]

XLLoop to open sourcowa biblioteka pozwalająca implementować Excelowe funkcje zdefiniowane przez urzytkownika ( UDF - User-Defined Functions) na zcentralizowanym serwerze (serwerze funkcyjnym).

### Zastosowanie

|  |
| --- |
|  |
| Rys. 1.3 schemat działania XLLoop |

* Funkcje mogą być napisane w każdym języku (w tym Java, języki skryptowe itd.),
* Funkcje mogą być dodane szybko i dynamicznie bez konieczności restartowania Excela,
* Funkcje mogą być zarządzane oddzielnie i centralnie, co zapobiega kosztów zarządzania wieloma XLL'ami i gwarantuje wszystkim użytkownikom używanie tych samych funkcji,
* Dane mogą być dzielone pomiędzy excelowymi sesjami/użytkownikami ( np. obecne ceny giełdowe mogą być przechowywane na pojedynczym serwerze i wszystkie sesje Excela mogą uzyskiwać te dane za pomocą przykładowej funkcji GetStock).

### Posiada następujące właściwości:

* Rozszerzalna **Javowa** biblioteka serwerowa,
* Lokalna swrwerowa biblioteka napisana w wielu językach,
* Pop-up dla długo pracujących operacji ( z opcją anulowania)
* Wsparcie dla hostingu C++ XLL'i przez JXLL
* Skupiona refleksyjnie funkcja "adapter" dla wyjątkowo krótko rozwijanych metod java.
* Adapter framework'owy "bean scripting" - funkcje mogą być napisane w każdym wspieranym przez BSF języku (np. Ruby, Python, JavaScript),
* Lisp Function Handler dla wywoływania Lisp'owych wyrażeń podczas lotu,
* Funkcje mogą mogą być dynamicznie dodawane do Excela bez konieczności restartu,
* Funkcje mogą być rejestrowane jako Excelowe funkcje ( z pomocniczymi informacjami)
* Szybki binarny protokuł z wielo-serwerową skalownością
* JSON ponad protokołami HTTP(S)
* Może być konfigurowane do połączenia z wieloma serwerami funkcyjnymi (dostawcami)

### XLLoop składa się z dwóch głównych komponentów:

* Excelowy dodatek (XLL napisany w C++)
* Serwer i biblioteka napisana w Java ( i w wielu innych językach)

Dodatek i serwer komunikują się poprzez zwykłe oparte o schowek protokoły, wysyłające i odbierające (binarnie) zserializowane excelowe obiekty (nazywane xloper).

W modyfikacji JSON dodatek i serwer komunikują się ponad protokołami http(s), wysyłając i otrzymując JSON'owe zserializowane excelowe obiekty. XLLoop jest pod licencją Common Public License (CPL).

## RapidMiner

RapidMiner to najczęściej używana open sourcowa platforma do predykcyjnej analizy oraz eksploracji danych. Tworzony w nim kod generuje się za pośrednictwem intuicyjnego graficznego interfejsu użytkownika, który pomaga zaprojektować analizowany proces. RapidMiner posiada upublicznione biblioteki służące do rozszerzenia jego funkcjonalności, można je znaleźć na serwerze GitHub.com. Platforma jest otwarta do współpracy z mnóstwem źródeł danych w tym: Excel, Access, Oracle, IBM DB2, Microsoft SQL, Netezza Teradata, MySQL, Postgres, SPSS, Salesforce.com. RapidMiner umożliwia nowoczesną analizę w każdej skali, jest wręcz idealne dla Big Data.

|  |
| --- |
|  |
| Rys. 1.4 GUI Rapidminera |

# Zarys pracy

Moja praca dyplomowa pod tytułem "Integracja Java z MS Excel" ma na celu zbadanie możliwości rozbudowy funkcjonalności programu MicroSoft Excel 2007 o własną implementację UDF tzn. funkcji zdefiniowanej przez użytkownika. Funkcja komunikuję się z Excelem za pomocą biblioteki XLLoop, która tworzy lokalny serwer nasłuchujący protokoły wysyłane z Excela. Następnie wywołuję metodę zaimplementowaną w Javie, która wywołuje proces z RapidMinera zapisany w formacie xml. W moim projekcie stworzyłem proces, posiadający sieć neuronową, która rozróżnia gatunki irysów na podstawie parametrów a1-a4. Następnie podaje procentową wartość, który gatunek z pośród trzech najlepiej pasuje do próbki i zapisuje wynik predykcji. Wynik przekazywany jest w zwracanej wartości metody i przesyłany do Excela.

Aplikacja posiada GUI, po uruchomieniu którego Serwer lokalny nasłuchuje, uruchomiony dodatek, który przekazuje Excelowi zaimplementowane metody (w naszym przypadku np. "=RapidMinerFunctions.dobierzIrysa()").

# Bibliografia

1 - Computerworld: [Polski Word 2.0 i Excel 4.0 już w sprzedaży](http://www.computerworld.pl/artykuly/39139/Polski.Word.2.0.i.Excel.4.0.juz.w.sprzedazy.html). [dostęp 2009-01-31]

2 - https://pl.wikipedia.org/wiki/Microsoft\_Excel#cite\_note-1 [dostęp 2016-01-01]

3 - https://www.java.com/pl/about/ [dostęp 2015-12-27]

4 - http://xlloop.sourceforge.net/ [dostęp 2015-12-22]

5 - https://rapidminer.com/ - zebrane i opracowane materiały przez autora pracy i przetłumaczone [dostęp 2016-01-02]