WS 2019
SPBA – Software Praktikum mit Bachelorarbeit

Topics @ Research Group
Multimedia Information Systems (MIS)

Wolfgang Klas (wolfgang.klas@univie.ac.at)
Gerald Quirchmayr (gerald.quirchmayr@univie.ac.at)
Belal Abu Naim (belal.abunaim@univie.ac.at)
Nour Jnoub (nour.jnoub@univie.ac.at)
Thematic Scope of SPBA

**Application Domains**
- Blockchain Applications
- IoT/Home Automation
- Fact Checking
- Individual Activity and Mood Recognition
- Real-time Crowd Sourcing

**Media Types**
- Audio
- Video
- Text
- Image
- Sensor data

**Resources**
- Sensors, like:
  - Smartphone
  - Social Media
  - Real-time Crowd
  - Camera
  - Mic
  - ...

**Information Security**
- Legislative Compliance
- Attack (Pattern) Detection
- Attack Prevention
- Threat Information Handling

**Technologies/Methods**
- Blockchain Technology
- Machine Learning
- Natural Language Processing
- Image Processing
- Video/Audio Processing
Thematic Scope of SPBA

• Blockchain-based applications & Technologies
  • Blockchain technology in various configurations
  • Local Ethereum blockchain platform of BlockchainSci-Lab
  • MultiChain Blockchain platform
  • Ethereum, EthereumJ, Smart-contracts

• Semantic Web Technologies
  • Web services, RDF, LOD, Microformat, JSON-LD, etc.

• Languages, Platforms, Libraries
  • Android, Java, GO, Python, P2P, Bluetooth, Wi-Fi, FFMPEG
  • Machine Learning Toolkits, scikit-learn, NLP tool kits

• Information Security
  • Legislative Compliance, Attack (Pattern Detection), Threat Information
Technologies and Platforms

• Robotics
  • EV3 Middleware Framework
  • Lego MindStorm Roboter,
  • ROS & ROS packages für EV3 (http://wiki.ros.org/Robots/EV3 );
  • Raspberry PI 3
  • Java; C-based Arduino PL
  • Android, Open street maps (OSM)
If further information needed?

If interested in 1 topic from MIS group, send an email with subject ‘SPBA-ws19’ as soon as possible to the advisor of the topic:

NJx... Nour Jnoub: nour.jnoub@univie.ac.at
WKx... Wolfgang Klas: wolfgang.klas@univie.ac.at
BAx... Belal Abu Naim: belal.abunaim@univie.ac.at
GQx... Gerald Quirchmayr gerald.quirchmayr@univie.ac.at

Include your skills & motivation for the topic you want to work on!

If you are interested in more than one topic of the MIS group, send one email with subject ‘SPBA-ws19’ with the topics listed, which you are interested in, to all advisors (see above). We will synchronize internally and propose information.

**FINAL TOPIC SELECTION and ASSIGNMENT:**
Follow the procedure as outlined on Moodle (or announced) later in detail !!
WK1 Blockchain: Prüfungspass (Ethereum-based)

• Implementation of a “Sammelzeugnis” in terms of a digital, blockchain-based “Prüfungspass” for students of a course. Functions include:
  LV admission, to draw the place in the course, to withdraw in time, to submit assignments, to take a test, grading of assignments and tests, final grading according to overall rules, etc.

• Implement the logic by using Ethereum Smart Contracts and by using proof-of-authority

• Provided: local Ethereum Blockchain infrastructure @ BlockchainSci-Lab of CS Faculty

• Technologies: Blockchain Technology, Smart Contract, Web Technology
WK2 Blockchain: Prüfungspass (MultiChain-based)

- Implementation of a “Sammelzeugnis” in terms of a digital, blockchain-based “Prüfungspass” for students of a course. Functions include:
  LV admission, to draw the place in the course, to withdraw in time, to submit assignments, to take a test, grading of assignments and tests, final grading according to overall rules, etc.

- Implement the logic by using MultiChain Blockchain technology infrastructure @ BlockchainSci-Lab of CS Faculty

- Technologies: Blockchain Technology, Smart Contract, Web Technology
WK3 Demo of Blockchain application using Proof-of-Authority

• implement a demo application which illustrates the concept of proof-of-authority (in place of the very often used "proof-of-work" as, e.g., used in the Bitcoin Blockchain).

• For example, a possible application could be the implementation of the four-eyes principle (Vier-Augen-Prinzip) for officially approving documents by making use of two signers acting as proof-of-authorities.

• It may be based on a generic, configurable implementation to show different variations of the proof-of-authority concept, e.g., 1 signer, 2 signers, N signers.

• Technologies: Blockchain-IT infrastructure with initial configuration of proof-of-authority based on Ethereum Project
WK4 ROS-package for Lego MindStorm Robot

• Design und Implementierung des ROS-Package für die Steuerung des Lego MindStorm Roboters
• Lego MindStorm Roboter gemäß vorgegebenem Bauplan zu konstruieren
• Design und Implementierung einer Demoanwendung zwecks Test der Lösung im Rahmen einer Anwendung. Die Anwendung kann in Abstimmung mit dem Betreuer frei gewählt werden.
• Lego MindStorm EV3 & LejOS Betriebssystem für Lego MindStorm Roboter

• Technologies: EV3 Middleware Framework (von MIS, eine Basis Library für das Ansteuern des MindStorm Roboters); Lego MindStorm Roboter, Aufbau gemäß Vorgaben; Basis Library für das Ansteuern des MindStorm Roboter; ROS & ROS packages für EV3 (http://wiki.ros.org/Robots/EV3 ); Raspberry PI 3, recommended OS Ubuntu 15.10 bzw. 16.04
WK5 Finger gesture-controlled MindStorm Robot

- Komplettierung der Konstruktion eines Handschuhs unter Verwendung gegebener Bausteine
- Implementierung der Verarbeitung der Sensordaten des Handschuhs durch den Arduino Microprozessor
- Implementierung der Software zur Interpretation der Signale vom Handschuh und Steuerung des Lego-MindStorm Roboters
- Lego MindStorm Roboter gemäß vorgegebenem Bauplan zu konstruieren
- Implementierung der Software für das visuelle Monitoring (Status des Handschuhs, Befehle an Roboter, Status des Roboters) des Systems über einen separaten Bildschirm
- Test der Lösung im Rahmen einer Anwendung

- Technologies: Lego EV3 & LejOS Betriebssystem für Lego MindStorm Roboter; Lego MindStorm Roboter, Aufbau gemäß Vorgaben; EV3 Middleware Framework (von MIS, eine Basis Library für das Ansteuern des MindStorm Roboters); Raspberry PI 3 OS (z.B. OS Ubuntu 15.10 bzw. 16.04); Java; C-based Arduino PL
WK6 FactCheck - Precision Metrics

• Framework for the detection and resolution of conflicting structured data on the Web.

• Context-dependent comparison of structured data of various representations of one and the same real world object or artefact guided by so called precision metrics which is a flexible and sophisticated technique for logically comparing structured data values.

• Goal of the project is to design and implement an appropriate model for the representation of precision metrics, the construction of such precision metrics as well as the application of the metrics for evaluating the comparison of data values. Various precision metrics should be defined and compared using a test dataset of 900.000 entities. Results of the project are to be demonstrated by a running demo application.

• Technologies: web services, semantic web technologies, LOD, Microformat, JSON-LD; (a given prototype/demo available to start with!)
WK7 FactCheck - Merging data from different datacentres and distributing suitable data given specific demands

• Different Systems often use similar data entities for different purposes. Therefore, it is reasonable to make it possible that one system makes use of the data from another system.

• The goal of this project is the implementation of a demo application that gets fed data from different systems. The datasets are getting evaluated before inserting them in the database. If there already exists a similar entry (entity recognition), the data should get updated by overwriting old data or adding additional information to the dataset. The data provider should also be able to get datasets from the systems by their needs. Each request can have configurations to it which determine which data-fields should have specific requirements. The system then provides the best suitable data without giving any dataentry the consumer already has.

Technologies: various database technologies, Web technologies
BA1 – Audio Narrated Photo Presentations

• With the proliferation of mobile devices and applications that allow the users to take photos and record audio files that describe these photos, users become in need for tools that help them to organize their photos and the related audio files in professionally composed presentations. Linking an image with the most relevant part of an audio file that describes this image remains a non-trivial task for the users. This project aims at finding a solution for the described problem, i.e., an audio narrated photo presentations.

Technology: Android, Java, image processing, audio processing, NLP
Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
Visiting new places enthuses users to shoot photos. In addition to the photos, users can record a voice narrative or even a video on those visited places. Users usually possess several devices, i.e., mobile phones, cameras, and other devices which can be used to generate different kinds of digital content. Having generated different multimedia digital content that can be linked to a place or a certain activity, users require a helping tool that enables them to organize those generated multimedia content in one multimedia document, we will call this multimedia document "Digital Diary". To make an added value for our tool, it is expected that the tool performs some analysis tasks on the generated multimedia content to identify some subjects that those content encode. One of those encoded subjects are identified, the tools retrieve some additional data on those subject from open linked data provides, e.g., DBpedia, and include those retrieved data in the Digital Diary.

Technology: Android, Java, image processing
Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
BA3 - Interactive multimedia content composition system with using robots

• Social network platforms and new generations of smartphones cause a change in the role of the users from being primarily multimedia content consumers to becoming multimedia content producers and consumers. The vast amount of daily generated and exchanged user’s multimedia content using smartphones and social network platforms introduces new requirements and urges revisiting traditional solutions focusing on the user as an active producer and consumer for the multimedia content instead of being a (passive) consumer only.

• Since robots are equipped with different input and output devices, they can interact with users during the lifecycle of the authoring process to collect additional user’s preferences that can be used as input and accordingly can influence the quality of the authored document. This requires extending and adapting the authoring process by enabling the interaction with the user and updating the input of the different authoring phases according to the information that the user provides through the interaction with the robot.

Technology: Android, Java
Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
BA4 - Adding Graph-Data Service-Layer to Ethereum

• Blockchain provides a reliable infrastructure for persisting read-only data. Besides the basic data model i.e., block headers, which is required by blockchain to connect the blocks in a chain, blockchain technology allows storing and retrieving data encoded by means of custom data models.

• The goal of this project is to find a solution for storing the data in graphs and make use of the graph theory to connect the pieces of the graphs which are stored in the blocks and extract new hidden information using the available reasoning methods. To be able to fulfill the requirements of this project, your solution can make use of the available platforms and frameworks, i.e., ethereum, ethereumJ, GO language, Apache Jena, etc. In other words, it is not expected from you to implement a new blockchain system, but to extend the functionalities of one of the open source blockchain systems by adding a new graph-data service-layer.

Technology: Java, Ethereum, EthereumJ, Solidity, GO language.
Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
**BA5 - Semantic blockchain**

- Blockchain systems store transactions and other relevant data in blocks, usually these blocks are stored in the file systems as binary files. Such a storing method makes the process of querying and retrieving the stored data from the blockchain complicated and in many cases inefficient.

- The goal of this project is to develop a blockchain system that makes use of semantic web technologies such as RDF to store and interlink the data. Within this project, you are expected to implement a blockchain system that captures and stores information about some real property, the owner of that property. You should use RDF graphs to store and interlink the data.

Technology: Java, Jena framework.

Contact: Belal Abu Naim belal.abunaim@univie.ac.at
Service-oriented architecture (SOA) and cloud computing led to the existence of the acronym X as a Service (XaaS), e.g., Infrastructure as a service (IaaS), Platform as a service (PaaS), and so on. As a result of such architecture, many service providers make their services available on the internet. Services consumers can search and make use of the provided services in various ways and for various purposes. The goal of this project is built SOA system that makes use of such available services. Given an image that represents an object, e.g., an old-timer car, some famous artefact such as the "Mona Lisa", etc., the application should make use of the available services to:
1. Recognize the object that an image represents.
2. Collects additional information about the recognized object from the available open data services.
3. Compose a multimedia document about the subject image.

Technology: Java, Android, Spring framework, JavaScript, HTML5

Services: DBpedia, clarfai, etc.

Object recognition and generating enriched multimedia documents (https://clarifai.com/demo)

Contact: Belal Abu Naim belal.abunaim@univie.ac.at
BA7 - Distributed composition of multimedia content

• Multimedia content authoring has been a challenging research topic for many years. In view of the IoT, (X)aaS and SOA, multimedia content authoring can take a new form of distributed authoring. The goal of this project is to design and implement a new type of a distributed authoring process to author multimedia content. The process should make use of different devices, data sources, and services to collect multimedia content, compose a multimedia presentation, and produce a ready to present multimedia content on the target device.

Technology: Java, FFMPEG, Android

Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
BA8 - Benchmarking Consensus in Permissioned Blockchains

• Permissioned blockchains are gaining popularity in enterprise contexts. Contrary to open blockchain-based systems like Bitcoin or Ethereum, they operate in a configurable and closed environment with defined access restrictions. This scenario makes the use of traditional consensus mechanisms designed for permission-less networks with a huge number of anonymous nodes like Proof of Work unsuitable since the focus lies rather on performance than on fraud prevention. Therefore, these technologies use alternative consensus models with a lower number of validating nodes operating in a permissioned network allowing for higher transaction throughput.

• The goal of this project is to benchmark and analyse trade-offs between transaction throughput, latency as well as scalability and security of permissioned blockchain networks built on technologies available today.

Technologies: Ethereum(Clique, Aura), Quorum (Raft, Istanbul), Hyperledger Fabric, Tendermint
Contact: Belal Abu Naim  belal.abunaim@univie.ac.at
[In cooperation with external technology partner in Vienna]
BA9 - Scalability Solutions for Permission-less Blockchains

• Current permission-less blockchain systems like Bitcoin or Ethereum use distributed consensus mechanisms to ensure the integrity and immutability of their state. This leads to a bottleneck in the validation process and therefore to low transaction throughputs. Reality has shown that large applications running on permission-less blockchains can cause severe performance issues on the network, which is undeniably a major reason for preventing the mass adoption of critical decentralized apps. There are several strategies that try to solve these problems by executing state-modifying operations parallelly off the main chain, which leads to the enhancement of scalability.

• The goal of this project is to evaluate strategies like state channels and sidechains on the base of a specific blockchain like Ethereum, Bitcoin, etc.

Technologies: Ethereum, Bitcoin, NEM
Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
[In cooperation with external technology partner in Vienna]
BA10 - Evaluation of Specific DLT/Blockchain Technologies

• Growing popularity and adaption of distributed ledger technologies (DLTs) like blockchains have resulted in an extensive variety of different platforms. Many of them follow the principle of storing the transaction history as a sequence of blocks whereas others use for example approaches like directed acyclic graphs. All these technologies vary in terms of consensus-model, validation-process, privacy-settings, etc.

• Therefore, the goal of this project is the proper evaluation of platforms that are not yet as highly adapted as common blockchains like Ethereum or Bitcoin. Possible technologies can be taken from the list below.

Technologies: Cardano, Hashgraph, IOTA, Monero, EOS, NEO

Contact: Belal Abu Naim, belal.abunaim@univie.ac.at

[In cooperation with external technology partner in Vienna]
BA11 - Game-based Learning

• With the proliferation of personal computers and mobile devices, users have the chance to experience various types of applications including the digital educational games. Educational games have educational purposes. They are designed to help users to learn about certain subjects or develop certain skills. Digital educational games are interactive play that teach users goals, rules, problem solving, and interaction. On the one hand, this project aims at studying the concepts of Game-based learning and gamification. And on the other hand it aims at applying these concepts on a digital education game.

Contact: Belal Abu Naim, belal.abunaim@univie.ac.at
NJ1 - Clustering user's related opinions on social media platforms

Twitter or other similar social networks provide a huge amount of comments every minute. Clustering such comments have different applications and helps to have an overview of the current ways of thinking of people in a specific city or a country.

The Goal of the project is to use natural language processing techniques and machine learning approaches to extract features from text and cluster similar comments of users that may have same opinion or same content.

Results of the project are to be demonstrated by a running demo application.

Technologies: Python, (Natural Language Toolkit (NLTK), Scikit-learn, Keras)

Contact: Nour Jnoub nour.jnoub@univie.ac.at
NJ2 - Users’ reactions detection based on their opinions

Twitter or other similar social networks provide a huge amount of comments every minute. Classifying such comments plays an important role to understand the well-being of people in the community and have an estimation of their daily problems.

Goal of the project is to use natural language processing techniques and machine learning approaches to extract features from text and classify users with respect to their emotions.

Results of the project are to be demonstrated by a running demo application.

Technologies: Python, (Natural Language Toolkit (NLTK), Scikit-learn, Keras)

Contact: Nour Jnoub nour.jnoub@univie.ac.at
NJ3 - FactCheck – Semantic Analysis based on Linguistic Techniques

FactCheck is a framework for the detection and resolution of conflicting structured data on the Web. The FactCheck framework is an ongoing research at the multimedia research group. One of the central building blocks is the recognition of objects represented by means of structured data (encoded as Microformat or JSON-LD data) on different web pages, but constituting the same conceptual real world entity or artefact. This problem is also known as an entity recognition.

Goal of the project is to use natural language processing techniques to implement a suitable, alternative solution for entity recognition. Results of the project are to be demonstrated by a running demo application.

Technologies: Python, Natural Language Toolkit (NLTK), Scikit-learn and Google API.
Contact: Nour Jnoub nour.jnoub@univie.ac.at
NJ4 - A recommender system for buying commercial products

Most of the users nowadays like to buy their stuff online but to trust the sellers they mainly depend on the reviews to buy the products. The goal of the project is to build an application that assists the users to save their time by recommending the best item based on the reviews of other users. To ensure that for a such item different sources of different reviews should be collected and autonomously classified, clustered or fused to get a final result regarding the real users rate of such an item. This can be done by using natural language processing techniques and machine learning approaches to extract features and categorize similar reviews of users that may have same opinion or same content. Results of the project are to be demonstrated by a running demo application.

Technologies: Python, (Natural Language Toolkit (NLTK), Scikit-learn), Keras

Contact: Nour Jnoub nour.jnoub@univie.ac.at
NJ5 – The detection of spam comments on Web

Huge amount of fake accounts exist on the Web that may influence the quality of ratings, opinions or even distributing fake news.

The aim of this project is to extract user reviews from different sources like IMDB and define different rules or equations that can define the characteristics of such comments with respect to spatial, temporal or linguistic behavioural patterns. Then these rules should be encoded in a logic-based platform to detect such comments.

Technologies: Python and Answer Set Programming (ASP)

Contact: Nour Jnoub nour.jnoub@univie.ac.at
There is plenty of open source information on cyber threats provided in varying quantity and quality. So it is necessary to evaluate the quality of a source according to a given set of weighted parameters. The goal of this project is to develop a tool for open source threat intelligence sources evaluation based on an approach described in the paper “A Quantitative Evaluation of Trust in the Quality of Cyber Threat Intelligence Sources” (1).

Technologies: STIX (Structure Threat Information Expression), OpenIOC (Open Indicators of Compromise), standard programming languages

Contact: Gerald Quirchmayr, Gerald.Quirchmayr@univie.ac.at

(1) ACM ISBN 978-1-4503-7164-3/19/08; https://doi.org/10.1145/3339252.3342112
GQ2 – Attack Pattern Learning

Recent cyber attacks do frequently consist of a combination of attacks. Firewall logs do usually protocol single events. The goal of this project is to develop a tool that can combine these single events with respect to parameters such as target, time, and origin of an attack into a pattern. These patterns can then be used to detect and analyse future attacks. Neural network technology can serve as basis for learning form firewall logs.

Technologies: pattern learning with neural networks, standard programming languages

Contact: Gerald Quirchmayr, Gerald.Quirchmayr@univie.ac.at
GQ3 – File less Malware Detection

This relatively new operating type of malware first emerged in 2017. Early examples of this malware were Frodo, Number of the Beast, and The Dark Avenger. Try to simulate the malware in a virtual OS and implement a process to detect and remove it, using open source IDS technology as basis.

Technologies: IDS, standard programming languages

Contact: Gerald Quirchmayr, Gerald.Quirchmayr@univie.ac.at