37/17

readme alumni

Das Bulletin
der Alumni Informatik Universität Zürich
www.alumni.ch



Assessment und Matching – Wie komme ich zu meinen besten Treffern?

Konfigurierbare Asset Assessment und Matching Plattform 3
Software-Baukasten zur Realisierung von Vermittlungsanwendungen
Kornel Szabo

Address Matching 8
Ähnlichkeiten bei Daten und deren Strukturen für besseres Adressen-Matching
Michael Böhlen

Tinder for Artificial Intelligence 10
Analogien aus verschiedenen Profile Matchings für bessere Empfehlungen
Amancio Bouza

Rationaler Bauchentscheid 11
Profil-Assessment und Finden des besten Matches bei Teilwahrheiten
Christoph Hilber



editorial



Liebe Alumni-Mitglieder,
liebe Leserin, lieber Leser

In den vergangenen Jahren sind zahlreiche Online Vermittlungsplattformen entstanden. Täglich werden wir auf ettlichen TV- oder Online-Kanälen auf diese aufmerksam gemacht. Daneben nutzen auch länger schon Social Media Channels wie Facebook, Xing oder LinkedIn die Nutzerprofildaten für Ihre Communities oder Ihr eigenes Servicespektrum. Erfassen der Nutzerprofile (profiling, assessment) und Abgleich (matching) von diesen innerhalb der User Community sind zentrale funktionale Komponenten von solchen Online Plattformen.

In dieser Ausgabe des readme 37/2017 haben wir nun ebenfalls diese Konzepte und Entwicklungen aufgenommen und in vier Beiträgen aus unserer Alumni Community behandelt. Fragen wie „welche Taxonomien oder Assessment Kriterien müssen für eine Anwendungsdomäne berücksichtigt werden", „wie komme ich zu guten Matchings" oder „für welche anderen Anwendungsfälle ausser Nutzerprofil-Vermittlung resp. –matching können diese Konzepte auch zum Einsatz kommen" sind nur einige, die sich einem in diesem Kontext stellen.

Im Rahmen eines europaweiten Projektes, induziert durch das Involvement unseres Alumni Mitglieds Andri Färber, habe ich mit meiner Firma und weiteren Projektpartnern eine konfigurierbare Profil (oder eben Asset) Assessment und Matching Plattform entwickelt, anhand der schnell und kostengünstig entsprechende Anwendungsfälle realisiert werden können. Der Artikel gibt einen Überblick über das Projekt, dessen Ergebnisse und Marktpotentiale.

Amancio Bouza diskutiert in seinem Bericht, wie unter Anwendung von Artificial Intelligence Methoden auf unterschiedlichen Profilen bessere Empfehlungen und damit Matches generiert werden können.

Michael Böhlen zeigt auf, wie Adressen aus unterschiedlichen Quellen miteinander gematched und gemastered werden können. Dabei werden nicht nur Ähnlichkeiten in den Daten selber, sondern auch in deren Struktur fürs Matching herangezogen.

Christoph Hilber macht schliesslich als Abrundung der Diskussion deutlich, dass bei der Rekrutierung und Assessment von Job Kandidaten trotz ausgiebigem Datenmaterial und Tooleinsatz immer ein Restrisiko besteht, doch nicht den best match zu erwischen.

Unser Präsident Andri Färber berichtet wiederum über Vereins Interna wie beispielsweise die letzte Generalversammlung oder den Stand der Fusion unserer Alumni UZH Dachorganisation mit dem ZUNIV zur UZH Alumni.

Avi Bernstein, Direktor des Instituts für Informatik, berichtet über diverse Aktivitäten am und ums IfI. Forschungs- wie Industrieprojekte um die Topics Big Data, Maschinelles Sehen oder Blockchain-Technologien wurden lanciert und werden wohl für die Zukunft noch einige interessante Erkenntnisse und Partnerschaften hervorbringen.

Selbstverständlich berichten wir, untermalt mit Bildern, wiederum über vergangene Anlässe wie den traditionellen Fondueplausch im Le Dézaley vom Dezember sowie über die spannende Führung und aktive Begehung des Museums of Digital Art.

Zum Schluss danke ich allen, die an diesem readme 37/17 mitgearbeitet haben, sowie unseren Sponsoren und Partnern für deren Unterstützung und wünsche viel Vergnügen bei der Lektüre.

Herzlich, Euer
Kornél Szabó

Chefredaktor readme

Kontakt:
Kornél Szabó, szabo@alpnet.ch

Konfigurierbare Asset Assessment und Matching Plattform

Kornel Szabo

Im Rahmen eines europaweiten Projektes wurde eine konfigurierbare Asset Assessment und Matching Plattform entwickelt. Als Software-Baukasten kann diese beispielsweise dazu verwendet werden, schnell und kostengünstig Profil Assessment und Job Matching Anwendungen zu entwicklen. Der Artikel gibt eine Replik auf das Projekt sowie dessen Ergebnisse und Marktpotentiale.

Wie alles begann

Im 2014 wurde ich von einem Mitglied aus unserer Alumni Community angefragt, ob ich einen Software Provider kenne, der Erfahrungen in der Applikationsentwicklung und Durchführung von IT Projekten hat. „Na klar“, war meine Antwort, „und zwar wir, die Alpnet.“ Der Kontext der Anfrage war, dass in einem europaweiten Projekt ein Software-Baukasten entwickelt werden soll, um einfach, schnell und kostengünstig Anwendungsfälle für das Assessments und die Vermittlung von Skills- und Erfahrungsprofilen realisieren zu können. Ferner galt es, nebst daraus resultierenden Online Tools auch sogenante Offline Services als begleitende Massnahmen zu erarbeiten. Der Fokus lag dabei auf der Zielgruppe 50+ mit dem Fernziel, diese vermehrt in der Beschäftigung zu behalten und damit die soziale Inklusion zu unterstützen, die Gesundheit zu fördern sowie älteren Menschen ein (Zusatz-)Einkommen, auch über die Pensionierung hinaus, zu ermöglichen.

Da wir uns als Alpnet eh schon intensiver mit Fragen hinsichtlich Demographieproblematik und Pensionierungplanung beschäftigten (Rechner für Rente-oder-Kapital oder PK-Rechner inkl. Steuernberechnung), erachteten wir die Teilnahme an diesem Projekt als äusserst komplementär hinsichtlich unserer Absicht, dieses neue Geschätfsfeld weiter abzutiefen. Wir beschlossen, bei diesem Vorhaben unsere Erfahrungen sowie Skills einzubringen und unseren Betrag leisten zu wollen.

Über 40´000 Stunden Entwicklungsarbeit

Als erstes galt es, alle Projektpartner auf ein Ziel hin auszurichten. Dies bestand primär darin, eine IT Standard-Lösung als Online Service resp. Software Baukasten unter starkem Einbezug der später involvierten Stakeholders zu entwickeln. Die Online Lösung sollte zudem noch verstärkt werden um sogenannte Offline (begleitende) Services (bspw. Schulung, Coaching, Support & Beratung).

In einer ersten Phase wurde eine Anforderungs-Baseline erarbeitet, in der die Bedürfnisse der Zielgruppen und die Defizite aktuell existierender weltweit existierender Online Lösungen (rund 160) validiert und festgehalten wurden. Das dem Projekt zugrunde liegende Vorgehensmodell mit den avisierten Lieferobjekten ist in der nachfolgenden Abbildung kurz aufgezeigt.

Um am Schluss des Projekts nicht etwas erarbeitet zu haben, das keiner brauchen will, hat man bewusst von Beginn weg ein iteratives Vorgehen gewählt und appliziert. Via evolutionärem, iterativem Prototyping sowie Taxonomie- und Use-Labs mit potentiellen Nutzern und Nutzersegmenten wurden früh konkret fassbare Setups erarbeitet. Diese ermöglichten ein fundierteres Abtiefen der Spezifikationen und der Ziellösung.

So entstanden in der frühen Projektphase bereits GUI Mockups und Prototypen sowie erste Drafts einer Taxonomie, die für die Erfassung resp. Assessment der Profile auf Angebots- (persönliches Profil, Skills, Erfahrungen und Motivatoren resp. Passionen) und Nachfrageseite (Angebot von Jobprofiles resp. gesuchte Kandidatenprofile) sowie dann auch zum Matching von diesen dienten. Als wichtige Zwischenetappe bildete sich daraus eine Archetyp-Architektur und -Applikation, die als Basis für die zu realisierenden Pilot-Anwendungen fungierten.

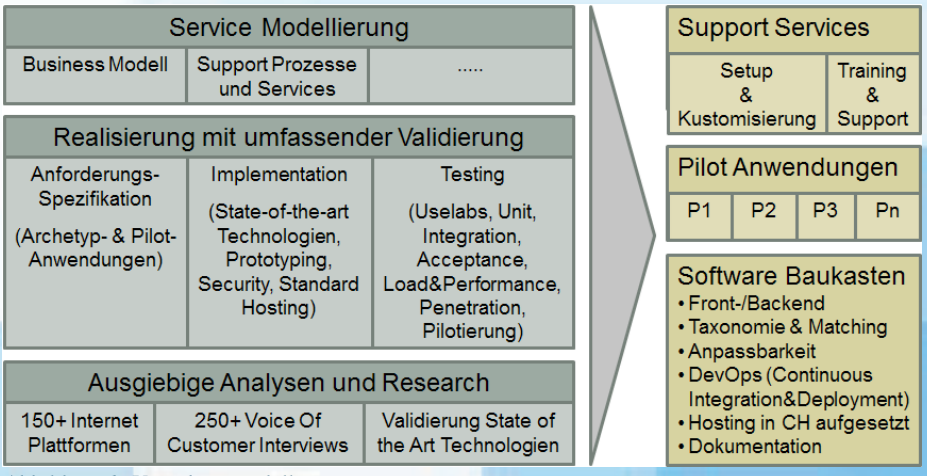


Abbildung 1: Vorgehensmodell

Fortsetzung auf Seite 4

Fortsetzung von Seite 3

Folgende 5 Pilot-Anwendungen resp. Anwendungsfälle, die innerhalb von 4-6 Wochen realisiert werden konnten, dienten dazu, den Software-Baukasten sowie die erarbeitete Assessment- und Matching-Taxonomie mit ausgewählten Nutzergruppen auf deren Funktionalität, Nutzbarkeit, Konfigurierbarkeit, Skalierbarkeit und Operabilität hin zu verifizieren:

1. Vermittlung von ehemaligen Führungskräften
2. Intra-organisationaler Erfahrungstransfer (eigene Erfahrungen zeitweise in anderen Unternehmensbereichen einsetzen)
3. Vermittlung und Reintegration von arbeitslosen 50+
4. Unterstützung und Vermittlung von Freiwilligenarbeit und Hilfsdiensten in Gemeinden
5. Coaching von KMUs bei Gründung, Weiterentwicklung oder Nachfolgeplanung

Aktueller Stand der Lösung: Konfigurierbarer Software Baukasten mit erprobter Taxonomie zum Assessment und Matching

Der durch das Projekt entstandene Software-Baukasten V1.3 wurde nach dem Projekt nochmals komplett überarbeitet hinsichtlich Funktionalität, User Experience (GUI Design, Usability, Interaktionsführung), Security Mechanismen sowie Last und Performance Verhalten.

Aktuell steht eine Version 2.1 mit einer Demo-Applikation online im Web zur Verfügung. An einer Version 2.2 wird bereits gearbeitet.

Die Implementation basiert auf einer 3 Tier Architektur. Sie beinhaltet eine Reihe von Core Services (bspw. User Management, Asset Management, Matching, Messaging), die mit anwendungsfallspezifischen Komponenten und Funktionalitäten angereichert resp. erweitert werden können. Als Basis für die Entwicklung dient eine Archetyp-Applikationen, die einem erlaubt, bei der Entwicklung bereits auf hohem Niveau starten und damit schneller zum Ziel, sprich fertigen Zielanwendung, gelangen zu können.
Es wurde bei der Validierung und Einsatz der technischen Komponenten von Beginn weg darauf geachtet, dass die Lösung europaweit bei einem beliebigen Hosting-Provider, der eine Java oder Virtual Machine basierte Anwendungsplattform unterstützt, betrieben werden kann. Im Rahmen des Projekts wurden beispielsweise 4 verschiedene Provider für die Pilotanwendungen eingesetzt. Dies ohne grossen Aufwand für die jeweilige Installation und Feinabstimmung.

Grosses Interesse

Schon während dem Projekt waren die Reaktionen auf das Vorhaben und die erarbeiteten Ergebnisse seitens Profil-Anbieter, -Nachfrager und potentieller Betreiber von solchen Plattformen äusserst positiv. Dass die Lösung, gestützt durch die entsprechende Taxonomie, Funktionalität und Interaktionsmechanismen, nicht nur die - bei heutigen Job-Portalen üblichen - Hard-Skills (was kann ich/habe ich gelernt), sondern auch Soft-Skills (Sozialkompetenzen, Soft Factors) als auch die persönlichen Ziele, Motivationen und Präferenzen (Motivation- oder Passion-Skills) inkorporiert und beim Assessment sowie Matching berücksichtigt, fanden alle als innovativ und sehr bereichernd. V.a. bei Personen 50+, bei welchen mehr deren Erfahrung als die initiale Ausbildung zum Tragen kommen, stiess dies auf sehr positives Echo. Wertvolles Erfahrungswissen sowie verborgene Talente werden mit der Software sichtbar gemacht. Nebst den Stimmen aus der europaweiten Pilotierung in der Schweiz, Österreich, Deutschland, Italien und Ungarn präsentierten wir unsere Lösung auch diversen Schweizer Firmen (Versicherungen, Personaldienstleistern, Telcos), Branchen- und Dach-Verbänden, Bundes- und Kantonal-Behörden sowie Vereinen.
Auch bei diesen war und ist das Interesse und Feedback sehr positiv. Ein grosser Versicherungskonzern bekundete u.a. auch die Absicht, in einem Folgeprojekt wieder als Pilot-Nutzer mitmachen und die bereits gewonnen Erkenntnisse weiter abzutiefen zu wollen.

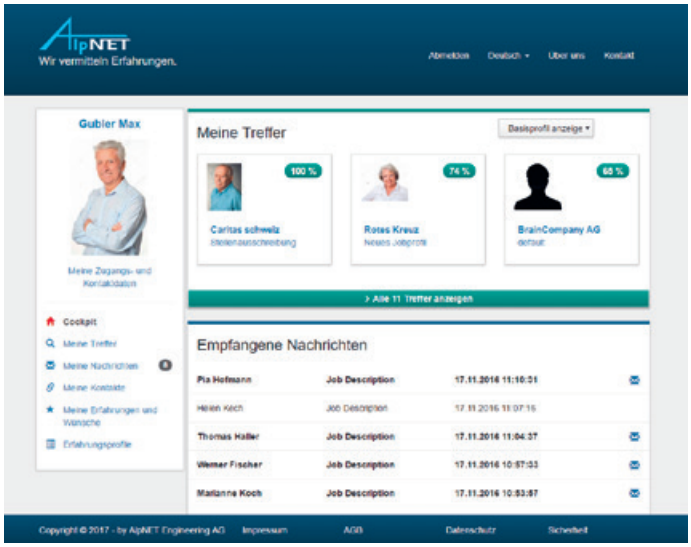


Abbildung 2: Cockpit

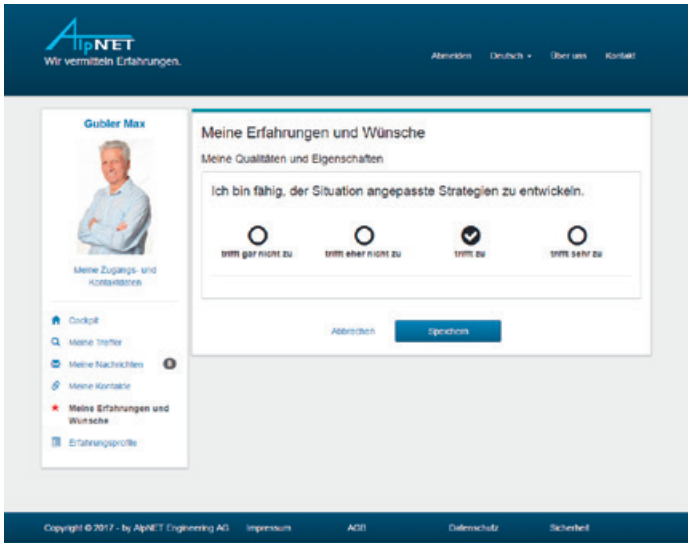


Abbildung 3: Profil Assessment

Zwischen-Fazit und Next Steps

Mit der aktuell vorliegenden Version 2.2 liegt ein Software-Baukasten vor, mit dem sich schnell und kostengünstig Anwendungsumgebungen umsetzen lassen (Piloten wurden innert 4-6 Wochen realisiert). Anbieter und Nachfrager von Assets (bspw. Personalprofile) können damit anwendungsfallspezifisch die Profile erfassen (assessment) und gegeneinander matchen lassen. Die jeweils zugrundeliegende Taxonomie mit ihren Kriterien und Gewichtungen, die Spracheinstellung für die Nutzerinteraktion sowie GUI Elemente lassen sich flexibel konfigurieren.

Das gewählte und gefahrene Vorgehensmodell im Projekt hat sich soweit bewährt. Wichtig ist, dass ein solches länderübergreifendes Vorhaben aber mit Partnern erfolgt, die einen gewissen praxiserprobten Reifegrad in der Durchführung von IT Projekten aufweisen.

Die zahlreichen Feedbacks aus den europaweit pilotierten Anwendungsfällen und Präsentationen in der Schweiz sowie die Validierung des avisierten Business Modells durch entsprechende Experten stimmen uns sehr zuversichtlich, das Vorhaben mit Elan weiter zu treiben. Das Bedürfnis- und Nutzen-Potential für unsere Lösung wurde uns in zahlreichen Gesprächen mit diversen Anspruchsgruppen mehr als bestätigt. Es gilt nun, die Lösung mit Kunden und Anspruchsgruppen weiter zu entwickeln und neue, noch ungeahnte Anwendungsfelder zu identifizieren.

Zu guter Letzt sei hier nochmals auf die Power unserer Alumni Community hinzuweisen, über dessen Triggering wir ein tolles und zukunftsträchtiges Vorhaben zusammen realisieren konnten.

Yes, our Alumni network rocks!



Dr. Kornel Szabo *studierte Wirtschaftsinformatik an der Uni Zürich und promovierte in Informatik am Institut für Informatik Uni Zürich. Kornél Szabó ist Gründer und CEO der Alpnet Engineering AG und verfügt über langjährige Erfahrungen im Programm- und Projektmanagement, in der Realisierung sowie Betrieb von heterogenen, standort- und länderübergreifenden Projektstrukturen und Anwendungsumgebungen.*

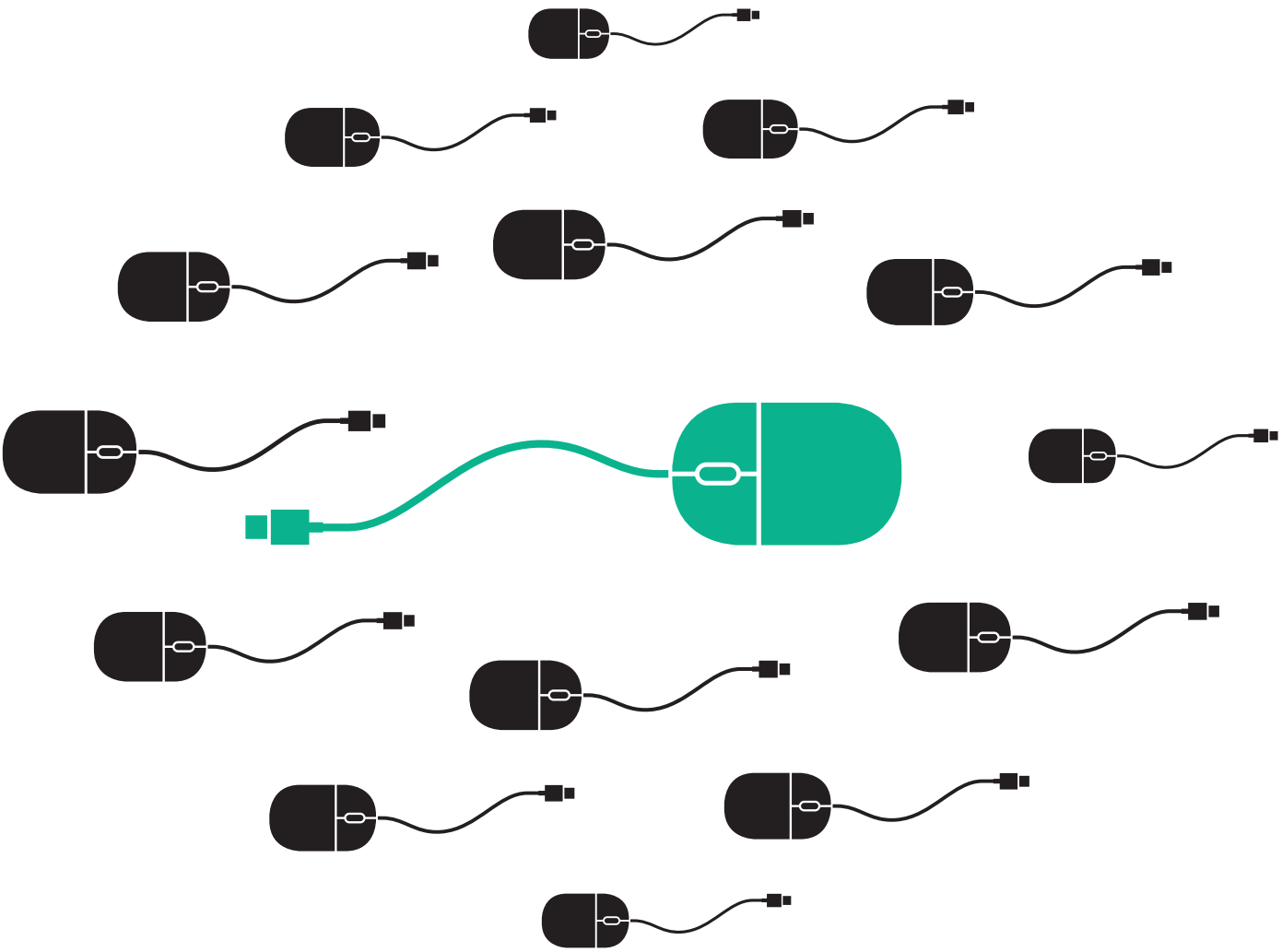
Kontaktadresse:

Dr. Kornel Szabo
Alpnet Engineering AG
Grundhaldenstrasse 60
CH-8303 Bassersdorf
Tel: +41 44 360 80 30
E-Mail: szabo@alpnet.ch
http://alpnet.ch

IT Jobs

What if you're right

adnovum.ch/jobs



AdNovum

High-End Software & Security Engineering

In eigener Sache: News aus dem Vorstand

Liebe Alumni-Mitglieder,
liebe Leserin, lieber Leser

Zur 20. Ordentlichen Generalversammlung waren wir bei Microsoft Schweiz in Wallisellen zu Gast. 48 Mitglieder folgten der Einladung und nahmen am Anlass teil. So viele waren wir schon länger nicht mehr! Die GV verlief, wie meistens, unspektakulär. Eine ansatzweise kontroverse Diskussion entwickelte sich rund um das Thema der lebenslangen Mitgliedschaft. Diese wurde im Vorfeld der GV durch die Mutter eines Absolventen angestossen, welche ihrem Spross eine solche Mitgliedschaft zum Uni-Abschluss schenken möchte. Der Vorschlag des Vorstands war es, die lebenslange Mitgliedschaft für CHF 1‘200 anzubieten. Dem Antrag wurde nach einigem Hin und Her mit 37 zu 8 Stimmen bei 3 Enthaltungen stattgegeben.

Aufgrund familiärer Verpflichtungen ist Christoph Kiefer von seinem Amt als Aktuar und „Datenaufbereiter“ für das who-is-who zurückgetreten. Nachdem schriftliche Aufrufe für seine Nachfolge nichts gefruchtet haben, erwies sich ein eindringlicher Appell an der Sommernachtsparty als erfolgreich. Es haben sich in der Folge gleich zwei Mitglieder für die Vakanz beworben. Der Arbeitsumfang im Vorstand hat in den letzten Jahren tendenziell zugenommen, während gleichzeitig die dafür vorhandene Kapazität der Vorstandsmitglieder - unserer zunehmend hektischen Zeit geschuldet - immer knapper wird. Der Vorstand entschloss sich deshalb, auf den statutarisch vorgesehenen Maximalbestand von neun Personen aufzustocken und beantragte der Versammlung die Wahl sowohl von Anja Cavigilli als auch von Oliver Carl. Beide wurden einstimmig gewählt. Da Abraham Bernstein, der neue Direktor des Instituts für Informatik, leider an der GV-Teilnahme verhindert war, sprang Martin Glinz, sein langjähriger Vorgänger, nochmals ein. Das ermöglichte uns, Martin Glinz vor versammeltem Publikum für seine vielen Beiträge und seine lückenlose Präsenz an der GV zu danken und ihn gebührend zu verabschieden. Nachdem uns Microsoft vor der GV in zwei Gruppen durch die neu gestaltete Bürolandschaft führte und uns an sehr interessanten Vorträgen teilhaben liess, schloss Microsoft den Anlass mit einem spannenden Vortrag von Marc Schöni über „Speech Machine Learning“ und einem anschliessenden Apéro. Nach einem intensiven Abend machten wir uns geistig und körperlich gesättigt zufrieden auf den Heimweg.

Die Vorbereitungen zur Fusion von ZUNIV und der Alumni Dachorganisation laufen derzeit auf Hochtouren. So wie es aussieht, wird an der Generalversammlung von Alumni UZH im Mai 2017 die Fusion beantragt. Der neue Name der zusammengelegten Organisation wird voraussichtlich „UZH Alumni“ sein. Als Folge der Fusion wird es ein einheitliches CI/CD für die Chapters und Mitgliedsvereine geben. Damit werden wir dann endlich ein „richtiges“ Logo erhalten und auch die Uni-Türme in unserer Kommunikation verwenden dürfen. Darüber hinaus wird sich für uns als „Alumni Informatik Universität Zürich“ wenig ändern.

Eine Änderung wünschen wir uns zusammen mit den OEC Alumni und den EMBA Alumni allerdings schon: Wir alle sind mit der Intraworlds-Plattform (erreichbar auf unserer Website alumni.ch via MyAlumni Menüpunkt) ziemlich unzufrieden. Lange Antwortzeiten und ein unbefriedigendes User-Interface machen Mitgliederbewirtschaftung, Event-Administration und E-Mail-Versände zur Qual. Es wurde deshalb auf der Ebene der Alumni-Dachorganisation eine Arbeitsgruppe zur Re-Evaluation der Informatik-Plattform gebildet, in der verschiedene Stakeholder vertreten sind. Je nach Ausgang der Re-Evaluation soll ggf. eine neue Plattform beschafft und eingesetzt werden. Falls jemand von Euch geeignete Software-Applikationen kennt, bin ich für „zweckdienliche Hinweise“ sehr dankbar.

Da wir dieses Jahr das 20-jährige Jubiläum feiern, soll die Sommernachtsparty etwas Besonderes werden. Wir begehen den Anlass am Samstag, 9. September, im Lake Side am schönen Zürichsee und haben - so viel sei schon verraten - diesmal Fabian Unteregger als Special Show Act zu Gast. Am besten markiert Ihr den Anlass jetzt schon dick und farbig in Eurer Agenda.

Nach dem Erfolg der letzten GV bei Microsoft meinte Alain Gut, dass IBM das mindestens ebenso gut kann und lud uns für die nächste GV an die Vulkanstrasse nach Zürich-Altstetten ein. Wir dürfen uns also jetzt schon auf eine spannende GV am Montag, 6. November, freuen.

Herzlich und bis bald an einem unserer Anlässe,
Euer

Andri Färber



Andri Färber
Präsident der Alumni Informatik
Universität Zürich

Kontakt:
Andri Färber
Tel.: +41 79 642 08 85
E-Mail: andri.faerber@alumni.uzh.ch

Address Matching

Michael Böhlen

Databases that have been built independently often store information about the same or related objects in the real world. To enable interaction between these databases, data items that refer to the same object must be identified. Residential addresses are data of particular interest as they allow to connect related pieces of information in different databases. A good matching of addresses is possible if besides the similarity of the individual addresses also the similarity of the structure of streets is considered.

Residential addresses often provide the only link between information stored in different databases. Matching residential addresses, however, is a challenging task. Consider Figure 1 with the databases from the Electricity Company and the Registration Office from the Municipality of Bolzano. Both databases cover the same geographic area, but an exact match on the address attributes obviously fails.

Electricity Company (EC)

address	bill
Hermann-v-Gilm 1	€ 121
Hermann-v-Gilm 3/A	€ 71
Hermann-v-Gilm 3/B	€ 63
Hermann-v-Gilm 6	€ 0
Siegesplatz 2/A	€ 98
Siegespl 3/-/1	€ 32
Siegespl 3/-/2	€ 51
Siegespl 3/-/3	€ 43
Friedhofpl 4	€ 143
Friedhofpl 6	€ 0
Untervigli 1	€ 117
Mariengasse 1	€ 161

Registration Office (RO)

address	resident
Gilmstrasse 1	Peter
Gilmstrasse 3	Hans
Gilmstrasse 3	Renate
Gilmstrasse 3	Max
Gilmstrasse 5	Arturas
Friedenspl 2/A/1	Markus
Friedenspl 2/A/2	Klaudia
Friedenspl 3	Igor
Cimitero 4	Linas
Cimitero 6/A	Franz
Cimitero 6/B	Romans
Untervigil 1	Andrej
Marieng. 1/A	Josef

Figure 1: Two Databases with Residential Addresses

The matching faces two key challenges. First, streets must be matched even if different databases use different names for the same street (e.g., due to misspellings, different coding conventions, or renamed streets). Second, addresses that are stored with different granularity must be matched correctly, although there is no one-to-one correspondence between them.
To get a good matching we need a structure-aware distance measure between two streets that includes the name and addresses of the streets. To get this we model the residential addresses of streets as ordered labeled address trees as illustrated in Figure 2.

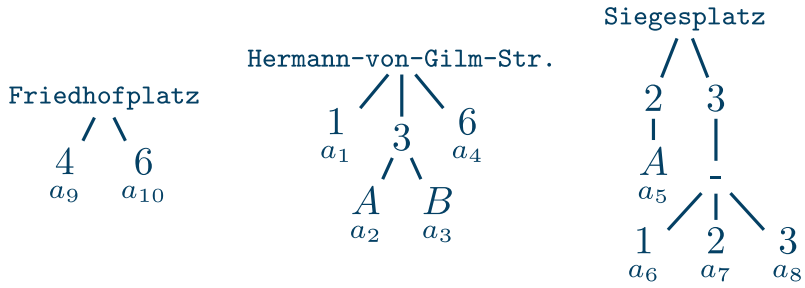


Figure 2: Example Address Trees

The root of address trees are the street names, while the rest of the tree represents house numbers, entrance numbers, and apartment numbers. The street distance depends on the structural similarity of the address trees and the similarity of the street names. This allows us to match streets even if the names are completely unrelated or if the structure of the address trees is ambiguous.

Step 1: Street Distances

To compute the similarity of two streets we define a street distance, which is based on two independent characteristics: the name of the two streets and the structure of the addresses of the two streets. Depending on the input data, more reliable matches can be expected from either the name distance (e.g., both address sets use the same language and similar coding conventions) or the structure distance between the address trees (e.g., the input sets use different languages). We weight name and structure distance, and combine the distances into a single distance between address trees.

Name Distance: The root node of an address tree is the name of the street. The name distance between two address trees is the minimum distance between their names. We use the q-gram distance [1, 3, 4] to determine the distance between street names. The q-grams of a street name are all its substrings of length q. Two street names

are similar if they have many q-grams in common.

Structure Distance: The structure distance of two streets quantifies the difference between the addresses of two streets. Such a measure can be defined as the structural distance between two address trees. We use pq-grams to measure the distance between two trees [2]. A pq-gram is a small, besom-shaped subtree consisting of an anchor node, p – 1 ancestors, and q consecutive children. Intuitively, the pq-grams are formed by shifting a pq-gram shaped pattern over the tree.

Step 2: Matching Streets

After the distances between all street pairs have been determined, the streets must be matched. We define the matching as a set of street pairs, where each street appears in at most one pair. Our goal is to compute a stable matching. Intuitively, a matching is stable if it is not possible to break up existing matches and form a new match, such that the new match is better than the old matches for both matching partners. Let D be the M×N distance matrix that stores the distances between streets str(A) = $\{\alpha_1,...,\alpha_M\}$ and streets str(B) = $\{\beta_1,...,\beta_N\}$. The distance between streets αi and βj is stored in row i and column j of D. Figure 3 shows an example of a distance matrix. The correct matches are shaded.

A constraint of the matching is that a street can have at most one matching partner. A threshold-based approach will not work since the same threshold may be too high for some streets (they are matched multiple times), but too low for other streets (they remain unmatched). A local greedy approach matches a street with its nearest neighbor and matched streets are not matched again. The quality of such a matching depends on the matching order and bad choices in the beginning may prevent good matches later. We therefore use a global greedy matching algorithm that matches each street to at most one other street and is stable.

Step 3: Matching Addresses

The final step is to match the addresses of matched streets. Two addresses should be matched if they refer to the same location. It is not enough to check whether the relative parts of the addresses are equivalent as the addresses may be stored with different granularity. For example, 'Gilmstrasse 3' should match both 'Hermann-v-Gilm 3/A' and 'Hermann-v-Gilm 3/B', but the entrance is not specified in 'Gilmstrasse 3'. Thus, there is no one-to one correspondence between all addresses of two streets since they might be stored with different granularity.

In summary, matching residential addresses is important to link information stored in independent databases and is a non-trivial problem. We presented an approach that considers both the structural similarity of the addresses of a street and the similarity of street names. This yields a flexible solution that produces high quality matchings.

References

1. Arvind Arasu, Venkatesh Ganti, and Raghav Kaushik. Efficient exact set-similarity joins. In Proceedings of the International Conference on Very Large Databases (VLDB), pages 918–929. VLDB Endowment, 2006.

2. Nikolaus Augsten and Michael Böhlen. Similarity Joins in Relational Database Systems. Synthesis Lectures on Data Management. Morgan & Claypool Publishers, 2013.

3. Nikolaus Augsten, Michael H. Böhlen, Johann Gamper, The address connector: noninvasive synchronization of hierarchical data sources. Knowl. Inf. Syst. 37(3): 639-663 (2013).

4. Sunita Sarawagi and Alok Kirpal. Efficient set joins on similarity predicates. In Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD, pages 743–754, New York, NY, USA, 2004. ACM.

5. Chuan Xiao, Wei Wang, and Xuemin Lin. Ed-Join: An efficient algorithm for similarity joins with edit distance constraints. In VLDB, 2008.

	β_1	β_2	β_3	β_4	β_5
α_1	1.0	0.7906	0.8186	1.0	0.9794
α_2	0.8720	1.0	0.9785	0.9201	0.9581
α_3	1.0	0.7290	1.0	1.0	1.0
α_4	0.8091	0.9817	1.0	0.7071	0.5946
α_5	0.9055	1.0	0.9765	0.3536	0.8339

Figure 3: Distance Matrix for Two Sets of Streets



Prof. Dr. Michael H. Böhlen *is a professor of computer science at the University of Zürich where he heads the database technology group. His research interests include various aspects of data management, and have focused on time-varying information, data warehousing and data analysis, and similarity search. He received his M.Sc. and Ph.D. degrees from ETH Zürich in respectively 1990 and 1994. Before joining the University of Zürich he was a faculty member at Aalborg University in Denmark and the Free University of Bozen-Bolzano in Italy.*

Kontaktadresse:

Prof. Dr. Michael Böhlen
Institut für Informatik
Universität Zürich
Binzmühlestrasse 14
CH-8050 Zürich
Tel: +41 44 635 43 33
E-Mail: boehlen@ifi.uzh.ch
http://www.ifi.uzh.ch

Tinder for Artificial Intelligence

Dr. Amancio Bouza

Stopp! Haben unsere berühmten Watsons, HALs, Marvins, C3POs und Mr. Datas dieser Welt etwa romantische Gefühle? Nein, zumindest noch nicht. Und doch geht es um Profile Matching von Artificial Intelligences



Dr. Amancio Bouza *arbeitet bei der ipt, Innovation Process Technology AG, eine Schweizer IT Consulting Boutique. Als Intrapreneur arbeitet er an der API Produktentwicklung von Unternehmen und den Anwendungen von AI und Machine Learning für die Digitale Transformation von Geschäftsmodellen. Zuvor gründete er ein Tech-Startup im Bereich Crowdsourcing mit. Er schloss sein Wirtschaftsinformatik Studium 2008 an der Universität Zürich ab. Danach doktorierte er bei Prof. Avi Bernstein und Prof. Harald Gall im Bereich von Empfehlungssystemen und Machine Learning.*

Kontaktadresse:

Dr. Amancio Bouza
Innovation Process Technology AG
Poststrasse 14
CH-6300 Zug
Tel: +41 79 387 69 35
E-Mail: amancio.bouza@ipt.ch
http://www.ipt.ch

Tinder ist eine mobile App, mit welcher weltweit über 50 Mio Benutzer nach neuen Bekanntschaften suchen. Dabei nutzt Tinder die Facebook-Profile ihrer Nutzer. Ein Nutzer bekommt die Profile anderer vorgezeigt. Er entscheidet, ob ein Profil interessant für ihn ist oder er doch schnell zum nächsten Profil springt. Wenn das Gegenüber das entsprechende Profil auch interessant findet, kommt es zu weiteren Interaktionen.

User Profile

Auf anderen Plattformen bestehen unsere User-Profile aus einer Reihe von Attributen. So besteht unser LinkedIn-Profil aus einer Reihe von Skills, Job-Positionen und Erfahrungen. LinkedIn vergleicht unsere Profile mit denen von Job-Profilen. Wenn beides matched, haben wir eine Job-Empfehlung im Posteingang. Dazu reicht ein effizienter Abgleich von Attribut-basierten Profilen.
Unser Kunden-Profil auf Amazon besteht unter anderem aus einer Reihe von Kategorien und Themen, die für uns spannend sind, und aus unseren gekauften Produkten und Bewertungen. Amazon ist ein Vorreiter für Empfehlungsmaschinen. Sie nutzen die Ähnlichkeiten von Kundenprofilen, um uns das zu empfehlen, was Kunden mit ähnlichen Profilen bzw. Geschmack bereits gekauft haben. Diese Empfehlungsmethode nennt man Collaborative Filtering. Collaborative Filtering basiert darauf, Kunden mit ähnlichen Profilen zu finden.
Soweit so gut. Amazon kennt unsere Kundenprofile. Aber was, wenn wir bei Exlibris einkaufen möchten? Auf Exlibris werden wir vergeblich nach einer Amazon-Einkaufshistorieimportfunktion suchen.

User Profile als AI Profile

Was wäre aber, wenn wir ein System intelligent genug machen könnten, welches unseren Geschmack basierend auf unseren Einkäufen und Bewertungen lernen und es auf Plattformen wie Amazon und Exlibris anwenden könnte? Diese Technologie existiert bereits: Machine Learning. Machine Learning ist ein erfolgreicher Ansatz für Artificial Intelligence (AI). Es ist ein Oberbegriff für die künstliche Generierung von Wissen aus Erfahrungen. Eine AI kann aus Beispielen lernen, erkennt daraus Muster und Gesetzmässigkeiten, und kann diese auf unbekannte Daten anwenden.
Unser Verhalten kann von einer AI gelernt werden. Durch die Erfahrung aus unserem Kaufverhalten und Bewertungen kann es mittels Machine Learning unseren Geschmack lernen und für uns nach interessanten Produkten suchen, egal auf welcher Plattform.
Solche AIs haben einen grossen Nachteil. Sie finden genau das, was sie gelernt haben zu finden. Solche Empfehlungsmethoden nennt man Content Filtering. Aber wir Menschen möchten gerne auch wieder mal etwas Frisches erleben, etwas Neues. Collaborative Filtering Methoden, wie die von Amazon, sind dafür hervorragend geeignet. Amazon vergleicht Warenkörbe und Bewertungen von Kunden. Je ähnlicher die Warenkörbe und Bewertungen desto ähnlicher der Geschmack. Doch wie vergleicht man Als bzw. künstlich generiertes Wissen zu Geschmäckern?

Vergleich von AI Profilen

AIs können auf zwei verschiedene Arten verglichen werden. Die erste Art lässt zwei AIs einen spezifischen Warenkorb von Produkten bewerten.

Je ähnlicher die Bewertungen, desto ähnlicher die beiden AIs. Die Herausforderung besteht hierbei in der Wahl der zu bewertenden Produkte im Warenkorb. Denn sie haben einen direkten Einfluss auf das Resultat und führen meist zu einem Bias.
Die zweite Art vergleicht das künstlich generierte Wissen bzw. die Muster und Gesätzmässigkeiten. Je nach verwendetem Machine Learning Algorithmus bedeutet dies ein Vergleich von Entscheidungsbäumen, probabilistische Graphen und Bayes'sche Netze, etc. Wir konnten zeigen, dass insbesondere die zweite Art zu besseren Ergebnissen führt als andere State-of-the-Art Empfehlungsmethoden. Wir können also Profile-Matching von AIs nutzen, um u.A. bessere Empfehlungen auf einer Plattform und sogar Plattform übergreifend zu erhalten. Das ist Tinder for Artificial Intelligence.

Rationaler Bauchentscheid

Christoph Hilber

Das Ziel ist klar und einfach formulierbar: Den Menschen finden, der ideal auf eine Position passt. Das ist Tagesgeschäft in der Rekrutierung, wo Beurteilungen und Abgleich mit Profilen eine der wichtigsten Aufgaben darstellt. Und doch gibt es Fehlbesetzungen. Eine Smart Watch, welche den Selektionsprozess unterstützt, gibt es leider noch nicht.

Super Job!

Ausgangslage ist die Tätigkeitsbeschreibung der Position. Wenn sie vorhanden ist, beschreibt sie die aktuellen, fachlichen Tätigkeiten. Hinzu kommen Internas wie spezielle Herausforderungen, Gründe für Neubesetzung etc. HR entwickelt ein attraktives Inserat für die externe Kommunikation, indem einige Anforderungen an die Persönlichkeit angefügt werden. Ansonsten wird das Inserat spannend tönen und Job sowie Firma von der besten Seite beleuchten. Legitime Teilwahrheit.

Ich bin der Beste!

Der Bewerber durchläuft einen ähnlichen Prozess. Er kehrt in sich selbst, macht eine Liste der Erfolge und Misserfolge in seinen bisherigen Jobs, seiner Stärken und Schwächen, seiner Wünsche und Ziele. Die positiven Elemente verpackt er in einen Lebenslauf, ergänzt ihn vielleicht mit ein paar unproblematischen Schwächen. Problemzonen werden weggelassen. Das übliche Motivationsschreiben umfasst hoffentlich drei gute Argumente, warum sich ein persönliches Gespräch lohnen würde. Legitime Teilwahrheit.

Assessment – die ganze Wahrheit!

Die Teilwahrheiten treffen aufeinander. Dieser Schritt ist vergleichbar mit dem M&A-Prozess bei einem Firmenkauf. Beide Seiten stellen sich möglichst gut dar und lassen das Gegenüber nach den Schwachstellen suchen oder nach Informationen fragen, welche in den Dokumenten schöngeredet oder verschwiegen werden.
Die fachlichen Themen wie technische Fähigkeiten oder betriebswirtschaftliche Zusammenhänge können im Gespräch oder durch Case Studies, Tests, vielleicht Schnuppertage, verifiziert werden.
Je höher und unternehmerischer eine Funktion ist, desto weniger ist dies jedoch möglich, da die Persönlichkeitsmerkmale wichtig werden wie z.B. der Denk-, Arbeits- und Beziehungsstil. Dem gegenüber stehen Themen wie Kultur, Offenheit oder Stimmung im Unternehmen. Diese Dimensionen werden in der Praxis über eine Gesprächsbatterie mit mehreren Peers, Vorgesetzten oder falls möglich mit Direktunterstellten ausgelotet. In diesen Gesprächen gewinnen beide Parteien Entscheidungsgrundlagen, denn auch der Bewerber lernt durch die Gespräche sein zukünftiges Umfeld näher kennen.



Christoph Hilber *ist lic.oec.publ. und seit 9 Jahren Personalberater mit seiner eigenen Firma P-CONNECT – Coaching – Searching - Consulting mit Fokus auf Industrie (MEM), Informatik, Telekom und Positionen VR, GL/Kader und Spezialisten. Die Dienstleistungen umfassen Outplacement, Executive Search, Change Management. Vorgängig war er in leitenden Linienfunktionen bei NCR/AT&T, diAx und Siemens.*

Kontaktadresse:

P-Connect GmbH
Christoph Hilber
Managing Partner
Stampfenbachstrasse 57
CH-8006 Zürich
Tel.: +41 044 211 05 00
E-Mail: christoph.hilber@p-connect.ch
http://www.p-connect.ch

Fortsetzung auf Seite 12

Fortsetzung von Seite 11

Abklärung dieser sehr persönlichen Merkmale. Oft werden Einzel-Assessments durchgeführt, wo der Bewerber einem Team von Psychologen ausgesetzt wird – sicher ein guter Stresstest. Eine weitere und schnellere Option sind On-line Assessments, in welchen Verhalten, Motivation und Kompetenzen gecheckt werden können. Beispiele dafür sind INSIGHTS MDI für Verhalten und Motivation, ASSESS by Scheelen für die Persönlichkeitskompetenzen.
Beide Verfahren bringen etwas Konkretisierung in die schwer greifbaren Dimensionen der menschlichen Seele. Alle Assessments dürfen jedoch nur einen Mosaikstein im Entscheidungspuzzle darstellen, sind diese doch nur Momentaufnahmen und abhängig von der momentanen Gefühlslage eines Bewerbers. Unter Stress denkt der Mensch anders als ohne Stress oder mit Longdrink unter der Palme. Trotzdem sind sie hilfreich, um allenfalls kritische Punkte in Gesprächen weiter zu verifizieren.

Matching – der Entscheid.

Smart-Watches können inzwischen sogar die Restlaufzeit des Lebens berechnen oder langweilige Gesprächspartner identifizieren. Genügen würde es, wenn sie anzeigen könnten, wenn der Gesprächspartner nicht die ganze Wahrheit sagt, denn dort könnte man weiter bohren. Solange dies nicht möglich ist, bleibt nichts anderes übrig, als möglichst viele und breit abgestützte Informationen zu kompilieren – systemisch und/oder mit gesundem Menschenverstand – und dann auf das Gefühl, die Bauchgefühle der involvierten Gesprächspartner, zu setzen – Q.E.C. – quod erat convincendum.

impressum

readme Nr. 37/März 2017

Herausgeberin:
Alumni Informatik

Chef-Redaktor:
Dr. Kornel Szabo

Redaktion:
Nicola Fantini
Bruno Kaiser
Andri Färber
Dr. Esther Kaufmann
Carl Oliver
Esther Rölli

Adresse:
Alumni Informatik Universität Zürich
Schönberggasse 15a
8001 Zürich

Homepage:
www.alumni.ch

Redaktionsadresse:
readme@alumni.ch

Redaktionsschluss Nr. 38/September 2017
27. August 2017

Titelbild:
pixabay.com

Auflage:
866 Exemplare

Produktion und Druck:
ETIDRUCK AG, Hausen am Albis

Anlässe

Museumsführung im Museum of Digital Art (MuDa)

Fast pünktlich und auch fast vollständig konnten wir unsere Reise in eine digitale Kunstwelt starten. Caroline Hirt, die Direktorin des Museums, erläuterte mit einer spürbaren Leidenschaft die Entstehungsgeschichte des Museums. Nicht nur die Suche nach einer geeigneten Lokalität hat sich aufwändig gestaltet und dann glücklich gefügt, auch die finanziellen Mittel wurden mit viel Unterstützung von Museumsbegeisterten ermöglicht. Der enorme Aufwand, ein neues Museum zu gründen, und der nötige Papierkrieg dazu ist jetzt bereits ein Jahr her und die Begeisterung der Mitarbeitenden ist unverkennbar. Was aber die Ausstellungen betrifft sind die Teilnehmenden zu Beginn noch skeptisch...
Das Museum ist klein, die Ausstellungsobjekte stellen meist "Etwas-mit-einem-Beamer" dar, bewegen sich und hie und da geben sie Geräusche von sich...
Die leidenschaftliche Präsentation der Künstler und deren Ausstellungen durch Caroline geht weiter und die Alumnis lassen sich auf dieses meist unbekannte Terrain ein. Wir werden nicht enttäuscht!

Jede der Ausstellungen hat ihren eigenen Charme und erwacht zu digitalem Leben, wenn man sich mit ihr auseinandersetzt. Das Kind im Betrachter wird auf- und herausgefordert und da lassen wir uns natürlich nicht zweimal bitten :-)
Hände, die Bewegungen festhalten. Tennisbälle, die sich an den Händen halten und tanzen. Ein kleines Windrad, dass den Orkan ruft. Ein Vogel, den man über Musik fliegen lassen kann. Eine Schaukel auf dem Mond. Genial, eindrücklich und definitiv ein Besuch wert!

Diese Ausstellung LAB212 ist bereits am 26.02.2017 beendet, somit sind alle Teilnehmenden gespannt, welche Überraschungen im April die Museumsbesucher erwarten. Für viele war dies wohl nicht der letzte Besuch im MuDA!







Fondueplausch im Le Dézaley



Am traditionellen Fondueplausch im Restaurant Le Dézaley am 13. Dezember 2016 nahmen 74 Personen teil, darunter erfreulicherweise auch 10 Vertreter unserer Sponsoren Goldwyn Partners Group, NSM Engineering, PRO ACADEMICS und Avaloq. Auch dieses Jahr haben wir über den ICU 12 Studenten zu einem ermässigten Preis eingeladen, um den Kontakt zur Studentenschaft auch auf diesem Weg zu pflegen.
Die Stimmung war einmal mehr hervorragend und der Lärmpegel entsprechend hoch. Dank speditivem Service und mehr Platz löste sich der Anlass dieses Jahr früher als gewohnt auf.

Neues aus dem IfI

Prof. Dr. Abraham Bernstein, Institutsdirektor

Liebe Alumni

Grosse Datenmengen sind in aller Munde: Die Politik setzt sich mit dem Datenschutz auseinander, Unternehmen sammeln das neue Öl der Datengesellschaft und selbst die NZZ am Sonntag widmet dem Thema eine Doppelseite.

Auch an der UZH ist Big Data ein grosses Thema. So haben wir es auch schon im Oec Magazin der Wirtschaftswissenschaftlichen Fakultät diskutiert. Schon dort wurde gesagt: Mit Big Data kommt auch Big Responsibility. Die neuen Datenmengen und Analyseverfahren führen auch zu neuen Verantwortlichkeiten. Wir müssen uns nicht nur fragen, was technisch machbar und rechtlich vertretbar, sondern auch was ethisch korrekt ist. Genau dies ist Gegenstand einer neuen Vorlesung zum Thema Big Data: Technology, Law and Ethics, welche wir zusammen mit der juristischen und medizinischen Fakultät anbieten. Auch die DSI nimmt sich des Themas an. So durften wir erst gerade im Januar einen äusserst interessanten Abend verbringen zum Thema Ethik in der modernen Kriegsführung. Professor Ron Arkin (Georgia Institue of Technology) argumentierte in seinem Vortrag „How NOT to build a Terminator“, dass man Roboter programmieren könne, ethisch korrekt zu handeln. Mir gab der Abend auf jeden Fall zu denken. Ich fragte mich, wie wir auch in anderen Bereichen ethisch handelnde Systeme entwickeln können.



Prof. Dr. Abraham Bernstein *studierte an der ETH Zürich Informatik und promovierte am Massachusetts Institute of Technology (MIT). 2002 folgte er dem Ruf an die Universität Zürich, wo er heute Direktor des Instituts für Informatik ist. Abraham Bernsteins Forschungsinteressen beinhalten das Semantische Web, Big Data, Data-Mining, heterogene Datenintegration, Crowd Computing und das Wechselspiel zwischen sozialen sowie technischen Elementen der Informatik.*

Kontaktadresse:

Prof. Dr. Abraham Bernstein
Institut für Informatik
Universität Zürich
Binzmühlestrasse 14
CH-8050 Zürich
Tel: +41 44 635 45 79
E-Mail: bernstein@ifi.uzh.ch
http://www.ifi.uzh.ch

In der Forschung sind wir ebenfalls aktiv im Bereich Big Data. Schon seit langem beschäftigt sich Renato Pajarola mit der Visualisierung grosser Datenmengen. In Kooperation mit Kolleginnen und Kollegen vom Geographischen Institut der UZH hat ihm der Schweizerische Nationalfonds ein Projekt zugesprochen zum Thema der interaktiven Visulisierung von Daten mit geographischem Bezug. Ich darf sogar von noch mehr Projektglück berichten: Michael Böhlen und ich haben vom neu geschaffenen Nationalen Forschungsschwerpunkt Big Data (NFP 75) eine Förderung unseres Vorhabens erhalten, das der Verarbeitung von Peta-Scale Daten unter Erhaltung der Privatsphäre dient. Hierbei untersuchen wir neue Methoden der hochvolumigen Datenverarbeitung, welche gewisse Garantien zum Datenschutz sicherstellen können. Und des Weiteren konnte ich zusammen mit Kollegen der Vetsuisse Fakultät ein weiteres NFP 75 Projekt zur Lebensmittelsicherheit lancieren. Daneben konnte das IfI auch schon verschiedentliche Industrieprojekte akquirieren: Das IfI ist also äusserst erfolgreich in dieser Thematik.

Wo wir schon bei den Industriekontakten sind: Das letzte Halbjahr war auch im Wissenstransfer äusserst erfolgreich. So hat Facebook das Projekt Zurich-Eye übernommen, an welchem Davide Scaramuzza massgeblich beteiligt und das aus dem Wyss Translational Center hervorgegangen war. Bei der neuen Einheit von Oculus geht es um Forschung und Entwicklung im Bereich des Maschinellen Sehens. Fast zur gleichen Zeit gewann das IfI-Spinnoff modium.io den ersten Platz des „Kickstart Accelator 2016“ im Bereich Future und Emerging Technologies. modium.io verwendet Blockchain-Technologien, um die Sicherheit von Medikamenten sicherzustellen.

Die letzten Monate waren sehr erfolgreich für das IfI. Wie immer war das Herbstsemester sehr intensiv, und ich konnte leider nicht selber an die Alumni GV kommen, was mir äusserst leid tat. Ich hoffe jedoch, dass das kommende Semester wieder mehr Gelegenheit für gegenseitigen Austausch bieten wird.

In diesem Sinne wünsche ich Euch allen einen guten Start in den Frühling.

Euer Abraham Bernstein

Unsere Sponsoren und Partner

Das vielfältige Angebot wäre ohne die tatkräftige und langjährige Unterstützung unserer diversen Sponsoren und Partner nicht möglich. Generell ermöglichen sie uns, sei es für Inserate im readme oder who-is-who, Events oder Emailversand, unsere Alumni-Aktivitäten in der gewohnten Art und Weise durchzuführen. Weitere Infos sind auf unserer Website unter der Rubrik „Sponsoren & Partner“ zu finden.

AdNovum

ausbildung-weiterbildung.ch
Das Schweizer Bildungsportal

avaloq
essential for banking

Goldwyn Partners Group

NSM Engineering AG

PROACADEMICSAG

Möchten Sie ebenfalls Sponsor oder Partner werden?
Nehmen Sie mit uns Kontakt auf via E-Mail an info@alumni.ch.

kalender

Dienstag, 24. Mai 2017
High-level Anlass mit OEC ALUMNI mit PartnerIn

Freitag, 16. Juni 2017
Informatiktage

Donnerstag, 29. Juni 2017
Alumni Golf-Turnier und Anfängerkurs mit PartnerIn

Samstag, 9. September 2017
Alumni Sommernachtsparty mit Special ShowAct und mit PartnerIn

Montag, 30. Oktober 2017
Goldwyn Partners Group Workshop: "Executive Videotraining – Bewerbung und Rekrutierung in Echtzeit üben!"

Montag, 06. November 2017
Alumni Generalversammlung (mit IBM-Programm)

Dienstag, 5. Dezember 2017
Fondue Plausch mit PartnerIn

Details zu den einzelnen Veranstaltungen werden jeweils vorgängig mitgeteilt. Änderungen sind jederzeit möglich. Siehe auch www.alumni.ch/events.


document
decodeURIComponent
undefined