

# «It's a bit of a fluke»

*From Henning Steier - Technology by Linguistic Search Solutions is supposed to detect money launderers and terrorists: In this interview, CEO Bertrand Lisbach reveals whether he is scared of Google and why he hasn't splashed out on marketing. – translated by Linguistic Search Solutions -*

The screenshot displays a search interface with three tabs: 'Quick Search', 'Standard Search', and 'Expert Search'. The 'Quick Search' tab is active. The search query is 'Hussein Abdurrahman'. The accuracy level is set to 'Broad'. The search results are displayed in a table format with columns for Name, Date of Birth, Nationality, and Score. The results are sorted by score, with the highest score being 96% for 'Hussein Abd al-Rahman Farid'.

Name	Date of Birth	Nationality	Score
▶ ● Hussein Abd al-Rahman Farid	● 1974	● Saudi Arabia	96%
▶ ● عبد الرحمن حسين	● 1973	● Saudi Arabia	93%
▶ ● Hossayn Faisal Abdel Rahman	● 10.07.1947	○ -	90%
▶ ● Hossein Mustafa Abdar Rahman	● 10.07.1971	● Saudi Arabia	87%
▶ ● Housain Abderahmane Mahmud	● 10.07.	⊗ United States	82%
▶ ● Hussain A.	○ -	⊗ Egypt	79%
▶ ● Hussein Abdurrahman	⊗ 22.12.	● Saudi Arabia	70%
▶ ● Hussein Khalid	● 10.07.1974	● Saudi Arabia	70%
▶ ● Abdoulrahmane Houssein	● 01.07.1975	⊗ Lebanon	69%

Traphoty: Example User Interface

"There are several ways of spelling a name like Suleyman Hussein," says Bertrand Lisbach. Sitting at his Notebook, the CEO of Linguistic Search Solutions clicks on the search button in a simple form. Some 20 results pop up on the screen - practically in real time. Including some whose letters differ considerably from the original entry, such as Souleymane Houssayn. "If names are transcribed from the Arabic alphabet into our Latin one, there can be some very different variations. Moreover, Arabic names are not divided into first and last names but are more

like a chain of names, which can have up to five different elements," explains Lisbach. His company Linguistic Search Solutions has developed a search technology to find people whose names are written and structured differently in various languages. Another example the 43-year-old likes to use is Yeltsin. The former Russian president's surname is written Jelzin in German, Yeltsin in English and Eltsine in French. If Yeltsin had wanted to open an account at the London branch of a big bank Linguistic Search Solutions technology could have been used to see what checklists he was on or if he already had an account with the Zurich headquarters. Regardless of the variant of Yeltsin's name in the particular sources being matched.

In the demo version that Lisbach presents, the name, date of birth, nationality, place of residence and so-called match score appear on the screen. The match score indicates the probability that the data belong to the person being searched for. Lisbach explains that any number of search criteria can be defined. Additional information can be called up by clicking on the name. The underlying algorithm weighs up certain criteria as being more important than others: If the address is not the same everywhere this then plays a less important role because somebody is more likely to have moved than to have changed their name.

Linguistic Search Solutions is based in Baar. The start-up was founded in April 2008 and the self-financed company now has ten full-time employees and 20 freelancers. Lisbach and the other owners sometimes have to sacrifice their salary, but this year they are hoping to make a profit and to be able to pay everybody adequately.



*Bertrand Lisbach*

#### **Swiss ICT Award 2010**

Switzerland's IT Oscars were awarded last night at the Swiss Museum of Transport in Luzern. The "Newcomer" prize went to Linguistic Search Solutions. Neteceera got the "Champion" prize and Namics received the "Audience" award. The "People" award went to Ascom CEO Riet Cadonau and to Amiando CCO Marc Bernegger in the "Newcomer" category. The jury, which was made up of industry representatives and journalists, gave the Special Award to Access for All, a Swiss foundation that advocates user-friendly technology for disabled people.

#### **20 Minutes Online: How did you come up with this business idea?**

**Bertrand Lisbach:** For one, I studied psychology and information science and worked intensively on computer-supported data analysis. And on top of that I lived for some time in the Middle East where I learnt Arabic and Hebrew. I

noticed that Hebrew and Arabic names were spelt differently depending on the language of newspapers. Later, during my time as a financial service provider consultant, I was given the task of improving the quality of search processes and of testing search solutions. Most of them were not able to find the common variations in the spellings of names. At the same time, there were many irrelevant matches. Because of these irrelevant matches many banks set up the system so it only did very precise searches. This meant that there were fewer matches but certain important information might not come up at all.

**I presume that you won't tell us which solutions you tested.**

You presume correctly.

**But you will talk about the prices of your solution?**

A desktop license costs 15,000 Swiss francs. Such a license is intended for asset managers who want to match a few stray customers with the checklists. A company license for many users starts off at 150,000 francs and there is a query fee on top of that. Depending on the query package this fee is only a couple of cents per search. Very big customers order a flat rate. The price for as many searches as you want from as many work stations as you want, including all the business processes, costs up to 1.5 million francs a year.

**An important field of application for your software Traphoty is the searching of lists of so-called politically exposed persons (PEPs) – politicians, military, religious dignitaries. This information is processed by companies such as World-Check and Dow Jones. Why have they not developed their own solution?**

I cannot speak for these companies of course, but maybe World-Check and Dow Jones Risk & Compliance see themselves less as technology providers than as providers of compliance data. That's their core competence. They consolidate hundreds of national and international checklists and process hundreds of thousands of PEP profiles. Banks use this information to combat money laundering and corruption or to observe sanctions. And this is where search technology comes into the picture. The information of the bank customers has to be matched reliably with the compliance data even if spellings vary. The names in the customer databases usually come from an identity document, whilst the compliance data are usually from the media. So it can easily be the case that the name of the same person is written differently. Before 9/11 there was little compliance information and relatively lax regulations. But the requirements of the regulating authorities have become stricter over the past decade, which is good for World-Check and Dow Jones. For us it's very fortunate that they are both our customers and are both supporting our work.

**And why do the companies need your solution?**

Most of the search technologies work with simple, purely mathematical algorithms. These cannot reliably detect whether a customer is on a sanction list. These non-linguistic search technologies are also inexact and often set off false alarms, which can be costly. Our approach is different from conventional

processes because we have looked into the reasons behind the name variations. And these are not of a mathematical but of a linguistic nature. This is why we work mainly with linguistics algorithms. Instead of having one simple mathematical algorithm our software has thousands of linguistic rules. This has an impact on the quality of the search of course. And it saves costs in compliance departments and also reduces the bank's risks.

### **How can such a rule be pictured in concrete terms?**

A simple rule could be for example that if a Russian name ends in –ev, like Gorbachev, the last letters could also be ov or ow, like in Gorbatschow. Or an example of a phonetic rule – in German a T before a Z is silent so there is no difference between a Schwartz with a T or without one. But the phonetic rules are different in English. For example U and EW sound the same, like in Stuart and Stewart. What sounds the same is often replaced with something else. The rules have to be formulated in such a way that each legitimate variation of a name matches with every other legitimate variation of the same name but not with other names. This is even more complicated than it sounds. There are also variations which we do not cover with rules but with thesauruses. Such as standard acronyms that are common when it comes to company names, such as AG for Aktiengesellschaft. There are also nicknames in the thesaurus. Bill and William are interchangeable in American English. Sasha and Alexander in Russian, and Paco and Francisco in Spanish. And what could the nickname Nell stand for? These are typical questions that our language experts deal with.

### **What does a typical business case look like for your company?**

A bank buys our software. Client advisers and compliance officers use it for individual searches. Screening systems use it for the automatic matching of transaction and customer data with checklists that have been supplied by World-Check, Dow Jones or Accuity. The IT department uses it to recognise duplicates and for data consolidation.

### **In theory you could post a website on the net, which uses sanctions data. That would be interesting for small companies.**

Maybe but we don't want to compete with data providers. Our aim is for compliance data providers to use our search technology on their own websites and thus improve online searches for their customers.

### **Customers can only use your search technology with their own databases or with databases they have bought. Why can't it be used all over the Internet?**

We do have some modules that can be used on the Internet, but you're right: It doesn't work with our big search solution Traphoty. That's because we have developed indexing that is optimised with personal data. This cannot simply be used on the unstructured database of the Internet. In future, we will probably expand our people search to news articles.

**What about Google Translate? The search engine already correctly translates your example Eltsine into the German Jelzin.**

I think that's more of a fluke. This is not true for other names such as Putin for example where the variation Poutine does not necessarily come up. This is not a criticism of Google's translation service, however, because it was not set up to find variations of names. Translations are usually irrelevant when it comes to proper names. The German Herr Schwarz does not appear as Mr Black in the English press. But the comparison is interesting. Linguistics and IT have already been used for decades in computer-generated text translation, the way that we are doing it in the specialist area of names. On top of this, Google offers the user the possibility of typing in entries in one language and then searches in others so that websites that would normally remain hidden can be found. The problem with language-dependent variations of proper names on which we are focussing has not yet been reliably solved in web searches, in my opinion. I could imagine working with Google and other search engines.

**Apart from search engines and financial service providers, who else might be interested in your solution?**

Our potential customers include the media, libraries, credit agencies, regulating authorities, social authorities and the police. The police often have to deal with information from different countries which might have country-specific spellings. Our software can help consolidate the data and make it findable.

**Other big providers such as Informatica, Basis Technology and IBM are already hard at work in your line of business. How will you survive?**

First, the global market is huge and not yet saturated with linguistic search technology – not even in Switzerland. There is space for several providers. Secondly, we hope to convince people with the quality of our product. Thirdly, we are trying to be attractive partners for software firms that already have a name and access to the market. That's why we have built our products in such a way that they can easily be incorporated into other applications. We can thank our partners for our installations in China and the Middle East. A welcome side effect of this cooperation-based business model is that we can do without our own marketing department. We don't have high customer acquisition costs and therefore we can concentrate on what we are really good at.

Article translated by

Linguistic Search Solutions AG

Source: <http://www.2omin.ch/digital/hardware/story/-Das-ist-eher-ein-Zufallstreffer--19798079>