Best Practices: How to Test Resume Parsing Software

1. Don't use fake resumes or fake data, and avoid using disguised resumes.

Our parser is designed to reject fake and disguised data such as "Employer 1." Avoid using a translator such as translate.google.com to translate a resume into another language for testing.

2. Don't accept vendor-supplied resumes for testing.

    *   Be aware that most vendor-supplied resumes have been handpicked to parse accurately by the vendor's own engine and poorly by other engines.

    *   Don't accept third-party resumes due to privacy, security and legal concerns. For instance, the European Union's privacy laws can subject you to liability of 4,000 Euros per CV. It's not worth the risk.

3. Test about 30-50 resumes per language or locale.

    *   Testing only a handful of resumes is not statistically valid.

    *   On the other hand, trying to test too many resumes (hundreds or thousands) is overkill and usually results in data overload and project failure.

4. Don't test resumes from just one source, industry or job type. Instead, test resumes from many sources and that are applicable to many industries and classifications.

5. Check with each vendor to ensure that you are setting up and running their software correctly.

6. Evaluate results individually, comparing to the actual resume. For instance, it's not accurate to assume: "Product A reported more phone numbers than Product B so Product A is better." That statistic has no probative value because unless you look at the actual resumes, you don't know whether Product B missed a valid candidate phone number, or whether Product A wrongly reported a National ID as a phone, etc.

7. Don't test only accuracy; also test completeness. You may not need certain types of data today, but what is more likely -— that going forward you will need the same amount or more data? We've been in business since 1996, and the reason our parser reports much more data than any other product is because over the years, our customers identified a need for it.

8. Test scalability and robustness, too. Accuracy is irrelevant if the product can't scale, and scalability and cost are irrelevant if the product is not accurate. For example:

   • Send a 5 MB text document to each vendor's parser. What happens? Does it hang for minutes? Crash?

   • Batch process 1,000,000 documents. What happens? Is there a memory leak? Does the processing time slow? Does it crash?

9. Test configurability. Assume that each resume must be parsed using a different set of configuration options and using a different skills taxonomy. How can that be accomplished using each vendor's software? Can the software be configured on the fly to use a new configuration and taxonomy for each transaction with no additional setup or resource overhead? Or does the software require a separate, persistent server application instance to be configured in advance for each different scenario?

10. Verify vendor claims. All vendors claim to be the most accurate, and that's clearly impossible, so ignore the claims and do your own testing. Some vendors hammer away at their scientific concepts, but are you buying academic technologies or real-world performance? Your customers are not shopping for algorithms; they just want the best performing product. Vendor emphasis on unverifiable concepts is probably a smokescreen for real-world shortcomings. No one has been doing resume parsing longer than Sovren's employees (who go back to the founding of the company in 1996), and there is no substitute for the experience we've gained over the decades. We've learned that no single approach to parsing (or searching, or matching) works best across all documents in all languages and all locales and all cultures. That's why our products use more than 30 different parsing paradigms/strategies, and yet require no pre-training.