

HEAR EVERYTHING YOUR STUDENTS ARE SAYING:

Why Text Analytics is Essential for Higher Education



CONTENTS

Executive summary	1..
No way to hear the message	2..
A wealth of comments	2..
Too much to read	2..
No way to normalize	2..
The solution: text analytics	3..
What is text analytics?	3..
The right tool for the job	4..
Simple methods are too limited	5..
Instructors follow an unsystematic process	5..
Administrators lack time to analyze comments	5..
Keyword frequency is too simplistic	6..
Word clouds are just pretty pictures	6..
Sentiment analysis is redundant	7..
What to look for in an ideal system	8..
Conclusions	9..
About eXplorance	10
Sources	11

© 2014 by eXplorance, Inc. All rights reserved.



EXECUTIVE SUMMARY

Students submit a wealth of comments with online course evaluations, but universities and colleges can't easily process them. This means administrators don't always hear the top-of-mind concerns of students.

To hear everything that students are saying, higher education needs a powerful, automated way to process natural language comments. This calls for a different form of computer software designed for text — not numbers — called “text analytics.”

Text analytics studies a body of text to find meaningful patterns or insights. This type of software isn't new and unproven: It's used today for everything from marketing to medical research. And the market is projected to triple in the next five years.

Course evaluations contain two types of items: quantitative and qualitative (open-ended). Text analytics can help analyze the unstructured data in open-ended comments.

Yet, simplistic forms of text analytics such as keyword frequency or word clouds are not powerful enough to hear the message in student comments. And sentiment analysis is redundant, since a course evaluation already reveals whether a student is satisfied or unsatisfied; the missing part of the message is why.

Higher education needs powerful text analytics software specifically adapted to automate the gathering, interpretation, and reporting of the messages contained in student comments.

For best results, this text analytics should be integrated with the institution's online course evaluation system.

NO WAY TO HEAR THE MESSAGE

Students submit many comments with online course evaluations, but universities and colleges can't easily make sense of them. There is a wealth of unstructured data available in comments, but administrators have no feasible way to listen to these messages.

A WEALTH OF COMMENTS

As universities move from paper to online course evaluations, they often discover that students submit more comments. This tendency has been discussed in numerous research papers, including the findings that students submit 4X to 7X as many words on online course evaluations as on paper forms. ¹

These comments provide valuable messages, in the form of qualitative feedback, otherwise known as natural language or unstructured data.

TOO MUCH TO READ

An instructor can read all the comments from a few hundred students, and get some sense of what they say. This may even be possible for a small department of only a few hundred students.

But when all the comments from thousands of students are compiled, no division-level or institution-level administrator can realistically deal with that much volume. Unfortunately, this means that an institution may not hear some major concerns of its students.

NO WAY TO NORMALIZE

A related problem is that different departments or divisions within an institution may have used different course evaluation forms over time. It may not be statistically valid to compare metrics gathered on quantitative items that are expressed differently.

And it is highly unlikely that open-ended comments can be normalized and compared across campus or across time. All the messages in that data are inaccessible.

THE SOLUTION: TEXT ANALYTICS

To really hear everything that students are saying, colleges and universities need a more powerful, automated way to process and compile comments.

Unstructured data like comments can't be processed the same way as the numeric ratings generated by other items on evaluation forms. Processing natural language requires a different type of computer software designed for text — not numbers — called “text analytics.”

Using this kind of software will help universities gain more value from online course evaluations, by hearing the messages in student comments, then being more proactive in responding to student concerns.

WHAT IS TEXT ANALYTICS?

Analytics means studying data to find meaningful patterns or insights. Traditional software works with numbers. For example, spreadsheets do math. Analytics software uses numbers, too. For example, Google Analytics captures metrics on website visitors and social media.

But text isn't numbers. It takes a different type of software to handle text, called “text analytics.”

A leading analyst firm defines text analytics as “the process of analyzing unstructured text, extracting relevant information, and transforming it into structured information that can be leveraged in various ways.”²

In other words, text analytics means studying a body of text to find meaningful patterns or insights.

This type of software isn't new and unproven. It's often used today to analyze social media and customer feedback, in content management systems, in fraud and e-discovery, and in medical research.

As more organizations discover text analytics, this field is expected to grow dramatically. One analyst projects the global text analytics market to grow from \$1.64 billion in 2014 to \$4.90 billion in 2019 — in other words, to triple over the next five years.³

THE RIGHT TOOL FOR THE JOB

Most course evaluation forms contain two types of items, designed to elicit two different types of feedback: quantitative and qualitative (open-ended) as shown in Figure 1.

Quantitative items gather ratings in numbers. For example, “On a scale of 1 to 5, rate how much the course provided opportunities for in-class discussion.” The response is a number in a structured format that is relatively straightforward to examine with traditional “analytics” software.

Qualitative or open-ended items gather comments in natural language. For example, a blank text box after an item, labeled “Your comments.” The response is freeform natural language, which can include local expressions, idioms, contractions, and typos.

These comments can be very precious. “While the numerical data from student evaluations provides useful overall information,” wrote researchers who studied three datasets worth of comments, “it is the responses to open-ended questions in surveys which often provide the richest source of data.”⁴

Tuning into these comments requires a different set of tools — called text analytics — that can compile, categorize, analyze, and report on this unstructured data.

An appropriate text analytics solution, specifically tuned to the needs of post-secondary education, can help administrators and educators hear what students are really saying in their comments.

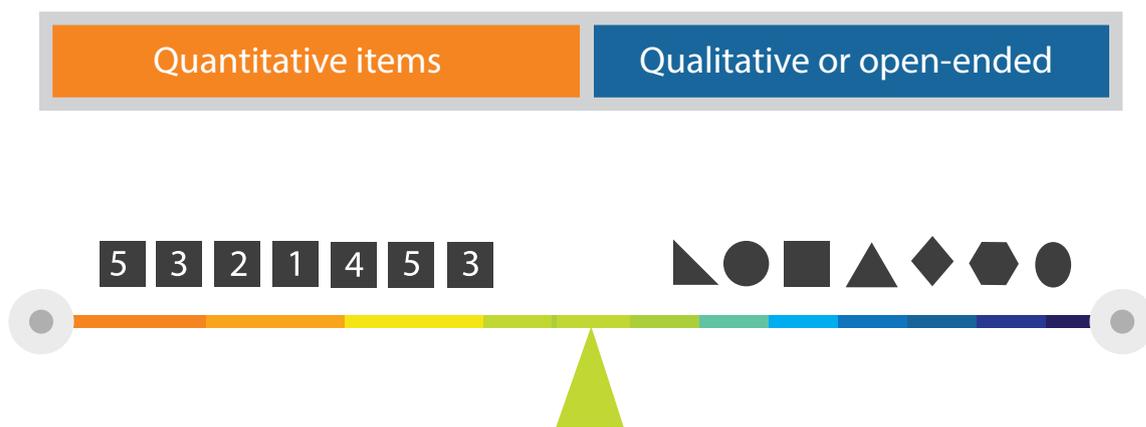


Figure 1: Various Feedback Types

SIMPLE METHODS ARE TOO LIMITED

Some simple approaches to analyzing text include:

- Keyword frequency
- Visualization or word clouds
- Sentiment analysis

But none of these is powerful enough to help educators truly hear the voice of the student from the comments.

INSTRUCTORS FOLLOW AN UNSYSTEMATIC PROCESS

Some instructors never look at student comments at all. In many cases, individual instructors do scan through at least some of their student comments to make their own interpretation.

According to a recent literature survey, many universities and colleges have no set policies on how to deal with student comments. “While many course evaluation forms include... space for general comments, institutional guidelines are not always explicit with regard to how this data is to be managed.”⁵

Such an unsystematic process makes it easy to miss the messages in a mass of student comments.

ADMINISTRATORS LACK TIME TO ANALYZE COMMENTS

In a very small department, it’s possible to scan through all the comments from students. But division- or institution-level administrators can’t begin to manually collate, review, and analyze thousands of student comments without devoting many hours to the task. As a result, they seldom do.

The same study referenced above found that “although qualitative data is collected, it is often not effectively interpreted, analyzed, or incorporated” into any decision-making process. And due to time constraints, instructors and administrators “rarely” review student comments.⁶

A powerful automated system enables administrators to compile results, produce statistics, and generate reports for every quantitative question on a course evaluation form. But the comments written by students — as freeform natural language, or unstructured data — are beyond the scope of these systems, and are not often analyzed rigorously.

KEYWORD FREQUENCY IS TOO SIMPLISTIC

Some educators are familiar with keyword frequency analysis, especially from search-engine optimization (SEO). In theory, this approach sounds good: count up how often a certain term appears, add up the frequency, and you'll see what the comments are all about.

But linguistic researchers warn that this simplistic number crunching provides no useful context for these terms, and yields little insight.

WORD CLOUDS ARE JUST PRETTY PICTURES

Since everyone appreciates visual feedback, the concept of “word clouds” was developed. Word clouds can take many forms; a simple version is commonly seen on blogs. For another example, Figure 2 shows a word cloud generated from this white paper with the online service www.wordle.net.

Notice that words like “analytics” and “Analytics” (capitalized) and “student” and “students” (singular and plural) appear as different words in this cloud. This shows the limited capabilities of this utility, which can't even equate various forms of the same word.

This word cloud is simply a graphic showing the most frequent words. Most similar visualizations lack any context, do little analysis of the text, and provide next-to-no insight. This approach is clearly not suitable for processing thousands of student comments and gathering insights from them.



Figure 2: Typical Word Cloud

SENTIMENT ANALYSIS IS REDUNDANT

Another way to analyze text is to look at the prevailing sentiment in a mass of feedback, such as a daily stream of social media. The output from sentiment analysis is most often a ratio of positive:negative terms, or a reading on the likely emotional state of the commentator.

Some companies use sentiment analysis to check whether public opinion is running with or against them. But who cares if 65% of student comments sound positive? Educators want to know what the other 35% are saying, so they can do something about their concerns.

“Using simple sentiment analysis to understand the customer experience,” said a presenter at the Text Analytics World 2014 conference, “is no better than using black-and-white photographs to understand colour.”⁷

In any case, sentiment analysis is not needed with online course evaluations, since quantitative items already capture each student’s sentiments. The comment box can help uncover the themes expressed by students about their learning experiences, in a nuanced way that goes beyond a simple positive or negative score.

WHAT TO LOOK FOR IN AN IDEAL SYSTEM

An ideal system for text analytics in higher education must provide the following features:

- **Automated collation** of all open-ended comments from all students
- **Powerful analysis** that goes far beyond word frequency or sentiment analysis
- Text analytics engine from a **leading vendor** with a successful track record
- **Global dictionary purpose-built for higher education**
- Domain-specific dictionary that understands thousands of local expressions, student vernacular, synonyms, related concepts, misspellings, and typos
- **Dictionary updated** at least once a year to remain current
- **Domain-specific hierarchy of themes** (taxonomy) developed specifically to support continuous improvement of the teaching and learning experience
- Insightful, **easy-to-read reports** with automated tables and charts
- Ability to **cross-tabulate reports** in flexible ways
- Ability to **normalize open-ended comments across campus and across time**, even when gathered with completely different course evaluation forms
- **Reports automatically compiled** for each appropriate stakeholders: individual instructors, department chairs, and division- and institution-level administrators
- Reporting can be **tailored to institutional privacy policies**
- **Automatic escalation of any “red-flag terms”** selected by institution
- Gateway provided to export student comments for even more in-depth analysis with **advanced tools**
- **Complete integration with online course evaluation** system from leading vendor
- **Integration with campus IT infrastructure**, including Student Information System and Learning Management System

Fortunately, there is a new solution designed specifically to help post-secondary administrators handle thousands of student comments quickly and easily. Blue Text Analytics provides a proven text analytics engine and a purpose-built dictionary for higher education, completely integrated with the Blue online course evaluation system from eXplorance.

CONCLUSIONS

Higher education needs powerful text analytics software specifically adapted to automate the gathering, interpretation, and reporting of the messages contained in student comments. This software must go far beyond simple keyword frequency and sentiment analysis.

For best results, these text analytics capabilities must be integrated with the institution's online course evaluation system.

With Blue Text Analytics, colleges and universities can quickly analyze thousands of comments from open-ended questions and tune in on themes that are important to students that may not have been covered in other feedback instruments.

To find out how Blue Text Analytics can help you hear everything your student are saying, download the accompanying white paper "Analyzing Student Comments in Online Course Evaluations with Blue Text Analytics." <http://goo.gl/Brg1DG>

Or call eXplorance at +1 514-938-2111.

ABOUT EXPLORANCE

eXplorance is the leading provider of Learning Experience Management (LEM) solutions for higher education, corporate training, and organization management.

Blue by eXplorance includes modules for course and instructor evaluations, psychometric and knowledge tests, 360 reviews, and broad-based stakeholder surveys. Blue powers the creation of a cycle of continuous improvement within organizations via benchmarking, feedback from all stakeholders, sophisticated reporting, customized suggestions for improvement, and automated workflows — all designed for an elevated learning experience.

Founded in 2003, eXplorance is a privately held corporation based in Montreal, Canada. Some of the company's clients include Boston College, RMIT University, UDLAP, University of Groningen, University of Louisville, UMPQUA Community College, University of Pennsylvania, University of Toronto, and many other organizations such as Discover, Fidelity, and NASA. eXplorance is recognized as one of the 2014 Best Workplaces in Canada by the Great Place to Work® Institute.

For more information, visit www.explorance.com

Toll Free: 877.938.2111(North America Only)

Tel.: +1.514.938.2111

Fax.: +1.514.635.6264

info@explorance.com



SOURCES

1: Judy Donovan, Cynthia E. Mader, and John Shinsky, "Constructive student feedback: Online vs. traditional course evaluations," *Journal of Interactive Online Learning*, Vol 5, No 3, Winter 2006, page 285

2: Fern Halper, Marcia Kaufman, and Daniel Kirsh, "Text Analytics: The Hurwitz Victory Index Report," Hurwitz and Associates, 2013, page 3

3: "Text Analytics Market Worth \$4.90 Billion by 2017," press release from MarketsandMarkets, 9 April 2014, retrieved 15 April 2015 from www.prnewswire.com/news-releases/text-analytics-market-worth-490-billion-by-2019-254531111.html

4: David Collings and Christina Ballantyne, "Online student survey comments: A qualitative improvement?" November 2004, paper presented at the 2004 Evaluation Forum, Melbourne, Australia, page 2

5: Pamela Gravestock and Emily Gregor-Greenleaf, "Student Course Evaluations: Research, Models and Trends," Higher Education Quality of Ontario, 2008, page 24

6: Gravestock and Gregor-Greenleaf, page 15

7: Moritz Sudhof, a Senior NLP engineer from Kanjoya, quoted in press release from Text Analytics World 2014, 22 January 2014, retrieved 15 April 2015 from www.prweb.com/releases/2014/01/prweb11508271.htm



© 2015 by eXplorance Inc. All rights reserved.

www.explorance.com

