Market-proven applications of media search and speech analytics

Marsal Gavalda

San Diego, CA
March 2009
Technology

- Phonetic indexing and search
- Transcript sync
- Detectors and classifiers

Markets and applications

- Media search
- Speech analytics
Corporate background

Georgia Tech spinoff, incorporated in 2000
  • Initially called “Fast-Talk”

Novel phonetic-based audio indexing and search
  • Core technology enables search of spoken content

Main markets: government, contact center, rich media, legal
  • Successful products for each market

Nexidia owns 100% of technology and applications
  • Six patents issued, many pending
Phonetic indexing and search

LVCSR
Audio → transcript
Search expression + transcript → hits

Characteristics:
- Speed: slow (~1 xRT)
- Footprint: large (~100s MB)
- Precision: ✓ high
- Recall: low
- Hits: present or absent
- Thresholds: no

Phonetic Search Engine
Audio → phonetic index
Search expression + phonetic index → hits

Characteristics:
- Speed: ✓ fast (~180 xRT)
- Footprint: ✓ small (~10 MB)
- Precision: low
- Recall: ✓ high
- Hits: ✓ confidence score
- Thresholds: ✓ yes
Measuring accuracy

LVCSR

Accuracy based on correctness of output transcript

Word error rate:

Number of substitutions + deletions + insertions, divided by number of actual words

<table>
<thead>
<tr>
<th>HYP</th>
<th>You weren’t born just to soak up sun</th>
<th>57%</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYP₂</td>
<td>You weren’t born justice so cups on</td>
<td>57%</td>
</tr>
<tr>
<td>HYP₃</td>
<td>You weren’t born justice vocal song</td>
<td>43%</td>
</tr>
<tr>
<td>HYP₄</td>
<td>You weren’t born justice so courts on</td>
<td>57%</td>
</tr>
<tr>
<td>HYP₅</td>
<td>You weren’t boring just to soak up son</td>
<td>29%</td>
</tr>
</tbody>
</table>

Phonetic Search Engine

Accuracy based on precision and recall

Precision: number of correct hits above threshold, divided by total number of hits above threshold

Recall: number of correct hits above threshold, divided by total number of correct hits independent of threshold

Thresholds allow flexible selection of operating points along the precision vs. recall tradeoff

© 2009 Nexidia, Inc. CONFIDENTIAL. DO NOT DISTRIBUTE
Sample DET curve
Structured query

Boolean, set-theoretic, and temporal operators

- AND, OR, SUBSET, ANDNOT
- BEFORE, FIRST, LAST, NOTFIRST, NOTLAST
Advantages of phonetic indexing

Robustness

• No out-of-vocabulary problem
• Robust to telephony codecs, noisy channels
• Speaker-, gender-, dialect-independent models
• Models for >30 languages

Scalability

• Indexing speed
  – 180 x RT x CPU core
  – Index >1,000 h of audio in <1 h on an 8-core server
• Search speed
  – 500,000 x RT x CPU core
  – Search >1,000 h of audio in <1 s on an 8-core server
Transcript sync

Aligns transcript to audio

- Time-stamps words in the transcript
- Assigns confidence score

Robust to

- transcript gaps
- audio gaps

Operates at word and phoneme level
Detectors for

- Silence/voice activity
- DTMF (touch tone)
- Gender

Generic classifier framework for

- Music
- Language ID

© 2009 Nexidia, Inc. CONFIDENTIAL. DO NOT DISTRIBUTE
Media Search
Media Search

Video Search
- Audio indexing and search
- Metadata search

Video Enrich
- Auto-tagging
- Auto-thresholding

Video Syndicate
- Transcript sync
- Video clipping

Transcript Sync
- Movie editing

Sample customers
- MSNBC
- Thought Equity Motion
- Microsoft
- Avid
Results based on audio and metadata

Keyframe based on best audio hit

Proximity tags

Tag-based navigation within long-form video

User clipping of video

Audio results within file

Search within file
DUBAI – Saudi investor Prince Alwaleed bin Talal thinks Citigroup Inc’s shares are undervalued, but he doesn’t appear to have much company.

The prince said Thursday he will increase his holdings in Citigroup back to 5% and expressed his support for the banking giant’s management. But after a brief bounce, the bank’s shares resumed their slide, dropping 22% to $5 in recent trading amid mounting concerns about losses and doubts about management’s ability to turn things around.

Prince Alwaleed began buying Citi shares as he strongly believes that they are dramatically undervalued,” according to an emailed statement from his office. Prince Alwaleed’s holdings in Citi are currently less than 4%, according to the statement.

Citigroup, still weighed down by troubled assets and facing mounting losses on consumer loans as economies around the world slide into recession, is struggling to restructure its balance sheet and cut costs, with 52,000 job losses planned over the next year. The prince, a key long-term shareholder in the bank, gave Chief Executive Vikram Pandit a key vote of confidence in those efforts.

Mr. Alwaleed expressed “his full and complete support to Citi management, led by Mr. Pandit, and believes they are taking all necessary steps to position the company to withstand the challenges facing the banking industry and the global economy” according to the statement.
Video Syndicate

Create phonetic index

Synchronize video with producer’s rundown or close captions

Create story boundaries

Perform entity extraction

Create and validate metadata

Suggest title

Suggest frame

Long form video

Search rundown, content-owner website and external sites to build list of key terms and named entities

1:16  1:26  1:20  0:52  1:01

President Obama, Canada

Major rain storm causes major traffic delays in Atlanta today

President Obama visits Canada today

Michael Vick talks about 2006 season

Georgia Governor Perdue meets with labor unions

Construction slows traffic on Roswell Road

1  2  3  4  5  6

© 2009 Nexidia, Inc. CONFIDENTIAL. DO NOT DISTRIBUTE
Thought Equity Motion is world’s largest archive of licensed video content

- Aggregates video content from NBC, HBO, NCAA, Sony, Paramount, MGM, National Geographic, Smithsonian Networks, and hundreds of others

Business challenge: Efficiently accept larger volumes of content, both clips and long-form content, to increase partners, customers, and content

- Prior to Nexidia >50 domain-specific FTEs manually tagging video

Nexidia **Video Search** effectively searches video archives

- Robust to varying audio quality, asset type

Nexidia **Video Enrich** effectively tags assets based on spoken content

- Time-coded tags integrated into existing metadata-based search engine
- Tens of thousands of hours of video indexed and tagged
Video Enrich for

Content aggregator
• CBSNews.com, Wall Street Journal, Associated Press, Boston Globe, etc

Video Enrich indexes and tags tens of thousands of clips
• Results presented as matrix of issues x politicians
• Enhances discoverability and syndication

© 2009 Nexidia, Inc. CONFIDENTIAL. DO NOT DISTRIBUTE
"A process that used to take hours has literally been reduced to minutes and is probably one of the greatest productivity gains of any new NLE feature to come along in years."

Oliver Peters, *Videography* magazine
Select clips from our ‘Explore’ playlist for a rich video experience

1. Easily navigate the video using the transcript

2. Find important moments using keywords

3. Clip and embed your own video edits from the transcript
Benefits across rich media ecosystem

Production
- Script sync
- Search
- Script-based editing
- Spot dialog list creation

Asset management
- Access drives demand
- Search video assets

Publication
- Leverage production notes to create clips from long-form content
- Reduce editorial burden of clip creation for on-line use

Distribution
- Distribute rich metadata and search index with media files for downstream use

Playback
- Navigate content, especially long-form with “jump to” navigation
- Deep tagging enables related playlists and alerts
- Consumer defined clips from long form for embeddable players

Monetization
- Publish more content with less expense
- SEO, content syndication
- Video search integrated with text throughout site
Speech Analytics
Speech Analytics

Science

- Speech processing, phonetic indexing vs. LVCSR, structured mining
- Robustness to noisy channels and dialectal/speaker variations
- Detection/classification of voice, music, DTMF, language, gender

Engineering

- Speed, scalability, system architecture, real-time monitoring
- Coverage of languages, creation of language packs
- Deployment options: consulting, managed services, on-site

Business

- Definition of business objectives
- Sophisticated analysis
- Recommendations
Speech Analytics in practice

Business objectives
- **Customer experience**: call transfers, escalations, repeat callers, agitated customers
- **Cost reduction**: call volume, average handle time, talk vs. non-talk time
- **Sales increase**: cross-sell/up-sell conversion rates
- **Call categorization**: cross-industry and industry-specific call drivers
- **Legal**: script adherence, agent compliance
- **Tactical analysis**: understanding call spikes, forensic search

Analysis
- **Reports**: call driver distribution, trends, correlations, agent evaluation via analysis of means, custom metrics
- **Dashboards**: activity maps, organizational views
- **Topical analysis**: agent behavior, customer behavior, products, processes, competitors
- **Drill-down capability**: call playback at relevant segments, ad-hoc searches

Recommendations
- Process improvements, agent coaching
- Success metrics, operational savings, ROI
ESI (Enterprise Speech Intelligence)

Call Driver Dashboards

Organizational Dashboards

Talk Time Analysis

Correlation

Trend

Analysis of Means

© 2009 Nexidia, Inc. CONFIDENTIAL. DO NOT DISTRIBUTE
Call driver analysis

Combines analysis of several performance metrics:

- Call volume, handle time, total handle time, and talk vs. non-talk
Correlation analysis

Shows relationships among call categories in a single report
Trend analysis

Charts handle time, total handle time, talk vs. non-talk across time

- Choice of time range and time units
Handle time analysis

Histogram of call durations and talk vs. non-talk times
Analysis of Means compares handle times of agents / supervisors / sites

- Decision boundaries (in red) delimit statistically significant variations
Activity map

Area indicates call volume, color indicates handle time

- Green: below average, red: above average
## Case studies

<table>
<thead>
<tr>
<th>Industry</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td>Eliminated the need for an outsourced call categorization service saving $1.2M annually</td>
</tr>
<tr>
<td>Healthcare</td>
<td>Identified errors in call coding processes saving $500,000 annually by reducing contact costs</td>
</tr>
<tr>
<td>Hospitality</td>
<td>Identified additional revenue opportunity of $20M through implementation of new agent sales effectiveness process</td>
</tr>
<tr>
<td>Hardware Manufacturer</td>
<td>Delivering on CEO mandated improvements in CSAT resulting in better and faster analysis of order inquiries (root of 51% of negative experiences)</td>
</tr>
<tr>
<td>Software Manufacturer</td>
<td>Transfer analysis led to reduced expense, improved CSAT, better product feature insights and improved relationship with outsourcers</td>
</tr>
<tr>
<td>Call Centre Outsourcer</td>
<td>Automating a significant part of quality management function to deliver $15M annual savings</td>
</tr>
<tr>
<td>Wireless</td>
<td>Analysis of regular callers led to process and billing changes that are delivering $1.3M savings in year one</td>
</tr>
<tr>
<td>Electric Utility</td>
<td>Broad analysis of CSAT surveys delivered significant new insights to drive product direction and customer relevance</td>
</tr>
<tr>
<td>Energy</td>
<td>Regulations and compliance demand analysis of trading conversations, robust audit capabilities, ability to forensically search large quantities of audio</td>
</tr>
<tr>
<td>Internet Service Provider</td>
<td>Automating and rationalizing QA process, fixing call routing errors, delivered &gt; $1M in cost savings through the first 7 months of managed services delivery</td>
</tr>
</tbody>
</table>
## Passed

<table>
<thead>
<tr>
<th>Name</th>
<th>User ID</th>
<th>Test Number</th>
<th>Location</th>
<th>Global</th>
<th>Pronunciation</th>
<th>Fluency</th>
<th>Speed</th>
<th>Status</th>
<th>Status Change</th>
<th>Application Date</th>
<th>Reviewed Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yarhese, Sam</td>
<td>96100500</td>
<td>1</td>
<td>Sutherland Cochin</td>
<td>66</td>
<td>60</td>
<td>79</td>
<td>-3%</td>
<td>Passed</td>
<td>Manual</td>
<td>4/23/2008</td>
<td>4/23/2008</td>
</tr>
<tr>
<td>S. Yisah</td>
<td>59020662</td>
<td>1</td>
<td>Sutherland Cochin</td>
<td>66</td>
<td>61</td>
<td>76</td>
<td>+8%</td>
<td>Passed</td>
<td>Manual</td>
<td>3/20/2008</td>
<td>3/20/2008</td>
</tr>
<tr>
<td>Raj, Efrisa</td>
<td>36858248</td>
<td>1</td>
<td>SITEL Hyderabad</td>
<td>66</td>
<td>59</td>
<td>79</td>
<td>-19%</td>
<td>Passed</td>
<td>Manual</td>
<td>12/19/2007</td>
<td>1/30/2008</td>
</tr>
<tr>
<td>Harika, Dulpola</td>
<td>53567119</td>
<td>1</td>
<td>SITEL Hyderabad</td>
<td>66</td>
<td>61</td>
<td>77</td>
<td>+3%</td>
<td>Completed</td>
<td>Manual</td>
<td>12/19/2007</td>
<td>1/30/2008</td>
</tr>
<tr>
<td>Plonanda, Kenogen</td>
<td>23616151</td>
<td>1</td>
<td>Sutherland Philippines</td>
<td>66</td>
<td>58</td>
<td>80</td>
<td>-5%</td>
<td>Passed</td>
<td>Manual</td>
<td>4/2/2008</td>
<td>4/2/2008</td>
</tr>
<tr>
<td>Yeluz, Kristine</td>
<td>10662316</td>
<td>1</td>
<td>Sutherland Philippines</td>
<td>66</td>
<td>60</td>
<td>78</td>
<td>+1%</td>
<td>Passed</td>
<td>Manual</td>
<td>5/6/2008</td>
<td></td>
</tr>
<tr>
<td>Sahu, Madhu S.</td>
<td>85925304</td>
<td>1</td>
<td>Sutherland Cochin</td>
<td>66</td>
<td>61</td>
<td>78</td>
<td>-5%</td>
<td>Passed</td>
<td>Manual</td>
<td>4/20/2008</td>
<td>4/20/2008</td>
</tr>
<tr>
<td>Yankavala, Kresh</td>
<td>88770523</td>
<td>1</td>
<td>Knoss Hyderabad</td>
<td>66</td>
<td>60</td>
<td>85</td>
<td>-31%</td>
<td>Passed</td>
<td>Manual</td>
<td>1/28/2008</td>
<td>1/28/2008</td>
</tr>
<tr>
<td>Hilario Garcia, K.</td>
<td>80750961</td>
<td>1</td>
<td>Delta Vincent Santa ...</td>
<td>66</td>
<td>61</td>
<td>76</td>
<td>-3%</td>
<td>Passed</td>
<td>Manual</td>
<td>1/15/2008</td>
<td>1/17/2008</td>
</tr>
</tbody>
</table>

### Application Details
- Offset: 00:01 How may I help you today?
- Offset: 00:02 Yes, we do offer phone service.
- Offset: 00:05 The home phone bundle
- Offset: 00:06 combines both D S L and
- Offset: 00:08 voice home phone service.
- Offset: 00:10 EarthLink true voice
- Offset: 00:11 includes premium calling
- Offset: 00:13 features like voicemail,

### Phoneme Performance

#### Consonants:
- Stop: p b
- Affricate: ch jh
- Nasal: m
- Fricative: f v th d h s z sh zh ng
- Approximant: w r l
- Lateral Approximant

#### Vowels
- Close: iy ih
- Close-mid: ay
g- Open-mid: ch ah er ow oy
- Open: ae ah ao aw ay

© 2009 Nexidia, Inc. CONFIDENTIAL. DO NOT DISTRIBUTE
Concluding remarks

Solid technology
• Scalable, accurate audio indexing
• >30 languages

Market-proven solutions
• Enterprise Speech Intelligence and Language Assessor for **contact centers**
• Video Search, Video Enrich, Video Syndicate and Transcript Sync for **rich media**
• Audio Finder for **legal e-discovery**
• SDK for select customers