## Who are we?

<table>
<thead>
<tr>
<th>Name/Title</th>
<th>Experience</th>
<th>Vocabulary</th>
<th>Image Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrew LaBonte</td>
<td>10 years</td>
<td>390,000 terms</td>
<td>8 million, 2 websites</td>
</tr>
<tr>
<td>Senior Manager, Search Metadata Corbis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mary Forster</td>
<td>17 years</td>
<td>150,000 terms</td>
<td>60 million, 7 websites</td>
</tr>
<tr>
<td>Director of Search Strategy Getty Images</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What are we talking about?

• Search behavior
• Aspects of an image that are keyworded
• Benefits of controlled vocabulary
• Translation
• Who keywords and impact of culture
• Keywording investment

How do people search for images?

Some know exactly what they want.

“I need a close-up of a 30-something Asian business woman talking on a cell phone on a busy urban street”

Some don’t.

“I need to convey the importance of family in the face of adversity”

“What was Lady Gaga doing last night?”
Two basic search patterns

Many to few
• Enter a search
• Look at the results
• New search or refine
• Repeat

Driven by:
• Keyword search
• Filters

One to many
• Find an image of interest
• Want “more like this one”

Driven by:
• Keywords
• “Sets” within same event or shoot
• Visual search

Browse can be an entry point to this process.

What do people search for? (in creative stock)

Most often
• Man-made objects (passport, disco ball, skateboard)

Often
• People (baby, woman, family, red hair, businessman)
• Natural world (bird nest, cows, dandelion, beach)
• Concepts (alone, future, growth, hope, love, precision)

Sometimes
• Locations (Florence, Hawaii, China, Eiffel Tower)
• Activities (ballet, diving, gardening, rock climbing)
• Descriptive (broken, silver, round, pattern)
What do people search for? (in news, sports, entertainment)

**Most often**
- Names of people (Lady Gaga, Shakira, Tiger Woods)

**Sometimes**
- Sports terms (Nascar, NHL, speed skating, rowing)
- Team names (AC Milan, Minnesota Vikings)
- Locations (Athens, Iran, Gatwick, Colombia)
- Man-made objects (Chanel, space shuttle, catwalk)

**Seldom**
- Concepts (funny, security, stress, weird)

What makes people give up?

**Too many irrelevant results**
Visually weak results are not worth wading through

**Root cause: over keywording (spamming)**
Describing too much detail, including tangentially-related keywords
What makes people give up?

Too little results
Zero, or only a few

Root cause:
Keywords too general
(need pea, in addition to vegetable)
Lack of synonym control
Typos/misspellings

LITERAL KEYWORDS:
- two people
- man
- woman
- husband
- wife
- Caucasian
- 60-65
- boom box
- living room
- armchair
- sunglasses
- reading
- book

CONCEPTUAL KEYWORDS:
- active seniors
- contrast
- young at heart
- gender roles
- retirement
- rebellion
- retro
- cool
- uncool
- marriage
What’s in a picture?

COMMERCIAL
- one person
- adult man
- 30-35
- Caucasian
- running
- ramp
- ascending
- briefcase
- challenge
- opportunity
- difficulty
- economic recovery

EDITORIAL
Who: famous CEO Fred Metaphor
What: International Businessman Olympics
Where: Taal Monument, South Africa
When: June 8, 2010

Controlled vocabulary - what’s the big deal?

<table>
<thead>
<tr>
<th>CONTRIBUTOR KEYWORDS</th>
<th>CONTROLLED VOCABULARY</th>
<th>SEARCHABLE TERMS</th>
<th>NINE OTHER SEARCHABLE LANGUAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excite</td>
<td>Excited</td>
<td>Exciting</td>
<td>Opwinding</td>
</tr>
<tr>
<td>Excited</td>
<td>Exhilarated</td>
<td>Empolgação</td>
<td></td>
</tr>
<tr>
<td>Exciting</td>
<td>Exhilarating</td>
<td>Empolgação</td>
<td></td>
</tr>
<tr>
<td>Adrenaline</td>
<td>Excitation</td>
<td>Hochgefühl</td>
<td></td>
</tr>
<tr>
<td>Excite</td>
<td>Excited</td>
<td>Excitazione</td>
<td></td>
</tr>
<tr>
<td>Excited</td>
<td>Exhilarated</td>
<td>Podniecienie</td>
<td></td>
</tr>
<tr>
<td>Exciting</td>
<td>Exhilarating</td>
<td>Empolgação</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Emoción</td>
<td></td>
</tr>
</tbody>
</table>
Multi-lingual search: all based on translations

Translation scalability

Controlled Vocabularies store translations in a scalable, efficient manner. Once a translation is stored, it can be used again and again to support multi-lingual search.

Translations of free-text, such as Captions, are expensive and difficult to re-use, because free-text has so much variation.
Who can keyword?

- In-house keywording teams
- Outsourcers
- Contributors
- Crowd sourcing

Issues and impacts of who keywords

- Cultural awareness
- Language ability
- Oversight / Control
- Cost
Cultural awareness

Does it matter whether the person keywording is aware of the cultural phenomena being portrayed in an image?

Yes!

It’s easy to miss the main point of an image if you are not aware of the cultural phenomena it is portraying.

Cultural awareness – how do we bridge the gaps?

Communication
Contributors and editors should tell the keyworders what the point is. This information should be part of the image workflow.

Training
People can and must learn about other cultures. Outsourcers have shown they can do it!
Language skills

Strong language skills are necessary to add the right keyword. It’s easy to make mistakes in a language that has not been mastered.
• Skull vs Scull

Oversight: consistency

If you want consistent keywording, training is required. People do not keyword images the same way unless they are trained to do so.

Consistency of keywording gives search results:
• Predictability
• Completeness
• Confidence
Oversight: accuracy

Keywording oversight gets more accurate keywording.

Keyworders need facts about the image in order to be accurate.

People unaware of the need for factual information are less likely to restrict keywords to facts.

Cost

Outsourcing is typically less expensive than in-house.

Contributor keywords are “free”.

Crowd sourcing is either free or extremely inexpensive (ex: 5c for 25 images)
Cost versus oversight of keywording

Findability!

Cost
Contributors/Crowdsourcing
Outsourcers
In-house

Oversight

Keywording and value spectrum

CONTENT
Microstock
Royalty free
Rights managed
Celebrity
Portraiture
Historical

VALUE
Low value
Inexpensive to produce
High Volume
Unique
Art-Directed
Harder to reproduce

KEYWORDING
Contributors
Crowdsourcing
Trained keyworders
Subject experts
Bargain bin vs. department store

- Bargain Bin store
  - Organized
  - Merchandized
  - Easy shopping

Why should you care?

- Keyword what customers search
- Cover the right aspects of images, don’t over-keyword
- Controlled Vocabulary variant & multi-language search
- Awareness & language ability matter to keywording
- Match keywording investment to business model and image value
- Good search data = more search options
Questions?