

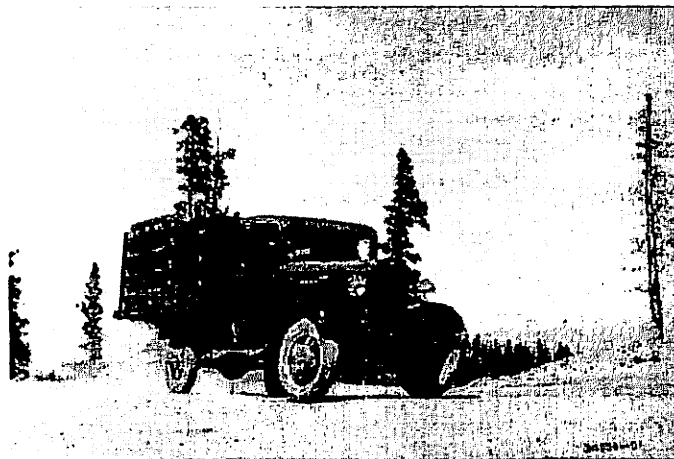
# Layered indexing of images

Kimberly A Schroeder

The General Motors Media Archives (GMMA) is undertaking one of the largest digitization efforts in the world. GMMA has developed a layered approach to visual indexing that dissects the objects, style and implication of each image, so that the indexing system can accommodate all potential approaches to the material.

The General Motors Media Archives (GMMA) project is one of the largest digitization efforts in one of the busiest Archives in the world. GMMA houses over 3,000,000 still photographic images and tens of thousands of motion picture films and videos. The collection spans over a hundred years, and material from every corner of the earth and beyond (GM developed the Lunar Rover).

The images are a rich history of the evolution of transportation, urban growth, fashion, design, and popular culture. Imagery presents many challenges in order to become retrievable. When quantity, cultural differences and historical context are added, it becomes even more complex.



**figure 1: 1935 Chevrolet Truck**

In order to achieve indexing consistency and precise research results, the GMMA has developed a layered approach to indexing imagery. The following are the three basic layers into which each image is dissected.

## Object layer

This first layer defines the bare components of an image. The first impression of Figure 1 is a truck with trees. In order to “read” the image, an indexer would have to know what is important to the audience. The most obvious object is the “1935 Chevrolet Truck”. This would also be the most important object for GM. The trees are secondary and are not distinctive to this image, for an automaker. Yet, if the same image was at the National Agriculture Library, the trees would be more

important than the vehicle. This same image might be indexed as “evergreens (native to American Southwest)”, and “truck”.

The following is a breakdown of this image, object by object:

### objects

- 1) 1935 Chevrolet Truck
- 2) evergreen trees
- 3) clouds
- 4) dirt roads
- 5) skylines/ landscapes

The above five things are all objects that one can see immediately when looking at the picture. Next, value must be added as an indexer, instead of an observer. Of the above five objects, which ones have value for retrieval?

### objects

**ranking (1 to 5,  
5 = least important)**

- |                         |   |
|-------------------------|---|
| 1) 1935 Chevrolet Truck | 1 |
| 2) evergreen trees      | 4 |
| 3) clouds               | 5 |
| 4) dirt roads           | 3 |
| 5) skylines/landscapes  | 2 |

Ranking these objects involves training in indexing, but also in research. How do clients ask for images? What are the most common requests? What do clients not ask for?

After looking at the explanation of ranking, an indexer can see that only the top three ranked objects are important for retrieval.

objects	ranking (1 to 5; 5 = least important)	explanation of ranking
1) 1935 Chevrolet Truck	1	This is the client’s core product
2) Evergreen trees	4	Trees are in background, not prominent
3) Clouds	5	Appear in most outside shots, not important, unless unusual (e.g. threatening)
4) Dirt roads	3	A way to differentiate from clean studio shots
5) Skylines/ landscapes	2	Images with a local flavor or skyline have a distinction over typical interior shots

Indexing for this image would not include the evergreen trees or the clouds.

Of course this is an oversimplification of GM's product descriptions. General Motors' product is extremely complex. The complete vehicle description would also include "one and a half ton Stake Truck with a conventional cab." GMMA has developed a hierarchical structure on automotive terminology that previously had no structure.

This is just part of defining the objects. The above list of five items are only the nouns of the picture. With visual and audio media, there is also action. The "verb" of the image is missing. Imagery is not passive. What are the five objects doing within the context of the image?

In evaluating the object ranked as the number one noun, the truck, an indexer should observe the following:

The truck is working hard.

It is kicking up dust, so it may be running fast.

The bed of the truck seems to be full.

An indexer needs to observe such cues in images. Referring back to original notes handwritten on the image envelope, the indexer sees that this shot was taken at Pike's Peak in Colorado. An educated GM indexer needs to know that GM conducted endurance races up that mountain for years. There is also an additional note that the truck was "fully loaded" which reemphasizes the endurance aspect. "Endurance race" is just another object in the image. It is the action that belongs in the object layer.

The history of this image was well documented, but this is unusual and every indexer needs to be aware of subtle implications in imagery. This awareness may allow the only opportunity to identify the image properly and reuse it. The value of collections increases as identity and retrievability increase. For instance, a shot of a group of men taken fifty years ago is of little value if no one knows who they are. The value and reuse of the images can increase fifty-fold if the group of men is found to be the first GM Board of Directors meeting.

### Style layer

The second layer is the purpose for which the image was captured. These are some general types of images in the GM Media Archives:

- 1) candid — documentary or special event
- 2) glamour — sales, studio or advertising
- 3) engineering testing — analysis and experimentation

These categories are also the style of photography. Candid is earthy and non-posed. Glamour is high-style, with spotless cars and pleasing composition. Engineering testing is gritty and detail oriented.

Keep in mind that researchers are dependent on the indexers' work. A video producer will not want a gritty engineering shot for a series on "The glamour of General Motors." A GM engineer will not want an image of Dinah Shore on stage at an auto show pointing to an automotive part, even if it is the right auto part. The styles of photography must be differentiated. This is more necessary with a collection as large as GM's. If a collection has only 20 pictures of each vehicle, it is not so necessary to differentiate as it is when there are thousands per vehicle.

This was accomplished throughout GMMA's indexing by building a hierarchical thesaurus with synonym capability. This gave researchers, clients and indexers the flexibility of both generality and specificity. Under the style categories, there are subtopics that add depth to this layer.

### Implication layer

This third layer is most elusive. It takes the objects, the style and together they become a story. This is where an indexers' level of competence is evident. Going back to the first image of the 1935 Chevrolet Truck, the objects have been defined, the style is candid and now the indexer asks:

What is the point of this image?

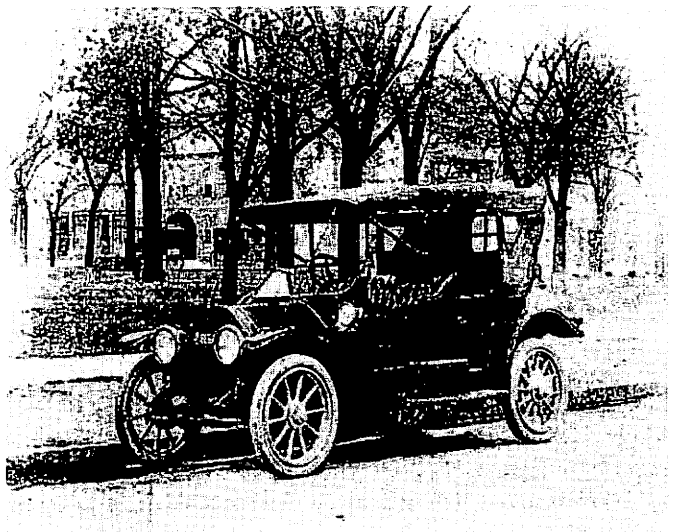
Why would someone want to look at it?

What is unique or informative about it?

The deep intent of the image was to show possible customers that Chevrolet Trucks were so durable they could climb up a mountainside fully loaded. Certainly, if GM products could do that, they could fulfill any other daily needs a consumer asked of them. This was a way to prove beyond a shadow of a doubt that Chevy Trucks were the best in class.

This is the most value-added layer that the researchers work from. It is attempting to apply tangibility to subjective interpretation, to "gut-feeling" and emotion generated from seeing an image. This is where our greatest intellectual challenge as indexers lies. It is what makes an image "art" instead of merely a "snapshot".

The following images are examples of Implication layers that can easily be missed.

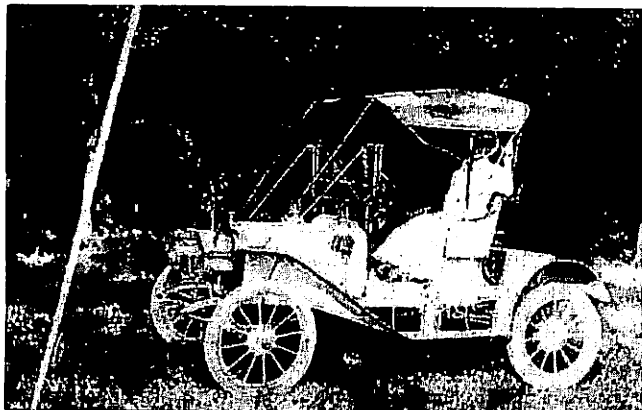


**Image A: 1912 Cadillac**

**object layer** 1912 Cadillac  
street scenes

**style layer** glamour

**implication description** first vehicle to have a self-starter (no crank). Changed the viability of vehicles as transportation. Allowed women to use automobiles independently, since great physical effort was no longer needed.



**Image B: 1909 Buick**

**object layer** 1909 Buick Model 10  
woman  
park setting

**style layer** glamour

**implication description**  
first Buick was in 1903, but this was one of the first Buicks produced by GM



**Image D: Knudsen Day Parade**

**object layer** parades; banners; men marching  
military uniforms; street scenes  
musical instruments

**style layer** candid

**implication description**  
This is a Welcome Home Parade for William Knudsen. He left his position as President of GM to run the United States conversion from commercial to military production in World War II. His annual salary was only \$1. Detroit set aside an entire day to welcome him home.

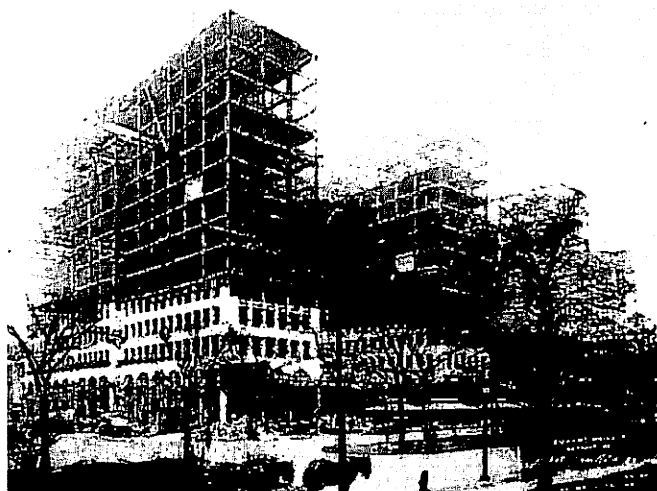


**Image C: 1910 McLaughlin Buick**

**object layer** 1910 McLaughlin Buick (Canadian)  
horses  
street scenes  
advertising art  
couples (people)  
weddings

**style layer** glamour

**implication description**  
GM's earliest export business. Earliest agreement for GM to be involved in supplying components to other companies.



**Image E: Construction of General Motors Corporate Headquarters**

**object layer** GM Building; construction; corporate headquarters; historic landmarks; horses

**style layer** candid

**implication description**  
This was the largest construction of a Corporate Headquarters in 1919. More specific to this picture, horses were hauling raw materials to the biggest automaker in the world's new building. Trucks did not yet have the capacity to replace horses.

All of the above have a deeper layer of meaning. These are milestones that exponentially increase the image value to GM.

- The 1912 Cadillac (Image A) is more important than the 1911.
- The 1909 Buick (Image B) is more important than the 1910.
- The 1910 McLaughlin (Image C) has an important distinction from earlier models.
- The Knudsen Day Parade (Image D) is not just a parade, but a day prominent in American history.
- The construction of the GM Building (Image E) shows the expansion of the auto market in 1919, but also how far it still had to go to usher in the end of the horse-drawn era.

This inherent significance may be something an indexer

immediately recognizes, it may be something that a product expert recognizes or it may be something that is recognized years after it has been indexed. The indexing process must accommodate all possibilities in order to serve clients and history itself.

---

*Kimberly A Schroeder's company Archive Impact is a consulting firm specializing in indexing, training and workflow issues on imaging projects.*

---

#### Acknowledgements

*Special thanks to Pam Gannon, Mies Martin and Margie Jankowski Friant; thanks also to the team at the GM Media Archives. All photographs Copyright 1978 GM Corporation*

---

---

# SOCIETY OF INDEXERS Annual Conference 1998

Grand Hotel, Grand Parade, Tynemouth, Tyne & Wear NE30 4ER  
9–11 October 1998

The Tynemouth conference is primarily a practical indexing conference. We are delighted that Nancy Mulvany, author of *Indexing books*, has agreed to speak.

Sessions will include talks and seminars on electronic indexing: latest developments including

- CD-ROMs
- the world-wide web and transnational indexing
- year 2000 compliance
- business and health issues
- indexer/ editor relationships

plus several short sessions where topics of burning interest to all indexers will be discussed, plus

#### Indexing workshops:

- beginners workshop— conducted by Ann Hudson
- advanced workshop — conducted by Pat F Booth

The conference will begin at tea time on Friday and end after lunch on Sunday. Workshops will take place on the Sunday afternoon and Monday morning after the conference. All the conference activities will take place in the Grand Hotel.

Although there will be limited free time during the conference itself, we may (weather permitting) arrange a trip to the castles on the coast north of Newcastle or to Hadrian's Wall on the Sunday afternoon if there is sufficient interest.

#### Accommodation and conference facilities:

The Grand Hotel, a Victorian spa hotel, was once the summer residence of the Duchess of Northumberland. It is situated next to Long Sands, five minutes' walk from the spectacular ruins of Tynemouth Priory and Castle and the headland at the mouth of the River Tyne. Accommodation is being booked in the Grand and its sister hotel, the Windsor, in Whitley Bay, just under two miles along the sea front. Guests at the Windsor are welcome to

stay in the bars at the Grand in the evening for as long as they like. There are taxi services to take guests back to the Windsor in the small hours. The Windsor is six minutes on foot from Whitley Bay metro station and the Grand 11 minutes from Tynemouth metro. Accommodation at both hotels is excellent, with all rooms ensuite.

We offer a considerably reduced rate for shared accommodation in twin-bedded rooms. Please note that we cannot guarantee rooms with a sea view. Lifts and disabled access are available at the Grand Hotel. Budget accommodation is available in guest houses in Tynemouth and Whitley Bay. If you are interested in this please let us know. Special rates (£20-£30 B&B) have been negotiated for those wishing to stay in the conference hotel over Sunday night. Please contact the hotel directly nearer the time, mentioning the Society of Indexers Conference.

#### Getting to Tynemouth:

Tynemouth is approximately 8 miles to the east of central Newcastle and is connected to Newcastle Central Station and Newcastle Airport by the Tyne & Wear metro. Train journey times from London, Manchester and Birmingham to Newcastle are (surprisingly) very similar and range between 2 hours 40 minutes to 3 hours 15 minutes (slightly longer on Sundays). The train journey from Edinburgh to Newcastle is about 1.5 hours. Newcastle airport has several flights a day to and from Heathrow, Gatwick and Stansted as well as all other major British cities. International connections include direct flights to Dublin, Amsterdam, Brussels, Paris and Bergen. For those coming by car there is parking immediately outside the hotels and in the surrounding streets (no restrictions).

For more information or booking forms, contact:  
Audrey Judkins, Treasurer, SI Conference 1998, 134 Elwick Road, Hartlepool, Cleveland, TS26 9LX, UK, Tel: 01429 274731