Research Project: Design, Delivery, and Dissemination (Week 2)

Welcome!!!

David Griesemer, J.D., M.A.
IFSH Library Manager and Webmaster
(708) 563-8160
dgriesem@iit.edu
Instructional Goal

The goal of this session is to provide you with a set of skills and knowledge that will enable you to locate, evaluate, and use information effectively, ethically, and legally.

To Meet the Instructional Goal We Will -

1. Review How Scientific Information is Produced, Organized, and Disseminated.

2. Learn how to Develop an Effective Research Strategy.

3. Utilize Endnote to Efficiently Manage Your Citations.

4. Learn how to Recognize and Avoid Plagiarism.
Part 1: Information Production, Organization, and Dissemination

Steps in the Food Science Information Cycle

1. Conduct an experiment, and record data in lab notebook.

2. Write up your findings in the form of an article or book.

3. Copyright then attaches to your written work.
Part 1: Information Production, Organization, & Dissemination

Definition of Copyright

A copyright attaches to any original work or authorship fixed in any tangible medium of creation now known or later developed.

Elements of Copyright

1. Originality

Work must owe its existence to the author and no one else.

2. Fixation

Work must be written down on paper, a computer file, or on audio or visual tape.

Facts, Ideas, and Titles are Not Copyrightable

Why are facts, ideas, and titles not copyrightable?
Part 1: Information Production, Organization, and Dissemination

Non-copyrightable Subject Matter

1. Facts, ideas, titles, and phrases are not copyrightable because they are not considered original.

2. The facts and data in your laboratory are not copyrightable because others could have discovered those facts as well.

3. Only how you express those facts in your research paper are copyrightable because that expression is original.
Part 1: Information Production, Organization, and Dissemination

- **Exclusive Rights of the Copyright Holder**
  1. Sole right to reproduce the work (copy the work).
  2. Sole right to disseminate work (give the work to others).
  3. Sole right to display the work (e.g. Internet & Television).

- **Licensing or Transferring Exclusive Rights**
  1. The copyright holder may license all of these exclusive rights or some of these exclusive rights.
  2. The copyright holder may transfer (e.g. sell) the entire copyright. So the buyer now has all the exclusive rights.
Duration of the Copyright

1. Single Author = 70 years after death of the author.

2. Multiple Authors = 70 years after the death of the last surviving author.

3. Work for Hire = work prepared by an employee for the employer
   
   i. If you are working for a food company they own the copyright.

   ii. The work for hire rule does not apply to university professors so the professor owns the copyright not the university.
Public Domain

70 years after the author’s death copyright ends and anybody can use the work without the copyright holder’s permission.

Rationale of Copyright Law

Moves science and society forward by giving authors a financial incentive to create new works that benefit humanity.

So everybody wins:

a. The copyright holder wins because they have 70 years to commercially exploit the work.

b. Society wins because at the end of 70 years the work goes into the public domain and copyright ends.
Fair Use Exception to Copyright

The Fair Use exception states that individuals may make use of certain portions of copyrighted material for teaching, scholarship, and research without the copyright holder’s permission.

Businesses (For Profit) Uses of Copyrighted Work

The Fair Use Exception does not allow for business uses of copyrighted work even if that material is for research.

For example, if you are working for a Food Company in the R&D Department the Fair Use Exception does not apply. You must have the permission of the copyright holder to use the work.
What the Fair Use Exception Allows

1. Copying a chapter from a book for educational purposes only.
2. Copying an entire journal article for educational purposes only.
3. Copying a sentence or paragraph from any work.

What the Fair Use Exception does not allows

2. Copying the same journal article, or from the same journal more than 5 times per year.
3. Copying that causes the author major economic harm (e.g. takes a lot of money out of the copyright holder’s pocket).
Part 1: Information Production, Organization, and Dissemination

- **Federal Government and Copyright Law**
  The Federal Government (FDA) cannot have a copyright in any work that its employees produce.

- **Application of this Rule to the FDA**
  This means that if all the authors on a paper are FDA authors, no copyright exists in the work, and you can use the work for any purpose – education or business.
Part 1: Information Production, Organization, and Dissemination

Steps in the Food Science Information Cycle

1. Conduct an experiment, and record data in lab notebook.

2. Write up your findings in the form of an article or book.

3. Copyright attaches to your written work.

4. Scientists then tries to get the work published:
   a. If a journal article, then by a scholarly journal.
   b. If a book or book chapter, then by reputable publisher.
Part 1: Information Production, Organization, and Dissemination

Why Scholarly Journals are Important

1. Scientists must publish in scholarly journals for promotion.
2. Students must use scholarly journal articles for research.

Defining Scholarly Journals (Peer Reviewed Journals)

Scholarly Journals are publications that are issued on a periodic basis (e.g. monthly, quarterly, or yearly) and are peer reviewed meaning that a group of scholars decides whether or not to accept or reject an article into the publication.
Example of a Scholarly Journal

- Journal has no artwork on the cover
- Journals often have the word “Journal” on the cover
- Published monthly, organized by volume and number
- Published by an Association

GARY L. ANDERSON, KRISHAUN N. CALDWELL, LARRY R. BEUCHAT, AND PHILLIP L. WILLIAMS

1Department of Environmental Health Science, University of Georgia, Athens, Georgia 30602-2302; and 2Center for Food Safety and Department of Food Science and Technology, University of Georgia, 1109 Experiment Street, Griffin, Georgia 30223-1797, USA

MS 02-4551: Received 13 December 2002/Accepted 28 March 2003

ABSTRACT

Free-living nematodes may harbor, protect, and disperse bacteria, including those ingested and passed in viable form in feces. These nematodes are potential vectors for human pathogens and may play a role in foodborne diseases associated with fruits and vegetables eaten raw. In this study, we evaluated the associations between a free-living soil nematode, *Caenorhabditis elegans*, and *Escherichia coli*, an avirulent strain of *Salmonella Typhimurium*, *Listeria monocytogenes*, and *Bacillus cereus*. On an agar medium, young adult worms quickly moved toward colonies of all four bacteria; over 90% of 3-day-old adult worms entered colonies within 16 min after inoculation. After 48 h, worms moved in and out of colonies of *L. monocytogenes* and *B. cereus* but remained associated with *E. coli* and *Salmonella Typhimurium* colonies for at least 96 h. Young adult worms fed on cells of the four bacteria suspended in K medium. Worms survived and reproduced with the use of nutrients derived from all test bacteria, as determined for eggs laid by second-generation worms after culturing for 96 h. Development was slightly slower for worms fed gram-positive bacteria than for worms fed gram-negative bacteria. Worms that fed for 24 h on bacterial lawns formed on tryptic soy agar dispersed bacteria over a 3-h period when they were transferred to a bacteria-free agar surface. The results of this study suggest that *C. elegans* and perhaps other free-living nematodes are potential vectors for both gram-positive and gram-negative bacteria, including foodborne pathogens in soil.

The agricultural impacts of plant and animal parasitic nematodes have long been recognized and, by virtue of their direct effects on fruit and vegetable production, have been extensively studied. Comparatively little is known regarding the impact of free-living microbivorous nematodes on produce production and safety, although these nematodes are the most abundant and widespread soil mesofauna. They play an important role in microbial degradation to plants. In addition, rhizobial nematodes feed a diet of *Escherichia coli* defecated viable cells (2). Nematodes are also able to protect bacteria present in their digestive tracts from chlorination (28). *Salmonella* ingested by *P. hispanica* is reported to have survived the chlorination of wastewater at a treatment facility (35). Thus, nematodes may protect, and disperse bacteria, including those ingested and passed in viable form in feces.
Example of a Scholarly Journal

Following the dispersal period for each test bacterium, it is shown in Figure 4. Worms dispersed 69 ± 13 Salmonella Typhimurium foci per worm over the 3-day dispersal period, a level similar to the determined for E. coli (56 ± 11 foci per worm). The dispersal level for L. weihenmayeri (93 ± 15 foci per worm) exceeded those for other bacteria; while the dispersal level for B. cereus (36 ± 10 foci per worm) was lower than those for the other test bacteria. In terms of differences in feeding and development of C. elegans on different species of bacteria, the nematode is capable of dispersing all four species of bacteria.

Exposure of worms that had fed on lawns of bacteria to 0.4 mM sodium hypochlorite for 3 to 6 min followed by two washes in K medium reduced the dispersal of cells on TSA. Dispersal rates were 36, 28, 39, and 9 foci per worm for E. coli, Salmonella Typhimurium, L. weihenmayeri, and B. cereus, respectively. Reduced phoretic transport following hypochlorite treatment may account for the reduced levels of dispersal. In addition, bacteria may have been excised through defecation during the 5- to 6-min hypochlorite treatment and during ca. 10 min of washing following the treatment. The suprornic nematode P. brasiliensis degrades viable plant pathogenic bacteria over a 27-h period following feeding (12). For C. elegans, defecation is limited in the first hour after feeding (17); however, excretion prior to deposition of worms on TSA cannot be precluded as a factor in the reduced level of dispersal following hypochlorite treatment. The persistence of dispersal following hypochlorite treatment is consistent with, although it does not unequivocally prove, the excision of viable bacterial cells in feces. The dispersal of viable E. coli by C. elegans after hypochlorite treatment was not unexpected, since Adams and Goll (2) have reported the defecation of viable E. coli by plant nematodes. Viable Salmonella are also reported to be harmful to humans. Rhabditid nematodes are omnivorous feeders, thus having the greatest potential to disperse pathogens in soils. Observations reported here demonstrate that C. elegans, a flashlight form, is strongly attracted to and consumes L. weihenmayeri and B. cereus, both soil inhabitants, as well as E. coli and Salmonella Typhi murium, which can be introduced into soils through the deposition of animal feces, untreated irrigation water, or runoff water from livestock feeding lots. C. elegans was actively attracted to all of the test bacteria but over time appeared to limit its exposure to E. weihenmayeri and B. cereus. Behavioral avoidance, even after an initial attraction, could decrease the potential for the vectoring of bacteria by rhabditid nematodes.

We also observed that C. elegans consumed less B. cereus, particularly if populations of the bacteria were large, than other test bacteria, including an avirulent strain of Salmonella Typhimurium. Reduced feeding, like the delayed avoidance of bacteria, may also reduce the potential for vectoring. C. elegans dispersed B. cereus, but in lower numbers (fewer foci per worm per unit of time) than it dispersed the other test bacteria. Thus, a reduced level of feeding appears to limit but not preclude the dispersal of B. cereus under laboratory conditions. In contrast to B. cereus, the level of feeding for Salmonella Typhimurium was equal to or above that observed for E. coli. This observation, along with the fact that some Salmonella strains may infect the gut of C. elegans (1), suggests the need for studies to determine the role of C. elegans as a vector for the transmission of Salmonella in soil to the surfaces of fruits and vegetables. We have explored the association between C. elegans and Salmonella Poona (5). Worms that had fed on Salmonella Poona were able to transport the pathogen in soil to cantaloupe rind placed on the soil surface. Further studies, particularly those involving soil media amended with manure, are needed to explore factors affecting this process. In addition, foodborne pathogenic microorganisms that are less favored but nonetheless consumed by bacterial or nematodes should also be investigated. The consumption and dispersal of these microorganisms could be amplified by factors imposed by the soil environment.

ACKNOWLEDGMENT
The nematode strain used in this work was provided by the Center for Nematology, Cornell University, which is funded by the NIH Center for Research Resources.

REFERENCES
Trade Magazines

Trade magazines report on industry trends and practices and are written for industry professionals not academics.

Using Trade Magazines in Your Research

Students may use a few trade magazine articles, but the vast majority of research should come from scholarly journals.
Examples of Trade Magazines

Trade magazines contain colorful artwork or pictures on the front cover.
Examples of Trade Magazines

Weighing Your Options
Accuracy and sanitation remain key for weighers, but many systems now perform multiple functions.

Weighing systems forever have played key roles in manufacturing. While their essential job has not changed, expectations about their performance—particularly sanitation and accuracy—keep getting tougher.

Inconsistency in weighing can be costly—either losing product by overfilling containers or risking fines by not meeting weighing and measurement standards. A large American pudding and yogurt manufacturer was having this exact problem (consistently) in fill weights during packaging.

The company’s fill weights were inconsistent and filler operators were not performing the SPC process correctly. The process needed a method by which it could monitor fill weights and attributes on a real-time basis, allowing the opportunity to make adjustments to the process immediately.

Mettler Toledo (www.mt.com), Columbus, Ohio, at first installed a single-station consisting of an ID50 industrial PC weighing terminal and a KA15 base—a stainless steel, high-precision scale base capable of handling harsh wash-down environments. FreeWeigh.Net software, a statistical process control software also from Mettler Toledo, was installed as the customer could track the filling application.

FreeWeigh.Net records and tracks net weight and other quality data from samples taken off the line. It calculates production averages and statistical measures, and it automatically generates reports and control charts. By watching these charts, operators can make real-time adjustments on the production line.

After six months, the system paid for itself by reducing overfill costs, according to Mettler Toledo. Based on the success of the first phase, the customer completed Phase II of the project by installing additional systems. The final system included four ID50 workstations with KA15 bases and FreeWeigh.Net software with a client monitoring bundle, which allowed the manager to watch the filling process from a remote PC.

Customers are demanding better weigher performance, in terms of both speed and accuracy, says Brian Burt, sales manager for packaging systems in the Lititz, Pa., office of Hearn and Control Inc. (www.hearnandcontrol.com), which is based in Hayward, Calif. Heen and Control is the exclusive sales and service representative for Ishida weighers, inspection and packaging systems in North and South America.

But Burt adds: “There’s been increased customer focus on equipment sanitation. Ishida has made great efforts to provide our customers the most sanitary offering in the industry. Our R series construction features the latest materials and manufacturing practices to ensure the Ishida minimizes and eliminates opportunities for product buildup and contamination while also meeting the most stringent of washdown conditions.”

Some other customer demands, according to Burt:

- Flexibility: A good operator interface makes it easy to meet the changing demands of production requirements. With an operating life of 20-plus years for most weighing systems, flexibility at the time of purchase is an important consideration.

- Special weighers for every need: As with most lines of equipment, weigh systems have become specialized for particular products. Whether weighing packaged products flying down a line at high speeds or dairy products or raw or sticky products, weighing systems can be developed for most any product.

Thermo Fisher Scientific’s Process Instrumentation Division (www.thermofisher.com), Waltham, Mass., offers a broad line of weigh systems. High-speed check weighers offer high accuracy at speeds up to 600 cases or just per minute. “This makes it easier for our customers to justify a capital expense—faster production throughout and real savings in material giveaway,” says Rick Cash, marketing technology manager.

Cost-conscious plants have been able to refurbish the frames of prior-generation units and install only the controller units of new-generation check weighers for a “significant discount over the cost of a new system,” he adds.

And in the case of all three of these vendors and others, food safety technologies—such as metal detection and X-ray inspection—can be integrated with the weighing system to perform other quality checks.
Examples of Popular Magazines

- Popular magazines contain short articles on a broad range of subjects. Students may not use popular magazine articles for their scientific research.
Recap of the Food Science Information Cycle

1. Conduct an experiment, and record data in lab notebook.

2. Write up your findings in the form of an article or book.

3. Copyright attaches to your written work.

4. Scientists then tries to get the work published -
   a. If an article, then in a scholarly journal.
   b. If a book or chapter then, by a reputable publisher.

5. *When the work is accepted for publication, the scientist then transfers, or sells, the copyright to the publisher.*
Copyright Transfer Process

1. The author executes a sales contract that transfers the entire copyright to the publisher.

2. Once the transfer is complete, the publisher now has all the exclusive rights of the copyright holder:
   a. Reproduce
   b. Disseminate
   c. Display

3. If the original author wishes to copy, or display the work, they must get the permission of the new copyright holder.
Part 1: Information Production, Organization, and Dissemination

Displaying the Article in a Database

1. Large publishers like Wiley have their own databases so they have the right to display their articles on their database.

2. Database Vendors, who are not publishers, license the rights to display articles on their own database.

3. *Licensing* means that you pay the copyright holder money to copy, display, or disseminate the work.

4. Libraries then license databases from publishers and vendors giving students access to copyrighted material.
Part 1: Information Production, Organization, and Dissemination

Legally Using the Library’s Database

1. The License Agreement that a library signs states that you may only use the database for educational purposes.

2. The *Fair Use Exception* to the Copyright Act also states that anybody may make use of *portions* of copyrighted education.

3. The Copyright Act states that when you are using a database it is illegal for you to give your password to others.
Subject Classification

Before an article or book is displayed in a database it must be classified meaning that one or more words or phrases is assigned to the work describing what the work is about.

1. Subject Cataloging

The act of assigning these words or phrases is called subject cataloging.

2. Library of Congress Classification System (LCSH)

Subject catalogers use very specific words or phrases selected from the Library of Congress Subject Headings (LCSH) to describe what a book or journal article is about.
Example of Journal Article Classification

1. Subject cataloger looked at the title and abstract of this article to see what it was about.

2. Subject cataloger then selected very specific words and phrases from the LCSH list to describe the book.

3. Subject cataloger followed very specific guidelines in choosing the words and phrases.

4. Therefore, you have authority control meaning that can use the subject headings to identify other related material.

Antimicrobial activity against foodborne pathogens of chitosan biofilms of different molecular weights.

Authors: Kim, Kyung W.¹
Min, B.J.²
Kim, Young-Teck¹ ykim@clemson.edu
Kimmel, Robert M.¹
Cooksey, Kay¹
Park, S.I.³

Source: LWT - Food Science & Technology; Mar2011, Vol. 44 Issue 2, p565-569, 5p

Document Type: Article

Subject Terms: *ANTI-Infective agents
*FOOD pathogens
*CHITOSAN
*BIOPOLYMERS
*MOLECULAR weights
*LISTERIA monocytogenes
*ESCHERICHIA coli
*VISCOSITY

Abstract: Abstract: Antimicrobial activity against Listeria monocytogenes, Escherichia coli 0111
Salmonella typhimurium of chitosan biopolymer films (CBFs) prepared with four viscosities of chitosans (10, 40, 100 and 200 mPa-s) were investigated by agar diff
## Structure of a Database

Databases are composed of many records, and each record has many fields (e.g. Title, Author, Abstract, and Subject).

<table>
<thead>
<tr>
<th>Record</th>
<th>Title</th>
<th>Author</th>
<th>Abstract</th>
<th>Subjects (LCSH)</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>Effect of Deletion of Genes Involved in Lipopolysacch Core ....</td>
<td>Monadi, A. R.; Mirzaei, H.; Javadi, A.; Hosseinzade, N.;</td>
<td>In this study, the effects of Lactobacillus Bifidobacterium angulatom and</td>
<td>Immunology, Infectious Diseases</td>
</tr>
<tr>
<td>#2</td>
<td>A foodborne outbreak of enterotoxigenic E. coli</td>
<td>Wall, Daniel M.; Srikanth, C. V.; McCormick, Beth</td>
<td>When one considers the organism Salmonella ....</td>
<td>Pharmacology, Molecular Genetics, Tumor Biology</td>
</tr>
<tr>
<td>#3</td>
<td>Use of Organic Acids to Inactivate Escherichia coli ...</td>
<td>Park, Sang-Hyun; Choi, Mi-Ran; Park, Jeong-Woong;</td>
<td>This study was undertaken to investigate the antimicrobial ...</td>
<td>Biochemistry studies - Lipids, Food technology</td>
</tr>
</tbody>
</table>
Structure of a Database

- **Database Default Field**

  Each database has a default field which is really not a field, but is a search parameter that allows you to search all the fields at the same time.

- **Examples of Default Fields**

  The Galvin Library calls its Default Field “Keyword”
Examples of Default Fields

ScienceDirect calls its Default Field “All Fields”
Examples of Database Default Fields

Web of Science calls its Default Field “Topic”

PubMed calls its Default Field “All Fields”

Academic Search Premier calls its Default Field “Select a Field (optional)”

Note: What is important to remember is that whenever you search a database’s default field you are searching many fields at the same time, and that the database’s default field is not the same as the “Subject” field.
The Internet is a network of millions of computers linked by wire and wireless connections. The Worldwide Web, comprised of webpages, is only part of the Internet.
Internet and Worldwide Web Basis

- **How the Worldwide Web Works**

  The Worldwide web is a collection of billions of Hypertext Markup Language (HTML) pages that are linked together and transmit data through Hypertext Transfer Protocol (HTTP).

- **Example of IFSH Library Website HTML Code**

  ![HTML Code Example](http://library.iit.edu/ifs/ - Original Source)

  ```html
  <body>
  <script type="text/javascript" src="../js/wz_tooltip.js"></script>
  <div id="container">
    <div id="header">
      <a href="http://library.iit.edu"><img id="iit_logo" src="../images/iit_black_logo.gif" width="349" border="0" alt="Illinois Institute of Technology" /></a>
    </div>
    <div id="quicklinks">
      <form name="iitsearch" action="http://www.iit.edu/about/search_results.shtml" id="cse-search-box">
        <select name="quicklinks" onchange="urlJump(this);">
          <option value="">Quick Links</option>
          <option value="http://library.iit.edu/">
            Administrative Offices</option>
          <option value="http://www.iit.edu/graduate_admission/">
            Admission (Graduate)</option>
          <option value="http://www.iit.edu/undergrad_admission/">
            Admission (Undergrad)</option>
          <option value="http://athletics.iit.edu/">
            Athletics</option>
          <option value="http://iitbkstore.com/bkstore/content">
            Bookstore</option>
          <option value="http://www.iit.edu/graduate_college/bulletin/">
            Bulletin (Graduate)</option>
          <option value="http://www.iit.edu/ugaa/">
            Bulletin (Undergrad)</option>
          <option value="http://library.iit.edu/bursar/">
            Bursar</option>
        </select>
      </form>
    </div>
  </div>
  </body>
  ```

  **HTTP Links**
How a Web browser Works

A web browser like Internet Explorer or Firefox translates the HTML code into texts, graphics, and hyperlinks.
**Determining Webpage Reliability**

Since anybody can create a webpage on any topic, you must first determine that the webpage is a reliable source before you use it.

**Webpage Reliability Factors**

1. **Website Domain (Web Address)**
   a. *Government websites* (.gov) are usually reliable sources.
   b. *Educational websites* (.edu) are usually reliable sources.
   c. *Non-profit websites* (.org) may or may not be reliable sources. Other factors listed below need to be examined.
   d. *Commercial websites* (.com) need to be heavily scrutinized using the factors below. They exist to sell goods and services.
Webpages Reliability Factors

2. Authoritative Source

To determine if a webpage is authoritative, ask –

a. Is the author of the website a well known expert or scientists? If yes, then the website may be reliable.

b. If the website has no author, is the website published by a .gov or .edu? If yes, it may be reliable.
Webpage Reliability Factors

3. Timeliness

A website that has been recently updated, is more reliable than an older website.

4. References or a Bibliography

A website that contains references or a bibliography is always more reliable than a website with no citations.

5. Advertisements

A website that has pop-up advertisements or promotes a business product or service is probably not a reliable website. It needs to be heavily scrutinized.
Example of a Reliable Webpage

This webpage is from the University of Wisconsin so it is an educational website (.edu)

The author of this webpage is a professor

The webpage was recently updated

The website has references for the data on the table

Table 1. The process must deliver a heat treatment of sufficient lethality so as to ensure a safe, shelf-stable product. Minimum temperature/time combinations are given.

<table>
<thead>
<tr>
<th>Temp (°F)</th>
<th>Time (min)</th>
<th>Temp (°F)</th>
<th>Time (min)</th>
<th>Temp (°F)</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>140</td>
<td>12.7</td>
<td>152</td>
<td>3.1</td>
<td>162</td>
<td>0.3</td>
</tr>
<tr>
<td>142</td>
<td>10.1</td>
<td>154</td>
<td>2.4</td>
<td>164</td>
<td>0.2</td>
</tr>
<tr>
<td>144</td>
<td>7.9</td>
<td>156</td>
<td>1.9</td>
<td>166</td>
<td>0.2</td>
</tr>
<tr>
<td>146</td>
<td>6.3</td>
<td>158</td>
<td>1.6</td>
<td>168</td>
<td>0.1</td>
</tr>
<tr>
<td>148</td>
<td>4.9</td>
<td>160</td>
<td>1.2</td>
<td>170</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>3.9</td>
<td></td>
<td></td>
<td>180</td>
<td>0.1</td>
</tr>
</tbody>
</table>


Dr. Barbara Ingham
University of Wisconsin-Madison
btingham@wisc.edu
31 October 2011.
What Search Engines Do

Search engines like Google essentially *catalog* the Worldwide Web by trying to figure out what each webpage is about. The websites were ranked based on the search terms “food safety.”
Internet and Worldwide Web Basis

How a Search Engine Catalogs the Web

1. Search engines employ software called “spiders” or “robots” that crawl through various websites and webpages.

2. The spider then looks at words on and phrases found on different parts of the HTML document (web page).
   a. The spider looks at words in the website’s URL.
b. The spider looks at words in the website’s title.

c. The spider looks at words in bold and large type.

The text in hyperlinks are cataloged by the search engine.

The spider can’t index images or pictures that contains words only text.
d. The spider looks at the webpage’s keyword and descriptive metadata that is not displayed by the browser.
e. The spider looks at how many other webpages link to the subject webpage, and then looks at the content of all those other pages to help catalog the target webpage.
Internet and Worldwide Web Basis

**All elements crunched through an algorithm**

After Google looks at all of these elements it then uses a mathematical algorithm to assign different weights to different words and phrases and essentially catalogs the page – determines what it is about.

**Overview of Search Engine Process**

- **Spider Crawl**
  - Website
  - Index terms for each page are generated
  - Cataloged Web Pages Uploaded to Google’s Database

- **User Google’s a topic**
  - Google’s database generates websites

- **Food Safety Issue**
  - Learn about food safety | FoodS
  - www.foodsafety.gov
Internet and Worldwide Web Basis

- So why do we care how search engines work and how webpages are created?
Two Reasons Why We Care –

1. Webpage Reliability

   Not all web pages indexed by search engines are reliable. You must examine each web page using the criteria mentioned above.

2. Webpage Indexing

   a. Search Engines only index about 10-15% of the Web. For example, the 40 million articles in Web of Science, are not indexed.

   b. Google uses a mathematical algorithm and not a human cataloger to determine what each web page is about so the search results from Google are not as accurate as a database.
What does all this mean?

To find scholarly information that is directly related to your search terms, always search the academic databases before you search Google or Google Scholar.
Important Note on Copyright Law

All the Images Used in this Lecture are Under Copyright and I have not gotten permission from the copyright holder to reproduce, display, or disseminate these images.

Am I in violation of copyright law?
Important Note on Copyright Law

- **Fair Use Exception to Copyright**

  The Fair Use exception states that individuals may make use of *certain portions* of copyrighted material for teaching, scholarship, and research without the copyright holder’s permission.

- **Application to Library Instructional Module**

  Since these images are being used for teaching (educational) purposes, they may be used without getting the permission of the copyright holder. However, if these images where being used for commercial or profit purposes (e.g. business presentation), the copyright holder must give permission to use these images.