Name (print): ________________________________

- **Instructions:**
  - Keep your eyes on your own paper, and do your best to prevent anyone else from seeing your work.
  - Do NOT communicate with anyone other than the professor/proctor for ANY reason in ANY language in ANY manner.
  - This is exam is closed book, closed notes, no calculator, and no computer.
  - Turn all mobile devices off and put them away now. You cannot have them on your desk.
  - Write neatly and clearly. What I cannot read, I will assume to be incorrect.
  - Stop writing immediately when told to do so at the end of the exam. I will take 5 points off your exam if I have to tell you multiple times to do so.
  - Academic misconduct will not be tolerated and will be referred immediately to the Emory Honor Council. Penalties for misconduct will be a zero on this exam, an F grade in the course, and/or other disciplinary action that may be applied by the Emory Honor Council.

- **Time:** This exam has 13 questions on 11 pages including the title page. Please check to make sure all pages are included. You will have 75 minutes to complete this exam.

---

*I commit to uphold the ideals of honor and integrity by refusing to betray the trust bestowed upon me as a member of the Emory community. I have also read and understand the requirements and policies outlined above.*

Signature: ________________________________

---

<table>
<thead>
<tr>
<th>Page</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Points:</td>
<td>12</td>
<td>12</td>
<td>10</td>
<td>9</td>
<td>7</td>
<td>12</td>
<td>13</td>
<td>75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
</table>
1. (12 points) Describe how the Internet and World Wide Web work using all of the terms below. To receive full credit, your answer must demonstrate the relationships between these terms in the context of the functionality of the Internet and WWW rather than simply defining them.

<table>
<thead>
<tr>
<th>HTTP (HyperText Transfer Protocol)</th>
<th>packets (or packet switching)</th>
<th>web server</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL (Uniform Resource Locator)</td>
<td>HTML (HyperText Markup Language)</td>
<td>web browser</td>
</tr>
<tr>
<td>DNS (Domain Name System)</td>
<td>IP (Internet Protocol) addresses</td>
<td>web page</td>
</tr>
<tr>
<td>routers</td>
<td>Internet Service Provider (ISP)</td>
<td>client</td>
</tr>
</tbody>
</table>

**Solution:** A user generates a request for a webpage by typing a URL like http://www.mycompany.com into their web browser. This compromises the “client” portion of the client-server interaction upon which the web/internet are built. The “http” part of the URL indicates the protocols/rules which govern the format of transferred data. HTTP indicates that the information being transferred is a webpage which uses hypertext and written in HTML which is the language of webpages. Since the domain name in the URL is in a human readable format like “mycompany.com”, it must be translated to a unique IP address. This translation from human-readable domain name to numerical IP address is accomplished by utilizing DNS (Domain Name Server) computers. This unique IP address allows the request to be routed to the web server via specialized computers called routers. The routers control which of the many possible paths between client and server (and vice versa) information takes. The web server for mycompany receives the request and sends the data back to the requesting computer. The data transferred between the client and server is broken up into packets or smaller chunks of data for more efficient transfer and routing. Once the packets containing the HTML for the webpage are received by the client’s web browser, the browser renders the HTML into webpages containing text, images, and other media. Both the clients and servers are connected to the internet infrastructure via ISPs such as Time-Warner, Verizon, or Comcast who maintain the hardware, cables, and other infrastructure required to transmit the data.
2. Pick one of the following business scenarios and answer the following questions concerning
data that will be generated by your business. Indicate which scenario you choose.

- You are opening an online store from which to sell your Great Aunt Josie’s home-
made baked goods.
- You run a small business which specializes in data deletion and disposal of old
  computer hardware (such as hard drives) for businesses.
- You are a firm which is hired to manage new customer acquisition by a large multi-
national law firm.

(a) (3 points) List 3 types of relevant data that your business will need to collect.

Solution: Answers vary. For #3:
client info (name, state, etc)
potential legal issues
group/lawyer in firm most qualified to represent client

For #1:
consumer data: name, address, ID, etc.
transaction data: amount of sale, number of products purchased.
website statistics: how long did customers visit site, what page did they search for the most

(b) (3 points) For each of the pieces of data you listed above, specify what the source of that data might be.

Solution: Ch. 11, Sec. 5. For #3:
first two pieces of info could come from in-take survey (paper or electronic) or
in-person interview.
3rd piece of data could be obtained by cross referencing customer’s current legal issue with table/database/file of lawyers/groups and specialties from partner firm.

For #1:
consumer data: source would be customer setting up online account or making
a purchase w/ a credit card and shipping info.
transaction data: source would be from point of sale (online site, most likely)
traffic data: collected/supplied by webserver as users browse website.

(c) (6 points) List 3 reasonable policies with regard to data governance for your busi-
ness. Explain why you consider each policy to be “reasonable” in the context of your business.
Solution: Ch. 11, Sec. 5. For #3:

policy 1: data stored electronically must be encrypted. Since we’re dealing with legal data, we must take extra safeguards with it.

policy 2: offsite backup of all data nightly: can’t risk losing customer data if computers are compromised or a natural disaster happens.

policy 3: data retention of only 6 months after case is resolved: keep data for less time and destroy/erase electronic data sooner than usual since it is sensitive legal data.
3. (2 points) Label each part of the following URL with the appropriate terminology as discussed in class and in your textbook.

http://mathcs.emory.edu/~valerie/courses/fall11/155/syllabus.html

Solution: protocol, host name, domain name, top-level domain, path, file

4. (2 points) Explain why a company would need to maintain a data warehouse.

Solution: separate transactional data (needed to run business on day-to-day basis) from older data needed for analysis. centralize data from many points for analysis.

5. (3 points) Explain the relationship between DBMSs, tables, records, and fields.

Solution: dbmses are made up of databases. Each database is a collection of tables which hold information about a class of entities. records store information about 1 specific entity in a table. fields define the specific pieces of data which is collected about each record in the table.

6. (3 points) Explain how a bot-net could be used in a Distributed Denial of Service (DDoS) attack.

Solution: A bot-net is a network of computers which are zombies. This means they can be directed to do various tasks w/o the owner’s consent. In a DDoS attack, these computers are directed to make requests (legitimate or not) to a targeted server. This server may be overwhelmed by the requests and not be able to serve any information to any client (bot-net computer or otherwise).
7. (3 points) The Internet is a *distributed or decentralized network*. Explain what this term means. List at least two advantages of a distributed network over a centralized network structure.

**Solution:** Distributed/decentralized networks do not have one central point of either information storage or communication. Thus the network is more fault tolerant since one computer may go down and there are still routes existing between the destination and source. Additionally, these networks are more efficient due to the dynamic routing of information around congested points in the network. They also lack a central point which may be a target for hackers/terrorists/etc.

8. (3 points) Are cookies (either first- or third-party) a security risk, a privacy risk, or neither. Explain your answer.

**Solution:** Technically they may be both, although they are generally considered more of a privacy risk than a security risk. They do store log-in credentials (although in an encrypted/disguised manner), but if someone gained access to your computer through another security breach, info stored via cookies could be vulnerable. Privacy is usually more of a concern with cookies, particularly 3rd party cookies. Movements from website to website can be tracked via 3rd party cookies leading to a very complete picture of you as a web-surfer.

9. (3 points) Explain the difference between data, information, and knowledge in the context of database systems.

**Solution:** Ch 11, Sec. 2: “Data refers simply to raw facts and figures. Alone it tells you nothing. The real goal is to turn data into information. Data becomes information when it’s presented in a context so that it can answer a question or support decision making. And it’s when this information can be combined with a manager’s knowledge - their insight from experience and expertise - that stronger decisions can be made.”
10. (4 points) List two ways digital technologies have changed data collection practices. Give an example of each.

**Solution:** Examples will vary but needed to be specific.
- **AMOUNT** of data that can be collected, **DURATION** data exists/can be stored for, **SPEED** with which data can be collected, **KIND/TYPE** of data which is available/accessible.

Biggest error seen: listing 2 specific digital technologies that (may) have changed data collection in a specific area, but not talking about the ways in which general technologies have changed all data collection practices.

11. (3 points) Explain the difference between a virus, a worm, and a trojan.

**Solution:** Ch. 13, Sec. 3
- virus: replicates itself through another program; needs a host
- worm: self-contained; take advantage of security vulnerabilities to spread, but don’t need another program through which to spread
- trojans: malware which disguises itself as another (usually harmless or beneficial) program
12. (12 points) In class, we spent time covering information security and discussed both technological vulnerabilities of information systems as well as human vulnerabilities. 

*Note: your answers below must be distinctly different. That is, they cannot be small variations of the same vulnerability.*

(a) List 4 sources of technological vulnerabilities.

**Solution:** See Figure 13.1 in textbook. Lots of technological vulnerabilities listed, but needed to list examples of technology rather than human actions/behavior.

Biggest problem w/ answers: listing the same thing in different ways. For example listing “keyloggers, viruses, worms, and trojans” as your 4 vulnerabilities. These are all “malware” and examples of the same vulnerability. As specified in the instructions, examples needed to be different.

(b) List 4 sources of human vulnerabilities.

**Solution:** See Figure 13.1 in textbook. Lots of human vulnerabilities listed.

(c) For 2 technological and 2 human vulnerabilities you listed above, give an a policy an organization could put in place to try to remove these vulnerabilities.

**Solution:** Answers vary, but could be things like:
Set systems to require password changes every 30 days and ensure strong passwords.
Hire impartial security firms to verify credentials to prevent unauthorized access to secure areas.
Prevent installation of unauthorized software on employee computers.
Utilize black/white lists to prevent employees from visiting problematic websites.
Set software to “auto-update” and install patches as soon as they become available.
13. Consider the database illustrated on the next page which keeps track of student data. In this picture, a relationship of “Many” has been illustrated with an “M”. For parts b and c below, you may answer by clearly drawing on the illustration of the database on the next page.

(a) (2 points) Why does the system need a “Course Title” table?

**Solution:** Need to separate title information (rather than just including it as a field in the “Course List” table) to reduce redundancy. If the course title changes, we can simply change it in one place (“Course Title” table) rather than having to change multiple records in the “Course List” table.

Biggest problem seen: Describing the function of the “Course Title” table (e.g. “it functions as a foreign key and links the table to...”) but not describing why it exists/is useful to our information/data design.

(b) (5 points) You want to keep track of faculty data. Add a “Faculty” table to this database including appropriate data. Indicate a relationship between this table and the “Course List” table via primary and foreign keys.

**Solution:** See figure 11.1 in textbook.
+1 for table
+2 for (some of the) appropriate fields in table
+2 for relationship to “Course List” table.

Biggest problem seen: trying to link via CourseNum. This introduces a huge amount of redundancy when faculty teach more than one course.

(c) (6 points) Imagine you also want to keep track of student majors. Add a table (or tables) to accommodate this, assuming a student can have more than one major. Link it to other appropriate tables via primary and foreign keys.

**Solution:** Add two tables: “Major Enrollment” and “Major List”. The “Major Enrollment” table will contain the student’s ID number and a code for the major (e.g. something like “BIO” or a numerical code like “001”). Thus, a student can have multiple records in the table if they have multiple majors. The “Major List” table would contain the code for the major and a textual description of that major.

For the “Major Enrollment” table, the student ID would be the foreign key and linked to the “Students” table via that field. The major id would also be a foreign key in this table. For the “Major List” table, the major id/code would be the primary key and linked to the “Major Enrollment” table.
+2 tables
+2 fields in tables
+2 links

Biggest problem seen: Not allowing a flexible number of majors.
### Students

<table>
<thead>
<tr>
<th>STUDENT_ID</th>
<th>FIRST_NAME</th>
<th>LAST_NAME</th>
<th>CAMPUS_ADDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>U11-00001</td>
<td>Phil</td>
<td>D’Basket</td>
<td>Weaver Hall 411</td>
</tr>
<tr>
<td>U11-00002</td>
<td>Art</td>
<td>Major</td>
<td>101 Creative Commons</td>
</tr>
<tr>
<td>U11-00003</td>
<td>Anne</td>
<td>Teak</td>
<td>Stickley 014</td>
</tr>
<tr>
<td>U11-00004</td>
<td>Anita</td>
<td>Knappe</td>
<td>Highway View 303</td>
</tr>
<tr>
<td>U11-00005</td>
<td>Easton</td>
<td>West</td>
<td>N. South Hall 123</td>
</tr>
<tr>
<td>U11-00006</td>
<td>Jack</td>
<td>Hammer</td>
<td>Tule 543</td>
</tr>
</tbody>
</table>

### Enrollment

<table>
<thead>
<tr>
<th>STUDENT_ID</th>
<th>COURSE_INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>U11-00002</td>
<td>9301</td>
</tr>
<tr>
<td>U11-00002</td>
<td>9303</td>
</tr>
<tr>
<td>U11-00002</td>
<td>9306</td>
</tr>
<tr>
<td>U11-00005</td>
<td>9305</td>
</tr>
<tr>
<td>U11-00003</td>
<td>9301</td>
</tr>
</tbody>
</table>

### Course Title

<table>
<thead>
<tr>
<th>COURSE_NUM</th>
<th>COURSE_TITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI101</td>
<td>IS for Managers</td>
</tr>
<tr>
<td>MI201</td>
<td>Prog. for Mgmt.</td>
</tr>
<tr>
<td>EN101</td>
<td>Intro. to Lit.</td>
</tr>
<tr>
<td>CJ101</td>
<td>Fighting Cyber-Crime</td>
</tr>
</tbody>
</table>

### Course List

<table>
<thead>
<tr>
<th>COURSE_INDEX</th>
<th>COURSE_NUM</th>
<th>TIME</th>
<th>LOCATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>9301</td>
<td>MI101</td>
<td>MWF 9</td>
<td>Fulton 250</td>
</tr>
<tr>
<td>9302</td>
<td>MI101</td>
<td>TTH 10</td>
<td>Fulton 250</td>
</tr>
<tr>
<td>9303</td>
<td>MI201</td>
<td>MWF 11</td>
<td>Fulton 150</td>
</tr>
<tr>
<td>9304</td>
<td>MI202</td>
<td>TTH 1:30</td>
<td>Fulton 130</td>
</tr>
<tr>
<td>9305</td>
<td>EN101</td>
<td>TTh 3</td>
<td>Gasson 265</td>
</tr>
<tr>
<td>9306</td>
<td>CJ101</td>
<td>MWF 3</td>
<td>Robsham 621</td>
</tr>
</tbody>
</table>