Lecture 19:

Big Data
Updates

- Tuesday: Friends Release
  - Class: Playtesting / spamming friends

- Thursday: Postmortem
  - Must use actual data from actual players
Today

- Survey results
- More real-world examples
- Data collection recommendations
Survey Results!
Most Disliked Topic

- Paper Prototyping: 60%
- Logging: 20%
- Game Mechanics: 5%
- AI: 5%
- PCG: 5%
- Overview: 5%

Big Data
Most Liked Topic

- Data Visualization: 31%
- Level Design: 17%
- AI: 17%
- Game Mechanics: 7%
- Procedural Content Generation: 7%
- Playtesting: 7%
- Logging: 7%
- Learnability and UI: 7%
- Big Data: 6%
Group Activities

- Activities could be improved: 31%
- This is great!: 35%
- I wish he would continue lecturing: 17%
- Other: 17%
I see what you did there…

- I hope to see this kind of analysis next Thursday
Deliverables

- First Prototype: Oct. 7
- Second Prototype: Oct. 23
- Friends Release: Nov. 4
- Newgrounds Release: Nov. 18
- Kongregate Release: Dec. 2
A Venn Diagram

Games with Published Analytics

Great Flash Games

Big Data
Bioware

Georg Zoeller
The Game Design Initiative at Cornell University
Interesting uses of analytics

- Tool for spatial event visualization
- Visualizing crashes
- Incentivizing developers
- Visualizing complaints
- Visualizing the effect of fog of war
- Uncovering human nature
- 3D Visualizations
As of March 2010…

- 1,100,000 games
- 1,141 unique machines
- 250 GB of data
- 38,000,000 datapoints
- 50,000 bugs
  - 4597 unique users
  - 582 machines

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Active Servers

List of all servers that are currently capturing data for this project:

<table>
<thead>
<tr>
<th>Server Host</th>
<th>Version</th>
<th>Started</th>
<th>Status</th>
<th>Connections</th>
<th>Last Update</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="mailto:bye-dm@sky.net">bye-dm@sky.net</a>(19026)</td>
<td>1.0</td>
<td>8/23/2009 1:27:18 PM</td>
<td>listening</td>
<td>0</td>
<td>10/20/2009 11:54:51 AM</td>
</tr>
</tbody>
</table>

Currently Active Games

The following game sessions are currently active on the server:

<table>
<thead>
<tr>
<th>RunId</th>
<th>User</th>
<th>Host</th>
<th>Exe</th>
<th>Started</th>
<th>Length</th>
<th>Platform</th>
<th>Build</th>
<th>Action</th>
<th>Close</th>
</tr>
</thead>
<tbody>
<tr>
<td>96212851172</td>
<td><a href="mailto:jfederline@bioware.adea.com">jfederline@bioware.adea.com</a></td>
<td>10.0.15.42</td>
<td>Game</td>
<td>10/29/2009 11:40:38 AM</td>
<td>08:00:07:23</td>
<td>XBox 360</td>
<td>10035 R on-us</td>
<td>Claim</td>
<td>Close</td>
</tr>
<tr>
<td>133243560374016</td>
<td><a href="mailto:sniderout@bioware.adea.com">sniderout@bioware.adea.com</a></td>
<td>10.0.13.43</td>
<td>Game</td>
<td>10/29/2009 11:39:38 AM</td>
<td>08:00:07:23</td>
<td>XBox 360</td>
<td>10035 R on-us</td>
<td>Claim</td>
<td>Close</td>
</tr>
<tr>
<td>227853679347620</td>
<td><a href="mailto:jasonhill@bioware.adea.com">jasonhill@bioware.adea.com</a></td>
<td>10.0.15.49</td>
<td>Game</td>
<td>10/29/2009 11:37:28 AM</td>
<td>08:00:07:23</td>
<td>XBox 360</td>
<td>10035 R on-us</td>
<td>Claim</td>
<td>Close</td>
</tr>
<tr>
<td>147942852946555</td>
<td><a href="mailto:lees@bioware.adea.com">lees@bioware.adea.com</a></td>
<td>10.0.13.36</td>
<td>Game</td>
<td>10/29/2009 11:09:16 AM</td>
<td>08:00:38:45</td>
<td>XBox 360</td>
<td>10035 R on-us</td>
<td>Close</td>
<td></td>
</tr>
<tr>
<td>2640536545083993</td>
<td><a href="mailto:mappeters@bioware.adea.com">mappeters@bioware.adea.com</a></td>
<td>10.0.14.36</td>
<td>Game</td>
<td>10/29/2009 10:31:34 AM</td>
<td>08:01:16:47</td>
<td>PS3</td>
<td>1077 R on-us</td>
<td>Close</td>
<td></td>
</tr>
<tr>
<td>2732135199554</td>
<td><a href="mailto:jiang@bioware.adea.com">jiang@bioware.adea.com</a></td>
<td>10.0.15.105</td>
<td>Game</td>
<td>10/29/2009 10:15:31 AM</td>
<td>08:01:32:30</td>
<td>XBox 360</td>
<td>3985 R fr-fr</td>
<td>Claim</td>
<td>Close</td>
</tr>
<tr>
<td>3131793309900</td>
<td><a href="mailto:dfurutan@bioware.adea.com">dfurutan@bioware.adea.com</a></td>
<td>10.0.14.63</td>
<td>Game</td>
<td>10/29/2009 10:14:53 AM</td>
<td>08:01:30:08</td>
<td>XBox 360</td>
<td>3985 R on-us</td>
<td>Claim</td>
<td>Close</td>
</tr>
<tr>
<td>09169595970105523</td>
<td><a href="mailto:paul@bioware.adea.com">paul@bioware.adea.com</a></td>
<td><a href="mailto:c95c7e67ab074a9@bioware.adea.com">c95c7e67ab074a9@bioware.adea.com</a></td>
<td>Game</td>
<td>10/16/2009 3:41:26 PM</td>
<td>12:25:06:53</td>
<td>PS3</td>
<td>10017 R on-us</td>
<td>Close</td>
<td></td>
</tr>
<tr>
<td>1997947827320599</td>
<td><a href="mailto:jacques@bioware.adea.com">jacques@bioware.adea.com</a></td>
<td>10.0.15.37</td>
<td>Game</td>
<td>10/5/2009 11:27:46 AM</td>
<td>20:00:20:15</td>
<td>PS3</td>
<td>10017 R on-us</td>
<td>Claim</td>
<td>Close</td>
</tr>
</tbody>
</table>

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Spatial Event Visualization

Session Information
A "session" is defined as a single game, from either a save game or a new one. It starts when you hit "load" and stops when you exit out of the main menu.

All Sessions in the Run: 343487261222820
These are all of the sessions played for a single run (executable start).
Select one to see the areas visited and events hit during the session.

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Logging Crashes

<table>
<thead>
<tr>
<th>MAC Address</th>
<th>00-1e-c9-59-f1-0c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hostname</td>
<td>rob.bioware.ad.ea.com</td>
</tr>
<tr>
<td>IP Address</td>
<td>10.0.15.145</td>
</tr>
<tr>
<td>RAM (GB)</td>
<td>1.01</td>
</tr>
<tr>
<td>CPU</td>
<td>intel(r) core(tm)2 duo cpu e6750 @ 2.66ghz</td>
</tr>
<tr>
<td>CPU Count</td>
<td>2</td>
</tr>
<tr>
<td>CPU Clockrate</td>
<td>2659</td>
</tr>
<tr>
<td>GPU</td>
<td>nvidia geforce 8800 gts 512</td>
</tr>
<tr>
<td>GPU VRAM (MB)</td>
<td>0</td>
</tr>
<tr>
<td>DirectX Version</td>
<td>4.09.00.0904</td>
</tr>
<tr>
<td>Screen Width</td>
<td>1024</td>
</tr>
<tr>
<td>Screen Height</td>
<td>768</td>
</tr>
<tr>
<td>Aspect Ratio</td>
<td>1.333</td>
</tr>
<tr>
<td>Timestamp</td>
<td>10/2/2009 1:17:22 PM</td>
</tr>
</tbody>
</table>

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Logging Crashes

This run crashed! This is the callstack.

```
void __cdecl GraphicsStructures::BuildInstructionTableANB(struct Animation::
bool __cdecl GraphicsStructures::BuildInstructionTableOBJ(struct Animation::
bool __cdecl GraphicsStructures::Object::BuildInstructionTable(double, struct
void __cdecl GraphicsStructures::Object::Animate(double) (00000030)
enum Multithreading::Task::Status __cdecl AnimationTask::Execute{
void __cdecl Multithreading::TaskGroup::ExecuteNextTask(void) (00000054)
enum Multithreading::Task::Status __cdecl Multithreading::TaskGroup::Execute
void __cdecl Multithreading::Thread::ExecuteTrapped(void *) (000000E0)
unsigned int __cdecl Multithreading::Thread::Execute(void) (00000040)
unsigned int __cdecl Multithreading::StandardThread(void *) (00000020)
XapiThreadStartup(00000040)
<Unknown>
```

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Heatmap of Crashes

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Game design initiative at Cornell University

Stability and Usage Metrics

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
## Developer Achievements

This page lists your developer achievements. You can filter the list using the dropdown below.

<table>
<thead>
<tr>
<th>Title</th>
<th>Requirements</th>
<th>Status</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-Head Of The Pack</td>
<td>At the end of a week (Sunday), be the person who logged the most hours across all modules within your department. (1 hour minimum)</td>
<td>achieved</td>
<td>09/01/05</td>
</tr>
<tr>
<td>Bug Hunter</td>
<td>File at least 20 bugs/feedback items through the My Feedback Interface.</td>
<td>achieved</td>
<td>00/09/22</td>
</tr>
<tr>
<td>Ceremonialist</td>
<td>Ceremonialist - Complete Um of Sacred Ashes (Side against Cult)</td>
<td>achieved</td>
<td>09/04/04</td>
</tr>
<tr>
<td>Chanter's Friend</td>
<td>Complete all of the Chanter board quests except for Allison's quest. Reloads are OK, no cheats.</td>
<td>achieved</td>
<td>09/03/15</td>
</tr>
</tbody>
</table>

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Developer Achievements

| Marketing Ploy | At the end of the month, have more hours of active gametime in Single Player than the entire Marketing department together. | achieved | 08/06/26 |

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Developer Achievements: Results

- *Doubled* developer play time for 1 month
- More bugs caught
- Developers got addicted
- Buggy achievements caused anger

Source: Georg Zoeller, “Game Development Telemetry” GDC 2010
Big Data
Optimizing Engines
Optimizing Art

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
How do I…

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
“bug”

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
The road less traveled

68% of players go right

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
Fog of War

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
3D Visualizations

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
More 3D Visualizations

Source: Georg Zoeller, “MMO Rapid Content Iteration” GDC Online 2011
Experimentation is everywhere
But so are ethical challenges

**The Washington Post**
Cornell ethics board did not pre-approve Facebook mood manipulation study

**The Guardian**
Facebook fiasco: was Cornell's study of 'emotional contagion' an ethics breach?

**PCWorld**
Privacy group files FTC complaint over Facebook's 'emotional contagion' study
Institutional Review Board

Cornell University
Office of Research Integrity and Assurance
Milgram Experiment
Stanford Prison Experiment
IRB Approval

Institutional Review Board for Human Participants

Notice of Course Activity Approval

To: Erik Andersen
Date: September 05, 2014
Protocol ID#: 1408004901
Protocol Title: CS-4154

The above referenced Course Activity Project was reviewed by Cornell’s Human Research Protection Program (HRRP) and approved by the Institutional Review Board for Human Participants in accordance with the ethical standards of the Belmont Report: http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html

You or your students must ensure that the welfare of the research participants is protected and that methods used and information provided to gain subject consent are appropriate to the activity. You and your students should familiarize yourself with and conduct the research in accordance with the ethical standards of the Belmont Report.

Please give a photocopy of this approval notice to each student in your class who will be conducting a human participant research project. Acceptance of these terms by students constitutes an understanding that data collection (and allusions to conclusions drawn from these data) from the project(s) covered solely by this approval may never be used for theses, dissertations, articles, or public presentations.

If you have any questions, please contact the IRB office at irbhp@cornell.edu or 254-5162.
Don’t publish data

An Epic Study of Epic Players In Epic’s Epic

An honors thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science
To: Erik Andersen  
Date: September 05, 2014  
Protocol ID#: 1408004901  
Protocol Title: CS-4154  

The above referenced Course Activity Project was reviewed by Cornell’s Human Research Protection Program (HRPP) and approved for the inclusion of human participants in class assignments. This approval does NOT cover students doing research for theses, dissertations, journal articles, public presentations, or other means of disseminating generalizable knowledge gained from these assignments. Such projects require the individual student to complete his or her own Initial Approval Request form before beginning recruitment and data collection.  

You or your students must ensure that the welfare of the research participants is protected and that methods used and information provided to gain subject consent are appropriate to the activity. You and your students should familiarize yourself with and conduct the research in accordance with the ethical standards of the Belmont Report: http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html  

Please give a photocopy of this approval notice to each student in your class who will be conducting a human participant research project. Acceptance of these terms by students constitutes an understanding that data collection (and allusions to conclusions drawn from these data) from the project(s) covered solely by this approval may never be used for theses, dissertations, articles, or public presentations.  

If you have any questions, please contact the IRB office at irbhp@cornell.edu or 254-5162.
No identifiable data

Please enter your name:
Avoid causing harm

Wow, you’re terrible at this.

Great job!

A   B
IRB Approval

Institutional Review Board for Human Participants

Notice of Course Activity Approval

To: Erik Andersen
Date: September 05, 2014
Protocol ID#: 1408004901
Protocol Title: CS-4154

The above referenced Course Activity Project was reviewed by Cornell’s Human Research Protection Program (HRPP) and approved for the inclusion of human participants in class assignments. This approval does NOT cover students doing research for theses, dissertations, journal articles, public presentations, or other means of disseminating generalizable knowledge gained from these assignments. Such projects require the individual student to complete his or her own Initial Approval Request form before beginning recruitment and data collection.

You and your students must ensure that the welfare of the research participants is protected and that methods used and information provided to gain subject consent are appropriate to the activity. You and your students should familiarize yourself with the ethical considerations described in the ethical standards of the Belmont Report: http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html

Please give a photocopy of this approval notice to each student in your class who will be conducting a human participant research project. Acceptance of these terms by students constitutes an understanding that data collection (and allusions to conclusions drawn from these data) from the project(s) covered solely by this approval may never be used for theses, dissertations, articles, or public presentations.

If you have any questions, please contact the IRB office at irbhp@cornell.edu or 254-5162.
You are visitor number

10775
Congratulations! You are player number

10775

to lose on this level
Big Data
Privacy Policy

LandShark Games Pte. Ltd. and its Affiliates ("LandShark" or "we") are dedicated to protecting the privacy rights of our online users ("Users" or "you"). Affiliates shall mean subsidiaries, parent companies, joint ventures and other corporate entities under common ownership. This Online Privacy Policy ("Privacy Policy") has been crafted to inform you as to the ways we collect, store, use, manage, and, in certain circumstances, share the information you provide in connection with any LandShark game or application ("Service"). Please note that the scope of this Privacy Policy is limited to only information collected from LandShark through use of the Service.

By using the Service, you agree to the collection and use of your personal information as outlined in this Privacy Policy. We may amend the Privacy Policy from time to time, and we encourage you to consult the Privacy Policy regularly for changes.

Information Collected and Used

LandShark’s primary objective in collecting user information is to provide and enhance the Service and to enable users to easily navigate and enjoy the Service. When you install any of the LandShark games or applications, we may collect and record the information you provide us voluntarily through www.facebook.com ("Facebook") and/or other social network(s) in accordance with the policies and terms of Facebook and/or other social network(s).

The types of personal information collected in accordance with the activities listed above may vary depending on the activity. This “personal information” is defined as information about you that can be used to contact or identify you. Personal information may include: (1) your first and last name, (2) your profile picture or its URL, (3) your user identification number of the pertinent social network, which is linked to publicly available information such as name and profile picture (4) the login email you provided to the pertinent social network upon registration. Additionally, we may acquire information from you through (1) your access and participation in message boards on the Service (2) your email of a question to our customer support (3) your purchase of a virtual product of any type (4) other sources, including but not limited to newspapers and internet sources such as blogs, instant messaging services, other users of the Service to supplement your profile. Please note that because a third party site manages the purchases of virtual goods offered by the Service, we do not have access to your credit card details and billing information.

LandShark may also invite you to share non-personal information about yourself which may include but is not limited to: (1) your age or date of birth; (2) your gender; (3) information about the hardware you may own; (4) your favorite websites, friend information (such as friends of yours that are also connected to our Service) and fan pages; (5) your physical location and that of your access device (collectively referred to as “non-personal information”). If non-personal information is collected for an activity that also requires personal information, LandShark may combine your non-personal information with your personal information in an attempt to provide you with a better user experience, to improve the quality and value of the Services and to analyze how our Service is used.
What this means for you

- Protect the welfare of your players

- Do not
  - record identifiable information
  - give data to anyone outside the class
  - discuss your data or analysis in
    - public presentations
    - theses
    - journal articles
    - conference papers
Summary

- Release / postmortem next week
- Creative visualizations are awesome and effective
- Be ethical in your data collection
Group Activity

• What hypotheses do you want to test?
• How are you going to record data to test them?
• How will the data answer your hypotheses?
• How will you demonstrate this to the class?