Future-Proof Your Research: Designing for Replicability and Reproducibility
Isabel Brunkan, Zhuo Zhen (Advisor), Kate Keahey (Advisor), Levent Tokosz (Advisor)

Computer Science is Agile. Iteration after iteration, moving quickly to solve new problems, uncover new questions, find the next big thing. Hardware, software, libraries, datasets, experiments - technology becomes outdated almost as soon as it’s released. So why share code? Why share experiments? For replication, to verify results. For education, to train the next generation. For variation, to discover new insights. We seek to understand how to package experiments to encourage experiment exploration and longevity by replicating an AlexNet reproduction.

**The Challenge**

Research replicability is crucial but computer science research has a unique challenge due to resource variety, availability, upgrading and data and code accessibility.

- **How do you achieve...?**
  - Repetition: Same Experiment + Original Artifacts
  - Replication: Same Experiment + Recreate Artifacts
  - Reproduction: Same Experiment Idea + New Artifacts
  - Variation: Repeat/Replicate + Modification

**Goal:** Replicate the original AlexNet model

**In Practice:** Package an experiment reproducing AlexNet on the Stanford Dogs dataset.

**Hardware Implementation**

<table>
<thead>
<tr>
<th>Original AlexNet + ImageNet</th>
<th>Chameleon + Stanford Dogs</th>
<th>Kaggle + Stanford Dogs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Images</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2 million (Training)</td>
<td>20,000</td>
<td>20,000</td>
</tr>
<tr>
<td><strong>Classes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,000</td>
<td>120</td>
<td>120</td>
</tr>
<tr>
<td><strong>GPU</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (NVIDIA GTX 580)</td>
<td>P100, M40, K80, RTX-6000, V100</td>
<td>1 Tesla P100</td>
</tr>
</tbody>
</table>

**Evaluation and Extensions**

- Choose stable, accessible datasets for replicability
  - Tensorflow’s Datasets API is faster and easier to use than Kaggle’s API, which requires credentials
- Experiment with your own data + transfer learning
  - Use additional data and pre-trained models to increase accuracy.
  - My frenchton puppy was classified as a chihuahua

**Conclusions**

1. **Package for the Future by Separating Container Setup and Experiment Scripts**
   - a. **Container Scripts:** Can be reused for different experiments and easily adjusted for different hardware, especially those that haven’t been invented yet.
   - b. **Experiment Scripts:** Easily adjust or add variation to the experiment without affecting your container set up.

2. **Increase Readability**
   - Code is hidden in scripts which can be expanded to direct focus as needed.

**Citations**

