# Formal Statement

## *Approximate Algorithmic Image Matching to Reduce Online Storage Overhead of User Submitted Images*

**Purpose:**

Current systems allow for duplicate data to be uploaded to the server by multiple users, requiring large amounts of storage space. The proposed research will target the issue of storing user submitted image data while reducing storage overhead through a series of algorithms and analysis methods.

**Example:**

Websites such as imgur, photobucket, as well as many others allow users to upload and share images free of cost. User A wishes to share an image with friends and uploads said image to the server and is provided a link that the public can access the image from. User B also wishes to share the exact same image (possibly a different resolution) and uploads it to the server. Both users have added identical – or near identical information to the server and have caused redundancy.

**Use of proposed system:**

As with the earlier example, user A wants to share an image and uploads it to the photo sharing website. The image will be processed, determined to be unique, and will subsequently be added to the server. User B also wishes to share the same image (possibly a different resolution) and proceeds to upload it to the server, it is processed, and is determined to be a duplicate of an image already on the server.

If the new image is lower resolution or the same resolution as the existing one, the site will show the user the most probable match and display the higher/same resolution image and the one they wish to upload. If the user determines it is the same image, the new image will be discarded and a link will be generated for the existing image. If the new image is a higher resolution, and the user states it is the same, the new image will be uploaded and the old one preserved to keep users from maliciously overwriting a previous share.

**Benefit:**

By using this system, storage costs can be reduced in addition to electrical costs as less server nodes are required to operate as storage space. In addition, the users receive better service as a higher resolution image alternative may be recommended to the one they wish to share.