Definition of Project:

The Mass Digitization and Preservation Initiative (MDPI) is a multi-campus Indiana University effort funded by the Office of the President to digitize roughly 350,000 audiovisual items by the university's bicentennial in 2020. In 2013, the university started digitizing roughly 325,000 items on a variety of audio and video formats, and in 2017, the university began working towards its goal of digitizing 25,000 film items.

Definition of Best Copy Selection:

Best Copy Selection on the MDPI project can be defined as the process of comparison between various copies, generations, versions, and elements of a given title with the goal of determining the item, or items, that are best suited for digitization by the university's digization vendor. There are two broad selection goals during this process – a final, complete version for immediate access and pre-print elements for digital preservation. On a more granular level, this process involves inspecting and comparing all of the items across several criteria:

1. Descriptive Identification: This involves recording gauge, generation, reel number, can size, footage, base, stock, color/black and white, image content type, aspect ratio, sound/silent, language, soundtrack type, sound content type, as well as edgecode and date.
2. Title Confirmation: This can normally be confirmed by a title card on the print itself; however, in the event of a missing title card or a non-existent title card, this confirmation requires a more in-depth examination of the item's content and comparison to another, already confirmed, copy of the suspected title.
3. Completeness: Some strategies to check for completeness include checking the first frame on an item and the last image on an item to check for missing footage at the head or tail, counting splices/breaks, and evaluating frames on either side of splices/breaks for evidence of missing footage. It is also helpful to wind through multiple copies of the same title using a gang synch footage counter. When using a gang synch, one aligns all reels being compared to the same frame in the gang synch. Typically, it is easiest to choose the first frame of the first scene after a hard cut. As one winds through the reels, if one or more is missing footage, this will become apparent as the frame alignment of the different reels will be lost in the gang synch – i.e. the scenes in the reel that is missing footage will end sooner than the reel that is not missing footage. Comparing the ends of scenes will also give one a frame accurate documentation of how much footage is missing. It is best practice to use a gang synch only with items in good physical condition as damage on the item increases the likelihood that the item could become misaligned on the gang synch, increasing the likelihood of the item sustaining further damage. It is also for this reason that it is not recommended that pre-print elements, particularly those designated for preservation, be run through a gang synch. It is also best practice to make sure that the gang synch is placed on the bench and that all of the reels on the inspection bench are spaced such that the film is running through the gang synch completely straight.
4. Color Fade: If the item is a positive, color fade will present as varying degrees of color loss. At the very beginning of color fade, the image might appear to be losing saturation in just the yellow or blue areas. At the height of color fade, the image will appear

completely pink or magenta. If the item is a color negative or internegative, the orange in the image will appear light, almost melon-colored, as opposed to a vibrant, well-saturated orange hue in a non-faded color negative or internegative. If the film is faded, it requires more intervention on the part of the Scan Operator and/or the Post-Production Manager to restore the film's original color.

5. Scratches: Abrasions on the surface of the base and the emulsion that can be assessed by casting a diffuse light over the surface of the film. These surface abrasions affect the user's viewing experience and require more intervention on the part of the Post-Production Manager to remove these scratches during the digital restoration process.
6. Warpage: Physical deformation of the celluloid base that results in the film not maintaining a flat, straight appearance. There are many different types of warpage. The most common types found on the MDPI project include roping, which has the appearance of twisting; cupping, which has the appearance of the edges of the film raising towards one side of the film, most commonly the emulsion; and edgewave, which has the appearance of the edges of the film raising and lowering in wave-like patterns. This warpage affects the Scan Operator's ability to maintain a consistent focus across the entirety of the frame and requires the use of a different gate on the scanner, called the Warp Gate. This warp gate necessitates that our digitization vendor's scanner be run at a significantly slower pace, 6 frames per second.
7. Shrinkage: Condition where the space between the perforations on a film item decreases. Different organizations have different rules regarding shrinkage, but IULMIA's is that if a film is more than 0.75% shrunken, it cannot be run on equipment. Our digitization vendor has a sprocketless scanner, which vastly eliminates the risk of damage to film items, particularly in the area of shrinkage, but shrinkage can still affect the framing of images – i.e. if the film is shrunken it is more likely that frame lines will be visible in the captured frame, without further cropping of the frame.
8. Conservation concerns: These include mold, foreign objects, inspects, and chemical decomposition issues such as "vinegar syndrome". All of these issues can cause problems in the digitization process, particularly mold, which is a health concern for the people working with the materials and can eat away at the emulsion on the film, and "vinegar syndrome", which can lead to physical abnormalities such as warpage, shrinkage, plasticizer exudation, channeling, brittleness, and hockey pucking. It should be noted that "vinegar syndrome" alone does not affect an item's suitability for selection and/or scanning. Instead it is the existence of the accompanying physical abnormalities that make it more difficult to scan an item and/or contribute to image degradation that affects the quality of the scan.
9. Dirtiness: Unless there are conservation concerns associated with cleaning an item, all items on the MDPI project are ultrasonically cleaned, meaning most routine dirt and dust will be removed before scanning; however, there are certain types of grime that cannot be removed, or that are difficult to remove, that affect image quality. These include fingerprints, stains, rust, and excessive adhesive residue.
10. Edge damage/Perforation damage: Nicks, tears, or other abrasions along the edges of the film or along the perforations. Since our digitization vendor uses sprocketless scanners, this type of damage, unless excessive, will not normally affect scan quality;

however, the presence of edge and perforation damage can be indicative of other types of physical damage caused by equipment such as scratches, breaks, tears, and soundtrack damage. Also, if a scanner with sprockets is being used, it is paramount to repair all perforation damage before running the film through the scanner to prevent further damage.

11. Soundtrack damage: On optical tracks this can include scratches, dirt, focus issues, and tears. On magnetic tracks this can include warpage, flaking, and binder degradation, also known as sticky shed syndrome. All of these issues can affect not only sound quality, but also the basic ability to scan the item in the first place.
12. Image degradation: Anything that affects image quality. This includes water damage, redox blemishes, plasticizer exudation, channeling, silver mirroring, and ferrotyping.
13. Sound quality: When possible, soundtracks are listened to, to check for issues such as hiss, crackle, wow, and flutter that could have an adverse effect on the viewing experience.

These criteria are then weighed against each other across the multiple copies and generations in order to choose the item, or items, exhibiting the least amount of issues and the least amount of issues that will create an adverse viewing experience for the user.

<u>Guiding Principles/Selection Criteria of Best Copy Selection:</u>

At the beginning of the MDPI Project, the Film Archivist and the Assistant Film Archivists developed a set of guidelines for performing best copy selection. These guidelines intended to maintain best practices of film inspection/evaluation and preservation research while also balancing the short timeframe and resource limitations of the MDPI project. In order to meet demand for the MDPI project in a timely manner, it was decided that the Assistant Film Archivists would only evaluate the first 100' of each reel. Within this first 100', the Assistant Film Archivist would record and evaluate as much physical, conditional metadata as would normally be gathered during a full, exhaustive print condition report. Examples of metadata collected include: gauge, generation, can size, footage, soundtrack type, edgecode, scratches, warpage, shrinkage, AD level, edge damage, perforation damage, title on print, text on leader, wind orientation, number of splices, color process, color fade, brittleness, stock, base, aspect ratio, dirt, image degradation, sound quality, and completeness. Additionally, if there were obvious stock changes later in the reel, the Assistant Film Archivist would wind to these stock changes and perform a conditional analysis of each stock on the reel. Further, if the Assistant Film Archivist felt it was necessary to analyze the entirety of the reel to make a selection decision, the Assistant Film Archivist would do so. Examples of reasons for evaluating a reel in its entirety include, but are not limited to, if the reel was unidentified and the Assistant Film Archivist needed to perform a more detailed content analysis or if there were multiple versions of a specific title and which version a particular item was needed to be determined.

Prints were selected so that there would be an immediate, complete, access copy for each title as well as so that the Post-Production Manager could use it as a reference in his digital preservation and restoration projects. If there were both black and white and color copies available for a particular title, a color copy was preferred over a black and white copy, even if the color copy exhibited color fade, with the assumption that the access file would be color

corrected either at the point of scanning or in post-production. Sometimes, in cases of severe color fade or an important title, both a color print and a black and white print would be sent. Color reversal prints, which normally did not exhibit any signs of color fade, were the first choice in terms of selection, assuming condition and completeness were relatively equal across all copies. Additionally, answer prints, if available, were preferred over projection prints because they were an earlier generation, and would thus normally have a higher image quality than projection prints. Answer prints were also generally circulated much less than the projection prints and would thus be in better physical condition. When evaluating answer prints; however, it was vitally important to evaluate them for not only completeness, but also timing and notations. For some titles, there would be multiple generations of answer prints as the filmmakers attempted to finalize the aesthetic of the final version. As this process took place, the filmmakers would routinely leave timing and editing notations for the lab on the answer print itself, indicating that it was not the final version used as a reference for creating the projection prints. The project would try to avoid sending answer prints that were not the final version, but sometimes this was the only complete, final edit of the film available, so sometime an answer print with notations was sent as a reference copy for the Post-Production Manager to digitally restore the film.

Elements were selected for digital preservation and restoration. Typically, this meant selecting one image element – an internegative, interpositive, a negative, or A/B Roll combination – and one track element – a separate optical track negative, separate optical track positive, or separate magnetic track. In terms of the image element, an intermediate, typically an internegative, was preferred if the title was in color or a negative if the title was in black and white. It was determined that the project would go back to the intermediate generation as opposed to the A/B Roll generation for most titles because of the limited timeframe and resources on the MDPI project. Going back to the A/B Roll generation would have meant scanning double the number of reels for each title, limiting the total number of titles that could be digitize, and would also mean exponentially more post-production time for each item to make it accessible. For particularly high-profile titles or if intermediate elements did not exist, A/B rolls were used for digital preservation. Internegatives were chosen as opposed to interpositives simply because interpositives did not exist for most of the titles in the collection. If a title had a complete interpositive, that would be selected instead because it is an earlier generation. In terms of soundtrack elements, optical tracks were preferred over magnetic tracks for purely conservation reasons. Typically, the magnetic tracks for a given title had high levels of vinegar syndrome and were exhibiting flaking, stickiness, and other condition issues that made them more difficult to work with and affected sound quality. If given the choice between an optical track positive and an optical track negative, the optical track positive would be selected because optical track positives generated a better sound quality on our digitization vendor's scanners.

These guidelines were used as general principles, but were adapted to fit the particularities and needs of individual collections.

<u>Workflow for Best Copy Selection:</u>

1. All known items for a particular title would be pulled from storage and delivered to the assigned Assistant Film Archivist.
2. All delivered items for a particular title would be organized by generation.
3. All items for a particular title of the same generation would be inspected together and compared to each other.
4. During inspection, descriptive and conditional metadata would be collected and recorded for all items along with documentation about which items were compared to each other and why each item was selected or rejected for digitization.
5. The most complete item in the best physical condition for each generational category would be selected, packed, and sent to the digitization vendor.

Importance of Best Copy Selection:

Best Copy Selection has many advantages, principally, it ensures that resources devoted to the digitization project are not wasted. Particularly when there are multiple copies, multiple generations, and multiple versions, it is vitally important to determine which item/s is/are in the best physical condition, are the most complete, and are thus best suited for digital preservation. If this process is not done and instead all items for a given title are sent for digitization, one will end up digitizing many duplicate, incomplete, or otherwise unsuitable items, resulting in a lot of wasted resources and wasted time.

Best copy selection also has the added benefit of producing a lot of conditional metadata about the collections, which produces substantially more intellectual control over the collections than existed before, particularly when the collections have previously only had the most basic level of processing. Additionally, best copy selection results in previously unidentified or misidentified materials being properly identified. This happens because, through the course of performing best copy selection, the Assistant Film Archivist becomes incredibly familiar with the collections they are working with and can use this knowledge to recognize materials others might not have been able to.

Tangentially, best copy selection can also result in better conservation of the collection. As an item is being handled for best copy selection, items that are on reels can be transferred to cores and items that are currently stored in non-archival containers can be re-housed in archival, non-reactive film cans.