# What is Enhanced Match Point?

Enhanced Match Point is a record-matching algorithm used by the MOBIUS INN-Reach union catalog to identify whether resources contributed by the local MOBIUS systems are already represented in the catalog. Records already represented in the MOBIUS union catalog may become the new “master record” or may be attached to the existing master record as holdings. Enhanced Match Point uses a multi-tiered matching algorithm to pinpoint matching records and replaces the previous matching algorithm, which looked primarily at the OCLC # located in the MARC 001 field.

# What is a Master Record?

The MOBIUS union catalog record display is similar to those in the local catalogs. Author, title, publication information, etc., appear toward the top of the display.



Description, subject headings, control numbers, etc. appear toward the bottom of the display.



In between these two sections, the holdings for MOBIUS’s local systems are displayed.



In a local Sierra catalog, the primary display is taken from a bibliographic record, while the holdings display is taken from “attached” item or holdings records. In INN-Reach, both the primary and holdings display are taken from bibliographic records. The bib record that appears in the primary display is called the Master Record; the records that appear in the holdings line are called Institutional Records.

# How Does Enhanced Match Point Work?

Enhanced Match Point follows a multi-tiered matching algorithm, designed to identify the best match for records to the MOBIUS union catalog. The matching algorithm is comprised of three steps: record matching, record verification, and Master Record determination.

## Step 1: Record Matching

### Has this record already been contributed (local system/record number)?

INN-Reach first looks at the bib record number (e.g., b20647052) and the local system (e.g., 6wash) contributing the record, to determine whether the record has previously been contributed to MOBIUS. If INN-Reach finds a matching local system/bib number combination, the matching process ends.

### Is there a match for this record?

Before the implementation of Enhanced Match Point, INN-Reach only looked for matching records based on OCLC # (located in the MARC 001 field). Enhanced Match Point uses a tier of sequential match points, listed below. When INN-Reach finds a match on any one of the match points in sequence, the process moves to the Step 2, Record Validation.

1. OCLC Number, MARC tag 001
2. LC Control Number (LCCN), MARC tag 010, subfield a (if the subfield repeats, only the first subfield is consulted)
3. LC Control Number (LCCN) (canceled/invalid), MARC tag 010, subfield z (if the subfield repeats, only the first subfield is consulted)
4. International Standard Book Number (ISBN), MARC tag 020, subfield a (if the subfield repeats, each is consulted)
5. International Standard Serial Number (ISSN), MARC tag 020, subfield a (if the subfield repeats, each is consulted)
6. Other Standard Identifier, MARC tag 024, subfield a (if the subfield repeats, each is consulted)
7. International Standard Book Number (ISBN) (canceled/invalid), MARC tag 020, subfield z (if the subfield repeats, each is consulted)
8. International Standard Serial Number (ISSN) (canceled/invalid), MARC tag 020, subfield z (if the subfield repeats, each is consulted)
9. Other Standard Identifier (canceled/invalid), MARC tag 024, subfield z (if the subfield repeats, each is consulted)
10. International Standard Serial Number (ISSN) (incorrect), MARC tag 020, subfield y (if the subfield repeats, each is consulted)
11. Standard Number, MARC tag 989, subfield a (if the subfield repeats, only the first subfield is consulted)

If no match is found on any of the above fields, the process moves to Step 3, Master Record determination.

## Step 2: Record Validation

In certain cases, INN-Reach may perform extra steps in verifying that contributed records that find a match to one of the above fields truly are for the same resource. INN-reach performs the following checks when a match is found. If any check fails, the process moves to Step 3, Master Record determination, without making the subsequent checks.

1. Large Print: INN-Reach checks to see whether or not the word “large” is contained in one of the following fields of both the incoming and matching records.
	1. MARC tag 245, subfield h (General material Designation)
	2. MARC tag 250 (Edition statement)
	3. MARS tag 300 (Physical Description)

If the word “large” is contained in one of the above fields in both records, or is not contained in one of the above fields in both records, proceed to the Reproduction check. If one record contains the word “large” and the other does not, the record do not match; proceed to Step 3, Master Record Determination.

1. Reproduction
	1. Do the records contain a MARC tag 533 (Reproduction Note)?
		1. If neither contain a 533, proceed to the GMD check
		2. If both contain a 533, proceed to 2.b.
		3. If one contains a 533 and the other does not, the record do not match; proceed to Step 3, Master Record Determination
	2. Are the 533 subfields b (Place of reproduction), c (Agency responsible for reproduction), and d (Date of reproduction) identical?
		1. If they are identical, proceed to 2.c.
		2. If they are not identical, the records do not match; proceed to Step 3, Master Record Determination
	3. Are the 533 subfields f (Series statement of reproduction) identical>
		1. If they are identical, proceed to the GMD check
		2. If they are not identical, the records do not match; proceed to Step 3, Master Record Determination
2. GMD: Are the 245 subfields h (General Material Designation) identical?
	1. If they are identical, proceed to the next check
	2. If they are not identical, the records do not match; proceed to Step 3, Master Record Determination
3. Title (not performed for a 989 field match): Are the normalized[[1]](#footnote-1) 245 fields identical?
	1. If they are identical, proceed to the Imprint check;
	2. If they are not identical the records do not match; proceed to Step 3, Master Record Determination
4. Imprint (not performed for a 989 field match)
	1. Are the normalized first word of the MARC tag 260 (Imprint), subfield a (Place of publication) or b (Name of publisher) identical?
		1. If they are identical, proceed to 5.b.
		2. If they are not identical, the records do not match; proceed to Step 3, Master Record Determination
	2. Are the first, normalized, non-bracketed dates in the MARC tag 260, subfield c (Date of publication) identical? (Not performed for records with Leader byte 7 “s”)
		1. If they are identical, the records match; proceed to Step 3, Master Record Determination
		2. If they are not identical, the records do not match; proceed to Step 3, Master Record Determination

## Step 3: Determination of the Master Record

### No Matching Record

If no matching record is found in Step 1, or if any of the validation checks fail in Step 2, INN-Reach designates the contributing record as a new Master Record.

### Matching Record Found

If a matching record is found in Step 1, and all of the validation checks pass in Step 2, INN-Reach then makes a determination as to whether the contributing record should replace the current Master Record, or be added to the existing Master Record as an Institutional Record. INN-Reach makes the following comparisons between the contributing record and the current Master Record.

1. Do the records contain 008 (Fixed Length Data) fields?
	1. If both do or both do not, proceed to the next check;
	2. If one does and the other does not, the record that does becomes the new Master Record.
2. Do the records contain 505 (Formatted Contents Note) fields?
	1. If both do or both do not, proceed to the next check;
	2. If one does and the other does not, the record that does becomes the new Master Record.
3. Do the records contain 520 (Summary) fields?
	1. If both do or both do not, proceed to the next check;
	2. If one does and the other does not, the record that does becomes the new Master Record.
4. Do the records contain 655 (Genre/Form) fields?
	1. If both do or both do not, proceed to the next check;
	2. If one does and the other does not, the record that does becomes the new Master Record.
5. Do the records contain 007 (Physical Description Fixed Field) fields?
	1. If both do or both do not, proceed to the next check;
	2. If one does and the other does not, the record that does becomes the new Master Record.
6. Do the records contain 880 (Alternate Graphic Representation) fields?
	1. If both do or both do not, proceed to the next check;
	2. If one does and the other does not, the record that does becomes the new Master Record.
7. What is the Encoding Level (Leader byte 17) of the records?
	1. If both records have the same Encoding Level, proceed to the next check;
	2. If one record has a lower Encoding Level, the record becomes the new Master Record.
8. What is the contributing local system?
	1. If both records are from MERLIN, Saint Louis University, or Washington University, proceed to the next check;
	2. If both records are not from MERLIN, Saint Louis University, or Washington University, proceed to the next check;
	3. If only one record is from MERLIN, Saint Louis University, or Washington University, that record becomes the new Master Record.
9. Which record was contributed first? – The record that was contributed earliest becomes the Master Record

# Appendix 1: Normalization rules for bibliographic headings

1. Take the first 150 characters of the Heading (excluding MARC tag and indicators, but including subfields and delimiter characters);
2. Strip non-filing characters (taken from the 245, second indicator or the SKIP fixed-length field)
3. Strip apostrophes and diacritics;
4. Replace the ampersand character (&) with the word “and”;
5. Strip all special characters, except “+”, “#”, “$”, “%”, and “@”;
6. Replace subfield delimiters and punctuation with “null” characters (i.e., “space”);
7. Collapse multiple sequential “null” characters into single “null” characters;
8. Take the first 125 characters of the resulting string.
1. See Appendix 1 [↑](#footnote-ref-1)