

## Menuju *Automated Data Creation*: Manajemen Metadata Bibliografi Nasional Indonesia (BNI) di Era Digital

**Ratna Pramudyawardhani**

Perpustakaan Nasional RI

ratnapramudyaw@gmail.com

### **Definisi *Automated Data Creation***

Menurut Schlesinger, S. & Gravzy, A. (2021), *automated data creation* adalah sebuah rangkaian proses pengumpulan data menggunakan mesin sampai pada pemasukan data ke dalam sebuah *database/ data repository/* atau *data warehouse* yang dapat menawarkan wawasan yang lebih signifikan bila dikelola dengan baik. Kemampuan pembuatan, pengumpulan, penyimpanan, dan analisis data otomatis ini memberikan peluang dalam memanfaatkan data penting *real-time* untuk digunakan di berbagai kasus. Tujuannya adalah untuk meningkatkan efisiensi organisasi, merampingkan biaya, meningkatkan profitabilitas, dan memastikan keselamatan dan kesehatan karyawan.

Dalam dunia perpustakaan, *automated data creation* memungkinkan pustakawan membuat, mengumpulkan, mengelola, dan menyebarkan data secara otomatis. Proses ini memerlukan bantuan *information technology* (IT) yang memadai dan *machine/ application user*, dalam hal ini, pustakawan dengan kemampuan IT

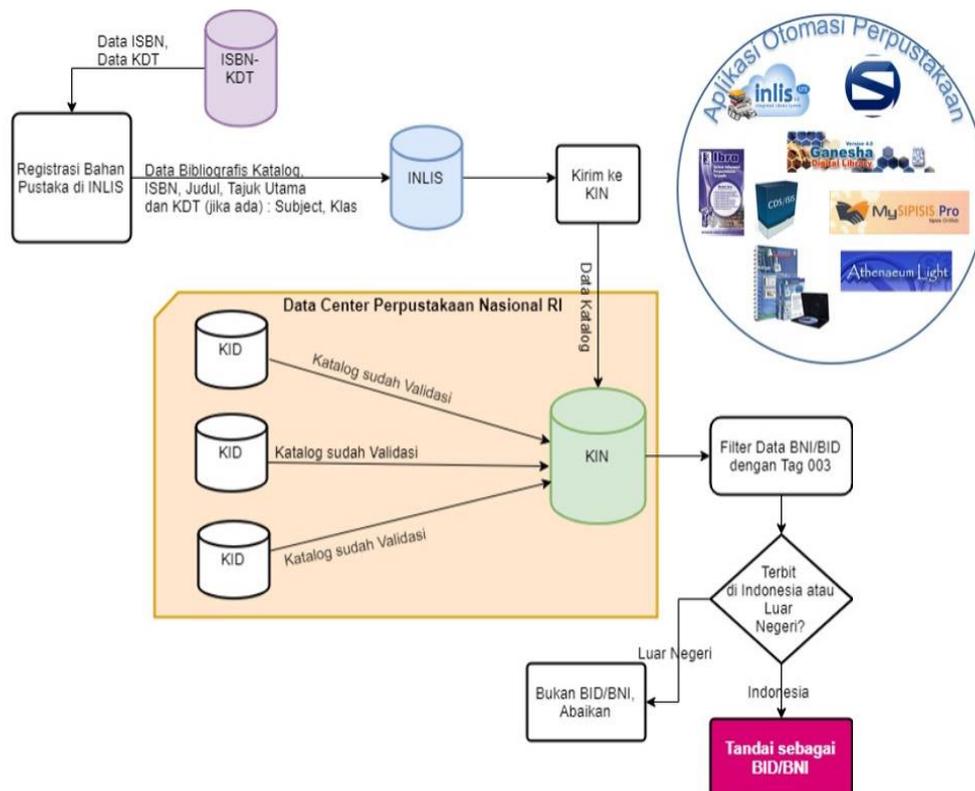
yang baik. Tujuan dilakukannya proses ini adalah untuk dapat mengolah data sebanyak mungkin tanpa memerlukan sumber daya yang terlalu besar, serta mengefektifkan kinerja organisasi.

Nilai otomatisasi sebagai bagian dari siklus pengumpulan data akan terus berkembang seperti halnya teknologi. Perangkat pengumpulan data akan berevolusi dari otonom (bekerja sendiri) menjadi terhubung sebagai bagian dari *Internet of Things* (IoT) dan terkoordinasi (bertindak bersama satu sama lain dengan intervensi manusia yang seminimal mungkin atau tanpa campur tangan manusia sama sekali). Semakin banyak informasi yang tersedia dari jenis perangkat pengumpulan data ini, manusia akan dapat mengontrol sumber daya yang tersedia dengan lebih baik dan menjadi lebih produktif.

Hal penting yang harus diperhatikan dalam *automated data creation* adalah pustakawan harus tetap meluangkan waktu untuk melakukan verifikasi data untuk memastikan kelengkapan dan akurasi data. Pada tahap ini, pustakawan

berfungsi sebagai validator data untuk mengurangi tingkat kesalahan pada data. Standar proses pemeriksaan data perlu diaplikasikan dengan menghilangkan proses entri manual yang sebenarnya.

dari pengolahan Koleksi Serah Simpan digunakan sebagai salah satu dasar penyusunan Bibliografi Nasional Indonesia dan Bibliografi Daerah. (2) Perpustakaan Nasional wajib menyusun dan menerbitkan/ memublikasikan Bibliografi Nasional Indonesia di dalam negeri



**Gambar SEQ Gambar \\* ARABIC 1. Siklus pengelolaan data BNI**

*Sumber: Pedoman Teknis Pengelolaan Bibliografi Nasional Indonesia Tahun 2022*

### **Automated Data Creation pada Pengelolaan Bibliografi Nasional Indonesia (BNI)**

Peraturan Pemerintah Republik Indonesia Nomor 55 Tahun 2021 tentang Pelaksanaan Undang-undang Nomor 13 Tahun 2018 tentang Serah Simpan Karya Cetak dan Karya Rekam Pasal 17 Ayat 1 sampai dengan Ayat 3 menyebutkan, “(1) Hasil

dan di luar negeri secara berkala. (3) Perpustakaan Provinsi wajib menyusun dan menerbitkan/ memublikasikan bibliografi daerah secara berkala.” Sesuai hal itu, Perpustakaan Nasional RI bertanggung jawab untuk melakukan pengawasan terbitan nasional guna melestarikan hasil budaya bangsa dan menjaga catatan peradaban Indonesia. Wujud pengawasan terbitan ini adalah berupa

pengelolaan dan penerbitan Bibliografi Nasional Indonesia (BNI).

Bibliografi nasional merupakan kumulasi atas produksi terbitan nasional (komersial dan non-komersial, termasuk terbitan pemerintah) dari suatu negara, yang diterbitkan secara teratur, dengan mengikuti standar internasional yang berlaku (IFLA, 2009). Dengan demikian dapat disimpulkan bahwa BNI merupakan kumulasi daftar terbitan nasional di seluruh wilayah Indonesia. BNI diterbitkan dalam kurun waktu tiga bulanan atau empat kali dalam setahun.

Pengelolaan BNI masuk ke dalam tugas dan fungsi Pusat Bibliografi dan Pengolahan Bahan Perpustakaan (Pusbiola), Perpustakaan Nasional RI. Manajemen metadata dalam pengelolaan BNI dilakukan melalui aplikasi Katalog Induk Nasional (KIN). Data terbitan diperoleh dari pelaksanaan Undang-Undang Nomor 13 Tahun 2018 tentang Serah Simpan Karya Cetak dan Karya Rekam setelah melalui proses pengajuan *International Standard Book Number* (ISBN) dan Katalog dalam Terbitan (KDT).

Secara alur, data BNI akan diregistrasi oleh Direktorat Deposit dan Pengembangan Koleksi Perpustakaan (Depobangkol) ke dalam aplikasi INLIS Perpustakaan Nasional RI. Data ini kemudian diteruskan

kepada Substansi Pengolahan Hasil Karya Cetak dan Karya Rekam Pusbiola untuk dilengkapi dan dimasukkan ke dalam aplikasi KIN, disunting metadata-nya, dan divalidasi. Data tersebut dipilah-pilah sesuai dengan kebutuhan data BNI.

Sementara itu, proses pengumpulan data terbitan yang masih tercecer di berbagai daerah idealnya dilakukan melalui integrasi antara aplikasi KIN dan aplikasi KIPI (Katalog Induk Perpustakaan Indonesia) yang ditujukan untuk pengelolaan Bibliografi dan Katalog Induk Daerah. Data mentah ini dikelola oleh masing-masing Dinas Perpustakaan dan Kearsipan Daerah (DPAD) pada aplikasi KIPI dan secara otomatis disetorkan ke dalam *database* aplikasi KIN secara berkala melalui *tools* OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) yang disematkan pada masing-masing *database*. OAI-PMH adalah sebuah mekanisme pertukaran metadata dari *database repository*.

Selanjutnya, data yang masuk ke aplikasi KIN akan disunting metadatanya dan divalidasi kualitas datanya oleh validator menjadi sebuah informasi terbitan yang sudah valid dan terverifikasi. Data ini nantinya menjadi acuan bagi proses *copy catalog* di perpustakaan lainnya dan menjadi bahan referensi bagi para pengguna BNI, seperti pustakawan, tenaga perpustakaan, civitas akademika, peneliti, dan masyarakat umum lainnya dalam mengadakan penelitian maupun akuisisi koleksi. Metadata

terbitan ini juga akan diteruskan kepada *Universal Bibliographic Control* (UBC), yakni sebuah konsep pengawasan bibliografi secara internasional yang diprakarsai oleh *International Federation of Library Associations and Institutions* (IFLA).

### **Tantangan Pustakawan Pengelola Bibliografi Nasional Indonesia (BNI)**

Mewujudkan siklus manajemen data yang baik pada proses pengelolaan BNI tidaklah mudah karena untuk memungkinkan tersajinya pertukaran data diperlukan standar data yang sama. Standar metadata yang digunakan oleh Perpustakaan Nasional RI adalah INDOMARC (*Indonesian Machine Readable Cataloging*). MARC pertama kali dikembangkan oleh *Library of Congress*, sedangkan MARC *Indonesia version* lebih dikenal sebagai INDOMARC, yaitu format MARC Indonesia. Format INDOMARC adalah wujud implementasi format ISO 2709 untuk Indonesia yang memungkinkan penggunaannya tukar-menukar metadata bibliografis dengan media terbaca mesin (Indreswari NS, 2020: 1).

Selain menggunakan INDOMARC sebagai standar metadata BNI, Perpustakaan Nasional RI juga menggunakan *Resources Description and Access* (RDA) sebagai standar deskriptif metadata data

bibliografisnya. RDA dapat mengakomodasi berbagai jenis bentuk terbitan dan koleksi perpustakaan di era yang sudah serba digital ini. RDA juga menyediakan titik akses yang lebih luas dan lengkap atas data-data terbitan dan katalog perpustakaan.

Kendala yang selama ini terjadi dalam pertukaran metadata bibliografis DPAD maupun instansi lain yang menyediakan data terbitan dengan Perpustakaan Nasional RI adalah masih adanya perbedaan standar metadata dan standar deskriptif data bibliografis. Masih banyak perpustakaan yang menggunakan standar deskripsi bibliografis berupa *Anglo-American Cataloging Rules 2<sup>nd</sup> Edition* (AACR2) dan standar metadata lain, seperti *Dublin Core*, *Metadata Object Description Schema* (MODS), dan *Metadata Encoding and Transmission Specification* (METS). Hal inilah yang menjadi tantangan bagi pustakawan Pusbiola dalam proses pengelolaan BNI.

Satu-satunya cara yang dapat dilakukan untuk mengatasi masalah ini adalah dengan konversi metadata *Dublin Core*/MODS/METS ke dalam metadata MARC, contohnya seperti yang banyak dilakukan adalah *Dublin Core to MARC crosswalk*. Secara keseluruhan, MARC memiliki elemen data yang lebih lengkap dibandingkan dengan *Dublin Core* yang hanya mempunyai lima belas elemen dasar. MARC mempunyai banyak ruas data yang berfungsi sebagai titik akses informasi bibliografis. Masing-masing ruas ini berisikan *tag* yang

ditandai dengan tiga digit kode angka. Pada setiap *tag*, terdapat sub-sub ruas data yang ditandai dengan simbol dolar (\$) dan ada juga indikator yang merupakan kode satu karakter untuk memberikan informasi tambahan mengenai subruas data. Dalam hal ini dibutuhkan sebuah aplikasi yang mampu mengidentifikasi pemetaan metadata (*mapping*) dan mengeksekusi metadata *crosswalk* pada proses *data harvesting* dari metadata lain ke dalam elemen MARC.

**Tabel 1. Mapping metadata MARC dan Dublin Core (*metadata crosswalk*)**

| MARC Fields                        | DC Element  | Implementasi on Notes     |
|------------------------------------|-------------|---------------------------|
| 100, 110, 111, 700, 710, 711       | Contributor |                           |
| 720                                |             |                           |
| 651, 662                           | Coverage    |                           |
| 751, 752                           |             |                           |
|                                    | Creator     | Creator element not used. |
| 008/07-10                          | Date        |                           |
| 260\$c\$g                          |             |                           |
| 500-599, except 506, 530, 540, 546 | Description |                           |
| 340                                | Format      |                           |
| 856\$q                             |             |                           |

| MARC Fields                  | DC Element | Implementasi on Notes  |
|------------------------------|------------|--|
| 020\$a, 022\$a, 024\$a       | Identifier |  |
| 856\$u                       |            |  |
| 008/35-37                    | Language   |  |
| 041\$a\$b\$d\$e\$f\$g\$h\$j  |            |  |
| 546                          |            |  |
| 260\$a\$b                    | Publisher  |  |
| 530, 760-787\$o\$t           | Relation   |  |
| 506, 540                     | Rights     |  |
| 534\$t                       | Source     |  |
| 786\$o\$t                    |            |  |
| 050, 060, 080, 082           | Subject    |  |
| 600, 610, 611, 630, 650, 653 |            |  |
| 245, 246                     | Title      | Repeat dc: title for each. Some applications may wish to include 210, 222, 240, 242, 243, and 247. |
| Leader06, Leader07           | Type       | See Appendix 2 for Leader-Type rules.  |
| 655                          |            |  |

Sumber: Development and MARC Standards Office (2008)

Pustakawan pengelola BNI bertugas untuk menyunting metadata, melengkapi, dan melakukan validasi terhadap data terbitan pada aplikasi KIN. Apabila disandingkan dengan berbagai tantangan yang dihadapi dalam proses pengelolaan BNI, maka diperlukan peran penting dari pustakawan sebagai *gate keeper* dalam mewujudkan data-data yang berkualitas. Oleh karena itu, pustakawan yang bertugas mengelola BNI harus terus mempertajam kemampuan literasi informasinya dalam penyuntingan dan validasi data bibliografis.

Pada saat proses konversi metadata ke dalam MARC selesai dilakukan, tidak semua *tag* dalam aplikasi KIN akan terisi lengkap karena adanya perbedaan jumlah elemen antara MARC dan metadata lain, maupun perbedaan standar deskripsi data bibliografi yang digunakan masing-masing perpustakaan. Dalam hal ini, pustakawan pengelola BNI harus memahami *tag* mana saja yang perlu diisi, dilengkapi, diperbaharui, dan disesuaikan dengan standar deskripsi data bibliografis yang digunakan oleh Perpustakaan Nasional RI, yaitu RDA. Apabila pustakawan tidak memiliki kemampuan literasi informasi yang cukup, maka data yang akan dihasilkan tidak lagi valid. Metadatanya pun tidak akan menyediakan titik akses yang memadai sesuai dengan standar RDA yang digunakan. Beberapa *tag* yang harus dilengkapi oleh pustakawan pada saat

penyuntingan dan validasi data BNI ditunjukkan pada tabel di bawah ini.

**Tabel 2. *Tag* yang perlu dilengkapi pada proses penyuntingan dan validasi BNI**

| Tag | Deskripsi metadata                           |
|-----|--|
| 003 | Control No. ID                               |
| 007 | Ruas tetap deskripsi fisik (keterangan umum) |
| 008 | Unsur data yang panjangnya tetap             |
| 015 | Nomor BNI                                    |
| 020 | Nomor untuk ISBN                             |
| 022 | Nomor untuk ISSN                             |
| 040 | Sumber pengatalogan                          |
| 041 | Kode bahasa                                  |
| 043 | Kode wilayah                                 |
| 082 | Nomor klasifikasi DDC                        |
| 084 | Nomor panggil                                |
| 100 | Entri utama                                  |
| 110 | Entri utama nama badan korporasi             |
| 111 | Entri utama nama pertemuan                   |
| 240 | Judul seragam                                |
| 245 | Judul dan penanggung jawab                   |
| 250 | Pernyataan edisi                             |
| 264 | Impressum                                    |
| 300 | Deskripsi fisik                              |
| 336 | Jenis isi                                    |
| 337 | Jenis media                                  |
| 338 | Jenis wadah                                  |
| 440 | Pernyataan seri/entri tambahan judul         |
| 500 | Catatan umum                                 |
| 504 | Catatan bibliografi                          |
| 600 | Entri tambahan subjek                        |
| 610 | Entri tambahan subjek nama badan korporasi   |
| 611 | Entri tambahan subjek nama pertemuan         |
| 650 | Entri tambahan subjek topik                  |
| 651 | Entri tambahan subjek wilayah                |

| Tag | Deskripsi metadata                  |
|-----|-------------------------------------|
| 700 | Entri tambahan nama orang           |
| 710 | Entri tambahan nama badan korporasi |
| 711 | Entri tambahan nama pertemuan       |
| 730 | Entri tambahan judul seragam        |
| 740 | Entri tambahan judul                |
| 850 | Badan yang memiliki                 |
| 852 | Lokasi                              |
| 856 | Akses dan lokasi elektronik         |
| 990 | Nomor deposit                       |
| 999 | Nomor induk                         |

### **Kemampuan yang Dibutuhkan Pustakawan untuk Mendukung Proses *Automated Data Creation***

Peran pustakawan di era digital telah mengalami perkembangan. Pustakawan di era ini, mau tidak mau akan dihadapkan pada fenomena banjir data dan informasi yang mengubah peran pustakawan menjadi ahli data maupun ahli informasi. Banyaknya data bibliografi dari berbagai daerah di Indonesia yang dihasilkan melalui proses *data harvesting* harus diolah dengan berpegang pada standar katalogisasi yang digunakan oleh Perpustakaan Nasional RI.

Data yang masuk ke dalam basis data pengelolaan BNI masih merupakan data setengah jadi yang perlu dipilah-pilah melalui tahapan *data cleaning*. *Data cleaning* merupakan salah satu bagian dari manajemen data yang bertujuan untuk menyaring data yang berkualitas dan siap diolah. *Data cleaning* merupakan proses

pendukung dalam pengolahan data yang dilakukan untuk menyunting, memperbaharui, melengkapi, dan membuat data menjadi lebih rapi. Ada kalanya pula data perlu dihapus. Menurut Elgabry (2019), ciri-ciri data yang berkualitas adalah sebagai berikut:

1. valid, menggambarkan realitas sebenarnya;
2. akurat, nilainya benar atau mendekati benar;
3. komplet, informasinya lengkap;
4. konsisten, sama persis di berbagai *dataset*;
5. seragam, punya satuan yang sama (misalnya semua uang ditunjukkan dengan rupiah, alamat ditulis dengan kode pos, dan lain-lain).

Untuk melakukan semua proses di atas, pustakawan yang bertugas dalam mengelola BNI perlu mengembangkan kemampuan manajemen data yang baik. Walek (2019) menjelaskan bahwa kemampuan yang dibutuhkan pustakawan dalam pengelolaan data adalah pengetahuan terkait data domain. Pustakawan juga perlu memiliki pengetahuan dan kemampuan dalam memfasilitasi berbagi data, *linked data*, siklus manajemen data penelitian, kontrol kualitas, pemrosesan data, perencanaan manajemen data, serta kemampuan dalam memahami dan memfasilitasi kebutuhan penyimpanan data. Pustakawan yang bertugas dalam mengelola BNI harus memiliki pengetahuan dan kemampuan dalam menerapkan metadata, metadata standar, ontologi, dan informasi terstruktur. Selain itu, pustakawan data perlu memiliki kompetensi dalam pemetaan data, katalogisasi, dan *data harvesting* sehingga secara umum, pustakawan

perlu memiliki penguasaan IT yang baik agar dapat bekerja sama dengan ahli IT maupun peneliti dengan baik.

### Daftar Pustaka

- Elgabry, O. (1 Maret 2019). *The ultimate guide to data cleaning: when the data is spewing garbage. Toward Data Science*.  
<https://towardsdatascience.com/the-ultimate-guide-to-data-cleaning-3969843991d>.
- Development and MARC Standards Office. (24 April 2008). *MARC to Dublin Core Crosswalk. Library of Congress*.  
<https://www.loc.gov/marc/marc2dc.html>.
- Indreswari NS, dkk. (2020). *INDOMARC*. Jakarta: Perpustakaan Nasional RI.
- Nasrullah, dkk. (2022). *Pedoman Teknis Pengelolaan Bibliografi Nasional Indonesia*. Jakarta: Perpustakaan Nasional RI.
- Peraturan Pemerintah No. 55 Tahun 2021 tentang Peraturan Pelaksanaan Undang-undang Nomor 13 Tahun 2018 tentang Serah Simpan Karya Cetak dan Karya Rekam [JDIH BPK RI]. (2021).  
<https://peraturan.bpk.go.id/Home/Details/163995/pp-no-55-tahun-2021>.
- Schlesinger, S. & Gavzy, A. (12 Juli 2021). *Automation of the data lifecycle: focus on data creation. RTInsights*.  
<https://www.rtinsights.com/automation-of-the-data-lifecycle-focus-on-data-creation/>.
- Walek, A. (2019). Data librarian and data steward - new tasks and responsibilities of academic libraries in the context of Open Research Data implementation in Poland. *Przegląd Biblioteczny*, 497-512.  
<https://doi.org/10.36702/pb.634>.