

Analisis saintometrika dalam penelusuran potensi penelitian *biomedicine* berdasarkan data publikasi keanekaragaman hayati Indonesia

Scientometrics analysis to investigate the potency of *biomedicine* research based on Indonesia biodiversity publication

IRENE MUFLIKH NADHIROH[✉], RIA HARDIYATI, MIA AMELIA, RIZKA RAHMAIDA, TRI HANDAYANI

Pusat Penelitian Perkembangan Ilmu Pengetahuan dan Teknologi, Lembaga Ilmu Pengetahuan Indonesia, Jl. Jend. Gatot Subroto Kav 10 Jakarta Selatan 12710, Jakarta. Telp./Fax. +62-21-5225206, ✉email: irene.muflikh@gmail.com

Manuskrip diterima: 26 Agustus 2017. Revisi disetujui: 15 September 2017.

Abstrak. Nadhiroh IM, Hardiyati R, Amelia M, Rahmaida R, Handayani T. 2017. Analisis saintometrika dalam penelusuran potensi penelitian *biomedicine* berdasarkan data publikasi keanekaragaman hayati Indonesia. *Pros Sem Nas Masy Biodiv Indon 3: 345-350*. Indonesia sebagai negara *mega biodiversity*, memiliki potensi kekayaan keanekaragaman hayati yang besar. Salah satunya adalah penemuan obat *biomedicine* yang berasal dari sumber daya genetik yang berasal dari Indonesia. Penelitian terkait dengan sumber daya genetik keanekaragaman hayati di Indonesia telah banyak dilakukan. Beberapa diantaranya terkait dengan proses penemuan obat baru dengan berbasiskan sumber daya genetik keanekaragaman hayati Indonesia. Saintometrika adalah pendekatan kuantitatif dalam penelitian mengenai perkembangan ilmu pengetahuan dan teknologi. Penelitian ini bertujuan untuk menganalisis potensi penemuan obat *biomedicine* berdasarkan publikasi ilmiah menggunakan pendekatan saintometrika. Salah satu penerapan saintometrika adalah mengungkap potensi dan tren ilmu pengetahuan dari data publikasi ilmiah. Banyak metode yang dapat digunakan terkait dengan penerapan tersebut, salah satunya metode *Co-word analysis*. Data yang digunakan dalam penelitian ini adalah data publikasi ilmiah internasional keanekaragaman hayati Indonesia terkait *biomedicine* yang berasal dari basis data Scopus. Berdasarkan data tersebut, mayoritas tahapan penelitian yang dilakukan masih pada tahap *drug discovery*. Oleh sebab itu, analisis lebih lanjut diperlukan untuk memberikan panduan dalam melakukan penelitian yang lebih efisien. Hasil penelitian menunjukkan adanya potensi besar pada beberapa spesies keanekaragaman hayati Indonesia, berdasarkan hubungannya dengan kata kunci terkait penemuan obat, seperti anti-kanker, antioxidant, dan lain-lain. Hasil ini dapat diaplikasikan dalam manajemen penelitian, yaitu dengan memberikan dana lebih besar pada penelitian tersebut, sehingga dapat menghasilkan produk yang bermanfaat.

Kata kunci: *biomedicine*, Indonesia, keanekaragaman hayati, saintometrika

Abstract. Nadhiroh IM, Hardiyati R, Amelia M, Rahmaida R, Handayani T. 2017. *Scientometrics analysis to investigate the potency of biomedicine research based on Indonesia biodiversity publication. Pros Sem Nas Masy Biodiv Indon 3: 345-350*. Indonesia as a mega biodiversity country, has great potential with a wealth of biodiversity. One of them is the discovery of biomedical drugs derived from genetic resources originating from Indonesia. Research related to genetic resources of biodiversity in Indonesia has been done a lot. Some of them are related to new drug discovery process based on genetic resources of Indonesia's biodiversity. Scientometrics is a quantitative approach in research on the development of science and technology. This study aims to analyze the potential discovery of biomedical drugs based on scientific publications using the approach Scientometrics. One application of saintometrika is to reveal the potential and scientific trends of scientific publication data. Many methods can be used related to the application, one of which is *Co-word analysis* method. The data used in this study is the international scientific publication data biomedical Indonesia related biomedical data originating from Scopus. Based on these data, the majority of research stages conducted are still at drug discovery stage. Therefore, further analysis needs to be needed to provide guidance in conducting more efficient research. The results indicate the great potential of some species of biodiversity in Indonesia, judging by their relation to *keywords* related to drug discovery, such as anti-cancer, antioxidant, and others. This result can be applied in research management, ie by providing more funds in the study, so as to produce useful products.

Keywords: *biomedicine*, Indonesia, biodiversity, scientometrics, *co-word*

PENDAHULUAN

Data publikasi ilmiah bidang keanekaragaman hayati (kehati) Indonesia telah berkembang dalam jumlah yang signifikan, berdasarkan Handayani et al. (2016) diketahui

bahwa rata-rata pertumbuhan publikasi bidang keanekaragaman hayati adalah 9% per tahun, sejak 1990 hingga 2015. Signifikansi ini menunjukkan bahwa bidang keanekaragaman hayati merupakan salah satu bidang yang banyak diminati. Bukan hanya bagi peneliti domestik,

tetapi juga peneliti dari luar negeri. Indonesia sebagai negara mega-biodiversitas telah menjadi daya tarik tersendiri bagi peneliti. Begitu banyak sumber daya alam keanekaragaman hayati yang menarik untuk dijadikan objek penelitian. Dengan begitu besarnya sumber daya alam keanekaragaman hayati, dengan ribuan bahkan ratusan ribu spesies yang dimiliki Indonesia, sangatlah besar upaya yang harus dilakukan untuk mempelajarinya. Bahkan tujuan besar yang harus dicapai bukan hanya untuk dapat mempelajarinya, tetapi juga mengambil manfaat secara ekonomis sehingga dapat dirasakan oleh seluruh masyarakat Indonesia. Untuk mengetahui posisi Indonesia dalam mengeksplorasi dan memanfaatkan sumber daya keanekaragaman hayati dapat dilakukan dengan studi saintometrika. Beberapa penelitian diketahui menggunakan pendekatan saintometrika untuk mengetahui perkembangan riset bidang keanekaragaman hayati serta dinamika yang terjadi didalamnya (Liu et al. 2011)).

Potensi didefinisikan sebagai sebuah kemampuan yang belum terlihat secara seutuhnya, bahkan terkadang belum terlihat sama sekali. Dalam konteks penelitian, potensi sebuah hasil penelitian dijelaskan dalam beberapa indikator. Biasanya terukur dalam sebuah nilai yang dapat didefinisikan, seperti *Technology Readiness Level* (TRL). TRL sendiri mengukur tingkat kesiapan sebuah teknologi untuk menjadi sebuah produk yang dapat dipasarkan. Sedangkan, penelitian merupakan proses investigasi yang sistematis terhadap sebuah objek untuk mendapatkan fakta. Penelitian sendiri memiliki resiko akan kegagalan. Einstein pernah mengatakan bahwa, “*if we knew what it was we were doing, it would not be called research, would it?*”. Hal ini menegaskan sisi ketidakpastian dari sebuah penelitian. Penelitian merupakan sebuah kegiatan dimana, peneliti (pada umumnya) tidak dapat menebak bagaimana hasil akhir yang akan didapatkan. Oleh sebab itu, kegagalan dalam penelitian merupakan sebuah penemuan. Bahkan terkadang terdapat “*unintended outcome*” yang menjadi sebuah produk/teknologi. Potensi yang cukup besar dari sumber daya keanekaragaman hayati adalah dalam bidang *biomedicine*, lebih tepatnya adalah potensi penemuan obat baru berbasis sumber daya keanekaragaman hayati Indonesia. Proses penemuan obat baru tidaklah mudah dan memerlukan waktu lama. Beberapa penelitian telah mengaplikasikan *text mining* dalam mengidentifikasi potensi obat bagi suatu penyakit berdasarkan fakta-fakta yang didapatkan dalam publikasi ilmiah (Blagosklonny 2002; Perez-Iratxeta et al. 2002). Penerapan *text mining* yang telah populer adalah identifikasi penyebab molekuler penyakit dengan menggunakan fakta dari database dan literatur. Sebagai contoh, peneliti *biomedicine* harus mengatasi gen set besar yang telah diusulkan sebagai gen kandidat untuk penyakit tertentu (Rebholz-Schuhmann et al. 2012).

Indonesia sebagai salah satu negara mega biodiversitas memiliki tanggung jawab cukup besar untuk dapat memanfaatkan sumber daya alam yang dimiliki untuk kebaikan umat manusia. Penelitian terkait *biomedicine* dengan menggunakan biodiversitas Indonesia pun telah banyak dilakukan. Tetapi, hasil penelitian yang telah diproduksi menjadi sebuah produk yang bermanfaat masih

belum signifikan. Oleh sebab itu artikel ini bertujuan untuk memberikan suatu pendekatan dalam menelusuri potensi *biomedicine*, terutama dalam hal penemuan obat baru, berdasarkan data publikasi ilmiah bidang keanekaragaman hayati.

BAHAN DAN METODE

Data

Data yang digunakan dalam penelitian ini merupakan kumpulan publikasi ilmiah internasional yang berhubungan dengan keanekaragaman hayati Indonesia yang berasal dari Scopus (www.scopus.com). Scopus merupakan basis data publikasi ilmiah internasional yang bereputasi tinggi. Jenis publikasi yang digunakan dalam penelitian ini dibatasi pada artikel yang diterbitkan pada tahun 1990-2015. Pemilihan tahun terbit dimulai dari tahun 1990 didasarkan pada penyelenggaraan Konvensi Keanekaragaman Hayati di Rio de Janeiro tahun 1992. Salah satu tanggung jawab negara yang ikut menandatangani konvensi tersebut adalah melaksanakan program pelatihan dan penelitian untuk perlindungan dan pemanfaatan keanekaragaman hayati secara berkelanjutan. Hal tersebut menjadi acuan bagi beberapa negara untuk meningkatkan penelitian keanekaragaman hayati. Penelitian ini menggunakan beberapa kombinasi kata pencarian yang digunakan dalam pengambilan data, yaitu nama kelas yang tercatat pada tujuh kingdom (Animalia, Plantae, Bacteria, Fungi, Protozoa, Chromista, dan Archaea) yang dikombinasikan dengan kata “Indonesia” pada *affiliation country*. Kombinasi kata pencarian tersebut dijaring pada bagian “*title, abstract, and keyword*”. Adapun bahasa komputer yang digunakan untuk mendapatkan data adalah “(TITLE-ABS-KEY (Name of class) AND AFFIL-COUNTRY (Indonesia)) AND DOCTYPE (ar) AND PUBYEAR > 1989 AND PUBYEAR < 2016”. Hasil penelusuran dengan menggunakan kombinasi kata pencarian tersebut menunjukkan bahwa jumlah artikel yang dihasilkan oleh kelas dari kingdom “plantae” cukup rendah. Hal ini disebabkan oleh pola penulisan artikel yang berhubungan dengan kingdom “plantae” jarang menyebutkan nama class pada bagian “*title, abstract, and keyword*”. Oleh karena itu, dilakukan penambahan kata pencarian dalam menjaring artikel keanekaragaman hayati dengan menggunakan tujuh nama kingdom (“animalia”, “plantae”, “bacteria”, “fungi”, “protozoa”, “chromista”, dan “archaea”) yang dikombinasikan dengan kata “Indonesia” pada *affiliation country*. Selain itu, penambahan data juga dilakukan dengan menjaring artikel melalui kata pencarian “plant” untuk kingdom plantae dan “animal” untuk kingdom. Total artikel yang didapatkan dari kata pencarian tersebut adalah sebanyak 2.432 artikel.

Metode

Penelitian ini akan menggunakan *keyword based network*, yaitu salah satu aplikasi *co-word* yang dapat digunakan dalam memahami informasi mengenai iptek yang terkandung dalam sebuah dokumen ilmiah, baik paten, publikasi ilmiah seperti jurnal ilmiah, prosiding

maupun buku ilmiah. Pada dasarnya *keyword based network* merupakan implementasi dari *text mining* yang merupakan salah satu alat dalam membuat data text (tulisan) yang belum terstruktur menjadi memiliki struktur yang lebih jelas. Dalam perkembangannya *keyword based network* banyak diimplementasikan dalam usaha memahami morfologi dari data paten, untuk menemukan peluang teknologi baru (Lee et al. 2009a,b).

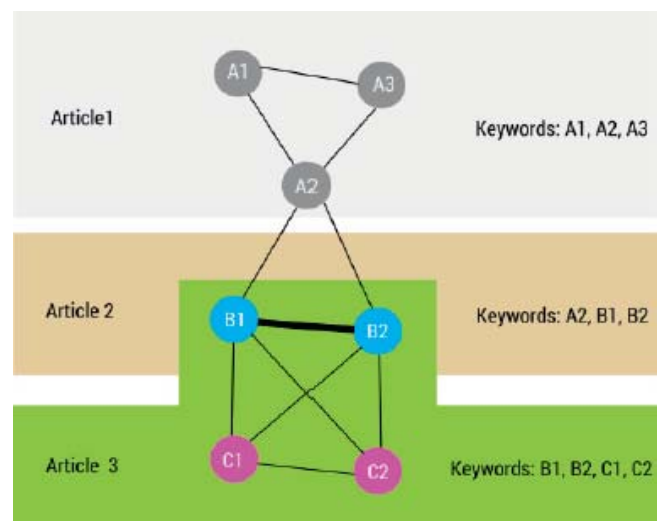
Keyword based network akan memberikan informasi bagaimana *keyword* pada publikasi ilmiah/paten saling berhubungan. Selain itu, jumlah insiden atau banyaknya *keyword* tersebut muncul pada publikasi ilmiah/paten di waktu tertentu merupakan salah satu indikasi bahwa *keyword* tersebut adalah isu yang paling banyak menjadi perhatian dalam publikasi. Dengan mengetahui informasi mengenai *keyword*, akan memberikan pandangan terhadap perkembangan ilmu pengetahuan pada bidang tertentu dan di waktu tertentu. Jika data ini dapat dilakukan penelusuran lebih jauh, hal ini dapat menjelaskan beberapa fenomena yang terjadi saat ini, misalnya terkait dengan tidak berjalannya program pemerintah atau akan adanya stagnasi. Walau demikian, itu menghasilkan output seperti itu perlu dilakukan penelusuran mendalam dengan studi literatur ataupun metode kualitatif lainnya.

Sebelum dilakukan analisis dengan *keyword based network*, akan dilakukan pemilihan terhadap artikel ilmiah yang memang membahas mengenai aplikasi *biomedicine*. Pemilihan dilakukan dengan melakukan penjaringan dari data publikasi keanekaragaman hayati yang telah dikumpulkan. Publikasi yang dipilih adalah publikasi yang mengandung beberapa kata pada kolom "title", "source title", "author keyword" dan "indexed keyword". Kata-kata yang digunakan sebagai penjaring adalah "drug", "drugs" dan "anti". Dengan dilakukan penjaringan, maka publikasi yang dilakukan analisis lebih lanjut adalah publikasi yang memang terkait dengan penelitian *biomedicine*.

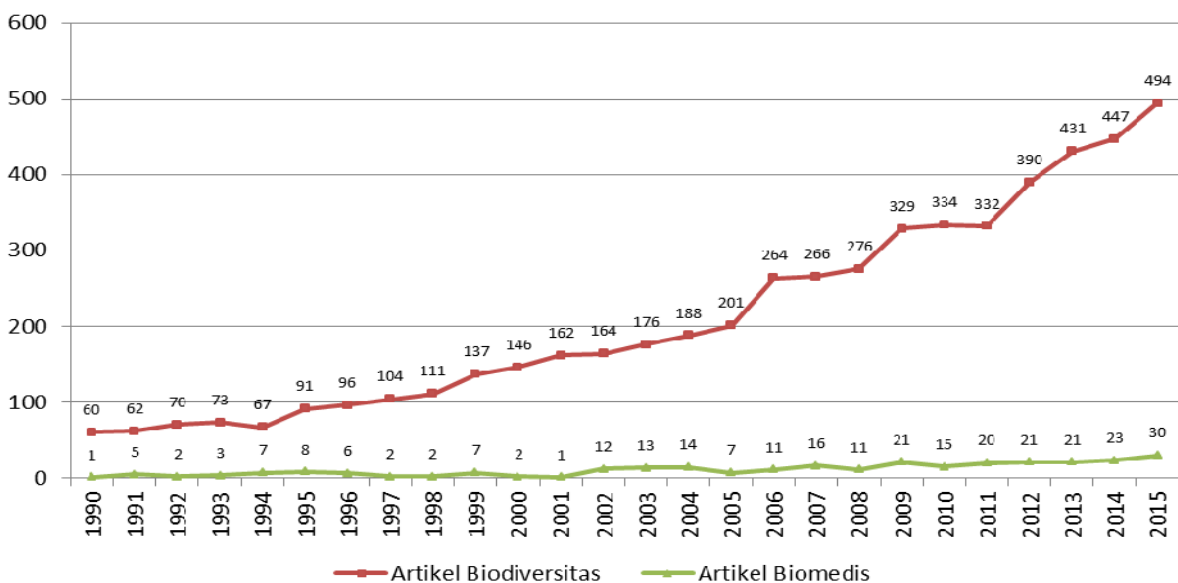
HASIL DAN PEMBAHASAN

Deskripsi penelitian *Biomedicine* berdasarkan publikasi bidang keanekaragaman hayati

Terdapat 389 artikel yang berhasil diidentifikasi dengan proses penjaringan yang dijelaskan pada bagian metode. Dari 389 artikel tersebut dilakukan pemilihan lagi dengan membaca satu per satu artikel tersebut. Terdapat 2 artikel yang dikeluarkan karena tidak sesuai dengan definisi penelitian *biomedicine* yang diharapkan pada penelitian ini. Kemudian terdapat 106 artikel yang harus dikeluarkan karena, mayoritas pada kelompok artikel ini merupakan penelitian obat untuk hewan, yang tidak masuk pada cakupan penelitian. Sehingga terdapat 281 artikel yang diikutikan dalam analisis selanjutnya.



Gambar 1. Formulasi *keyword based network*



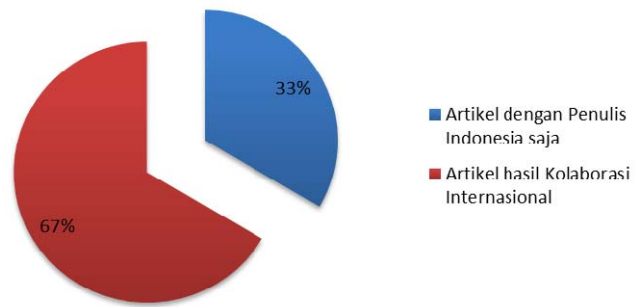
Gambar 2. Sebaran jumlah artikel biodiversitas dan *biomedicine* keanekaragaman hayati Indonesia, 1990-2015. Sumber data: Scopus.com (diolah)

Table 1. Sebaran artikel *biomedicine* penelitian keanekaragaman hayati Indonesia berdasarkan Nama Jurnal (1990-2015)

No	Nama Jurnal	Jumlah Artikel
1	Journal of Natural Products	49
2	International Journal of Pharmacy and Pharmaceutical Sciences	12
3	Chemical and Pharmaceutical Bulletin	8
4	International Journal of PharmTech Research	8
5	American Journal of Tropical Medicine and Hygiene	7
6	Malaria Journal	6
7	Natural Product Communications	6
8	Journal of Ethnopharmacology	6
9	Pakistan Journal of Biological Sciences	5
10	Phytochemistry	4
11	Marine Drugs	4
12	Journal of Natural Medicines	4
13	International Journal of Pharma and Bio Sciences	3
14	International Journal of Pharmacognosy and Phytochemical Research	3
15	Research Journal of Pharmaceutical, Biological and Chemical Sciences	3
16	International Journal of Phytomedicine	3
17	International Journal of ChemTech Research	3
18	Veterinary Microbiology	3
19	Phytotherapy Research	3
20	The Southeast Asian Journal of Tropical Medicine and Public Health	3

Sumber: Scopus.com (diolah)

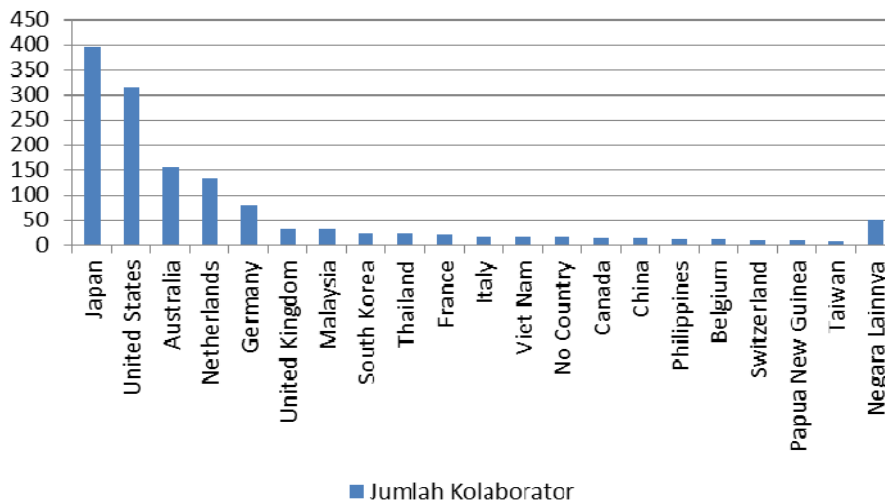
Jumlah artikel bidang keanekaragaman hayati mengalami pertumbuhan yang cukup signifikan (Gambar 2). Begitu pula dengan jumlah artikel bidang *biomedicine*, terlihat adanya peningkatan meskipun tidak begitu signifikan. Kesenjangan yang cukup besar antara keduanya menjadi indikasi bahwa penelitian bidang keanekaragaman hayati yang dilakukan oleh peneliti dengan afiliasi Indonesia masih berkisar pada eksplorasi spesies. Sedangkan pada penelitian pada bidang *biomedicine* masih belum banyak berkembang.



Gambar 2. Persentase Artikel berdasarkan Kegiatan Kolaborasi. Sumber: Scopus.com (diolah)

Mayoritas artikel mengenai aplikasi *biomedicine* berdasarkan keanekaragaman hayati di Indonesia dipublikasi di *Journal of Natural Product*, yaitu sebanyak 51 artikel. Sebanyak 12 artikel dipublikasi di *International Journal of Pharmacy and Pharmaceutical Sciences*. Tabel 1 memperlihatkan 20 besar jurnal dengan jumlah artikel *biomedicine* terbanyak. Terlihat bahwa artikel-artikel tersebut mayoritas dipublikasi pada jurnal terkait *pharmacy, natural product* serta *medicine*. Jumlah artikel yang tercakup pada 20 besar jurnal tersebut adalah sebanyak 167 artikel atau sebanyak 51% dari total artikel bidang *biomedicine* yang terjaring dari publikasi bidang keanekaragaman hayati.

Dari 281 artikel bidang *biomedicine* yang terjaring, terdapat 2.538 penulis. Sebanyak 1.144 penulis (45%) merupakan penulis dengan afiliasi Indonesia, sedangkan sisanya berasal dari berbagai negara. Jepang merupakan negara kolaborator terbesar dengan sebanyak 395 penulis Jepang yang berkolaborasi dengan penulis Indonesia, diikuti dengan Amerika Serikat sebanyak 315 penulis dan Australia sebanyak 156 penulis.



Gambar 3. Sebaran jumlah kolaborator berdasarkan asal negara. Sumber: Scopus.com (diolah)

Co-word analysis pada “antidiabetic” keyword dan relasinya

Salah satu penerapan *biomedicine* yang paling populer adalah dalam pencarian obat ataupun alternatif pangan sehat yang antidiabetes. Selain itu, untuk melihat aplikasi yang lebih jelas, maka kata kunci “antidiabetic” diambil sebagai contoh kasus. Gambar 5 memperlihatkan bagaimana *keyword network* yang telah terbentuk pada kata kunci “antidiabetic”. Beberapa kata kunci yang memiliki relasi dengan “antidiabetic” adalah “eugenia polyantha extract”, “1 (2,4,6 trihydroxyphenyl) 9 phenylnonan 1 one”, “glycosidase inhibitor”, “averrhoa bilimbia extract”, “catharanthus roseus”, “catharanthus roseus extract plus averrhoa bilimbia extract”, “butylated hydroxyl toluene”, “4 hydroxy 3 methoxy benzoic acid”, “1 (2, 6 dimethoxyphenyl) 9 phenylnonan 1 one”, “3,5 dihydroxy benzoic acid”, dan lain-lain. Kata kunci tersebut dapat diindikasikan sebagai potensi pengembangan obat bagi antidiabetes.

Pembahasan

Saintometrika sudah menjadi pendekatan yang umum digunakan untuk mengetahui perkembangan dari suatu bidang ilmu. Saintometrika dapat memberikan beberapa informasi seperti pertumbuhan jumlah artikel, hingga pemetaan aktor yang terkait di suatu bidang ilmu tertentu. Jeong dan Kim (2010) menunjukkan bagaimana struktur intelektual pada bidang *biomedicine*. Lievrouw et al. (1987) melakukan analisis triangulasi untuk mengidentifikasi *invisible actor* yang berperan dalam perkembangan bidang *biomedicine*. Pendekatan *text mining* dalam aplikasi *biomedicine* sendiri juga telah banyak dilakukan. Beberapa penelitian yang menggunakan pendekatan saintometrika pun telah banyak dilakukan. Chaman Sab et al. (2017) melakukan pemetaan terhadap potensi *biomedicine* di India pada tahun 2012 hingga 2016 (Chaman Sab et al. 2017). Zhu et al. (2013) menggunakan *biomedicine text mining* dalam aplikasi di bidang penelitian mengenai kanker. Rebholz-Schuhmann et al. (2013) juga melakukan penelusuran hubungan *gene-disease* pada penyakit diabetes melitus tipe 2 berdasarkan publikasi serta artikel terkait yang tersedia pada web.

Perkembangan ilmu pengetahuan dan teknologi di Indonesia dihadapkan pada tantangan untuk memberikan bukti atas investasi penelitian dan pengembangan yang selama ini telah diberikan oleh pemerintah. Sebagai seorang peneliti juga memiliki tanggung jawab atas penggunaan dana APBN dalam melakukan penelitian. Tantangan terbesar adalah dalam memberikan peningkatan nilai ekonomi dari hasil litbang. Analisis saintometrika dengan pendekatan *co-word* dan *text mining* dapat menjadi

salah satu pendekatan yang digunakan. Dengan melakukan analisis lebih mendalam pada data publikasi ilmiah serta mengidentifikasi penelitian yang potensial untuk diberikan insentif lebih, terkait meningkatkan TRL dari hasil penelitian tersebut. Selain itu analisis saintometrika juga dapat menunjukkan bagaimana evolusi yang terjadi pada suatu bidang penelitian, sehingga jika memang telah terjadi akumulasi pengetahuan yang dianggap cukup dapat dijadikan rujukan bahwa penelitian tersebut sudah cukup dan memiliki kapabilitas untuk diberikan insentif lebih untuk ditingkatkan TRL nya.

DAFTAR PUSTAKA

- Blagosklonny MV. 2002. P53: an ubiquitous target of anticancer drugs. *Intl J Cancer* 98 (2): 161-166.
- Chaman Sab M, Dharani Kumar P, Biradar BS. 2017. Mapping of Indian biomedicine Research: A Scientometric Analysis of Research Output During 2012 -2016. *Indian J Inform Sour Serv* 7 (2): 2231-6094.
- Handayani TH, Amelia M, Rahmaida R, Hardiyati R, Nadhiroh IM. 2016. Kajian Saintometrika Perkembangan Publikasi Ilmiah Keanekaragaman Hayati Indonesia sebagai Bahan Rekomendasi Kebijakan Arah Penelitian Keanekaragaman Hayati Nasional. Laporan Teknis Penelitian. PAPPITEK-LIPI.
- Jeong S, Kim H.-G. 2010. Intellectual structure of biomedical informatics reflected in scholarly events. *Scientometrics* 85 (2): 541-551.
- Lakitan B, Hidayat D, Herlinda S. 2012. Scientific productivity and the collaboration intensity of Indonesian universities and public R&D institutions: Are there dependencies on collaborative R&D with foreign institutions? *Technol Soc* 34 (3): 227-238.
- Lee S, Bozeman B. 2005. The Impact of Research Collaboration on Scientific Productivity. *Soc Stud Sci* 35 (5): 673-702.
- Lee S, Yoon B, Lee C, Park J. 2009a. Business planning based on technological capabilities: Patent analysis for technology-driven roadmapping. *Technol Forecast Soc Change* 76 (6): 769-786.
- Lee S, Yoon B, Park Y. 2009b. An approach to discovering new technology opportunities: Keyword-based patent map approach. *Technovation* 29 (6-7): 481-497.
- Lievrouw LA, Rogers EM, Lowe CU, Nadel E. 1987. Triangulation as a research strategy for identifying invisible colleges among biomedical scientists. *Social Networks* 9 (3): 217-248.
- Liu X, Zhang L, Hong S. 2011. Global biodiversity research during 1900–2009: a bibliometric analysis. *Biodiv Conserv* 20 (4): 807–826.
- Perez-Iratxeta C, Bork P, Andrade MA. 2002. Association of genes to genetically inherited diseases using data mining. *Nature Genet* 31 (3): 316-319.
- Rebholz-Schuhmann D, Grabmüller C, Kavaliuskas S, Croset S, Woollard P, Backofen R et al. 2013. A case study: semantic integration of gene-disease associations for type 2 diabetes mellitus from literature and biomedical data resources. *Drug Discovery Today* 0 (0). DOI: 10.1016/j.drudis.2013.10.024
- Rebholz-Schuhmann D, Oellrich A, Hoehndorf R. 2012. Text-mining solutions for biomedical research: enabling integrative biology. *Nature Rev Genet* 13 (12): 829-839.
- Zhu F, Patumcharoenpol P, Zhang C et al. 2013. Biomedical text mining and its applications in cancer research. *J Biomed Inform* 46 (2): 200-211.