

Buku Pegangan Open Data

Release 1.0.0

Open Knowledge Foundation

Buku pegangan ini membahas aspek-aspek hukum, sosial, dan teknis dari Open Data. Buku ini bisa digunakan semua orang, meskipun dirancang khusus bagi mereka yang ingin **membuka** data. Buku ini membahas mengapa, apa, dan bagaimana cara melakukan open data – mengapa membuka data, apa pengertian membuka data, dan bagaimana cara ‘membuka’ data.

Kami menyambut hangat komentar atas teks ini dan akan memasukkan umpan balik tersebut bagi perbaikan di masa depan. Kami juga menyambut sumbangan saran mengenai hal atau bidang-bidang tambahan yang perlu dijelajahi.

BAB 1

1.1 Pengantar

Tahukah Anda berapa banyak uang dari pajak yang Anda bayarkan digunakan untuk penerangan jalan atau penelitian kanker? Rute manakah yang tersingkat, teraman, dan paling menarik bagi Anda jika ingin bersepeda dari rumah ke tempat kerja? Dan apa yang mengambang di udara dan Anda hirup sepanjang perjalanan? Di mana Anda bisa mendapatkan peluang kerja terbaik dan pohon buah terbanyak? Ketika Anda bisa memengaruhi keputusan suatu topik yang sangat Anda pedulikan, kepada siapa Anda harus bicara?

Teknologi yang berkembang sekarang membuat kita mudah menjawab pertanyaan-pertanyaan tersebut. Banyak data yang Anda butuhkan untuk menjawab pertanyaan-pertanyaan tersebut berasal dari badan-badan publik. Namun seringkali data yang tersedia tidak dalam bentuk yang siap pakai. Buku ini mencoba membuka potensi informasi resmi maupun informasi lain agar layanan (publik) bisa meningkatkan kehidupan warga dan mendorong pemerintah dan masyarakat bekerja lebih baik.

Open data dan khususnya open government data – yaitu informasi, milik publik maupun bukan, yang bebas diakses siapapun dan digunakan untuk keperluan apapun – telah ada sejak dulu. Pada tahun 2009, open data mulai dikenal setelah beberapa pemerintah (misalnya pemerintah AS, Inggris, Kanada, dan Selandia Baru) mengumumkan inisiatif baru yang mengarah pada pembukaan informasi publik. Buku ini menjelaskan konsep dasar ‘open data’, terutama dalam kaitannya dengan pemerintah. Buku ini juga mencakup bagaimana open data menciptakan nilai-nilai dan memiliki dampak positif di berbagai area. Selain mengeksplorasi latar belakang open data, buku ini juga memberi informasi kongkrit mengenai bagaimana menghasilkan open data.

1.1.1 Sasaran Pembaca

Buku pegangan ini menyoal kalangan pembaca yang:

- belum pernah mendengar tentang open data dan mereka yang menganggap dirinya ‘data professionals’ musiman
- pegawai negeri dan aktivis
- wartawan dan peneliti
- politisi dan developer/pembuat perangkat lunak
- penggiat data dan orang-orang yang belum pernah mendengar API.

Kebanyakan dari informasi yang disediakan itu berfokus pada data yang diselenggarakan sektor publik. Namun penulis berniat memperluas pembahasan sejauh waktu mengijinkan. Anda boleh berpartisipasi membantu upaya kami tersebut.

Buku pegangan ini ditujukan bagi pembaca yang hanya tahu sedikit atau bahkan sama sekali tidak mengetahui topik ini. Jika Anda menemui istilah yang terasa asing, silakan buka Glosarium dan FAQ pada bagian akhir buku ini.

1.1.2 Credits

Credits and Copyright

Contributing authors

- Daniel Dietrich
- Jonathan Gray
- Tim McNamara
- Antti Poikola
- Rufus Pollock
- Julian Tait
- Ton Zijlstra

Existing sources directly used

- Technical Proposal for how IATI is implemented. The IATI Technical Advisory Group led by Simon Parrish
- *Unlocking the Potential of Aid Information*. Rufus Pollock, Jonathan Gray, Simon Parrish, Jordan Hatcher
- Finnish manual authored by Antti Poikola
- Beyond Access Report. Access Info and the Open Knowledge Foundation

Other sources

- W3C Publishing Government Data (2009) <http://www.w3.org/TR/gov-data/>

1.2 Mengapa menerapkan Open Data?

Open data, terutama open government data, adalah sumberdaya besar yang belum banyak dilakukan. Banyak orang dan organisasi mencari dan mengumpulkan berbagai jenis data untuk mengerjakan tugas mereka. Dalam hal ini,

peran pemerintah signifikan, bukan hanya karena kuantitas dan sentralitas data yang dikumpulkannya, melainkan juga karena menurut hukum, data pemerintah pada umumnya adalah data publik, sehingga harus dibuka dan bisa digunakan oleh yang lain. Mengapa itu penting?

Pada banyak bidang, keberadaan open data akan berharga. Contohnya sudah banyak. Ada pula berbagai kelompok dan organisasi yang mendapat manfaat dari ketersediaan open data, termasuk pemerintah sendiri. Pada saat yang sama, mustahil memperkirakan dengan tepat bagaimana dan di mana nilai itu bisa tercipta. Sudah menjadi sifat inovasi bahwa pembentukannya sering terjadi di tempat-tempat yang tak terbayangkan.

Namun kita masih bisa menduga pada bidang apa open data tersebut bernilai, misalnya pada:

- Transparansi dan kendali demokrasi
- Partisipasi
- Pemberdayaan diri
- Produk atau jasa yang baru atau yang diperbaiki
- Inovasi
- Peningkatan efisiensi layanan pemerintah
- Peningkatan efektivitas layanan pemerintah
- Pengukuran dampak kebijakan
- Pengetahuan baru yang merupakan perpaduan berbagai sumber dan pola data dalam volume besar

Ada contoh manfaat open data di tiap bidang tersebut.

Dalam konteks transparansi, proyek-proyek seperti “pohon pajak” di Finlandia dan “ke mana perginya uang kami” di Inggris, menunjukkan bagaimana uang pajak rakyat dibelanjakan oleh negara. Selain itu juga ada contoh bagaimana open data menyelamatkan Kanada dari penipuan pajak sumbangan sejumlah \$ 3,2 milyar. Contoh lain adalah keberadaan situs web seperti folketsting.dk di Denmark yang menelusuri kegiatan di parlemen dan proses pembuatan undang-undang, sehingga publik bisa melihat apa yang terjadi di parlemen dan siapa-siapa saja yang terlibat.

Open government data juga bisa membantu Anda membuat keputusan yang lebih baik dalam hidup Anda, atau membuat Anda lebih aktif di masyarakat. Seorang wanita di Denmark membuat situs findtoilet.dk yang menunjukkan lokasi semua toilet publik di Denmark, sehingga orang-orang yang memiliki masalah kandung kemih bisa memberanikan diri untuk pergi keluar lagi. Di Belanda, vervuilingsalarm.nl bisa memberi pesan pada publik jika kualitas udara di lingkungannya diperkirakan akan mencapai titik ambang tertentu pada esok hari. Di New York, masyarakat bisa mudah mengetahui tempat-tempat mana yang bisa mereka pakai untuk membawa anjingnya jalan-jalan, atau untuk mencaritahu siapa yang menggunakan taman yang sama. Layanan seperti ‘mapumental’ di Inggris dan ‘mapnificent’ di Jerman memungkinkan Anda mencari tempat untuk tinggal, dengan mempertimbangkan lama waktu yang dibutuhkan untuk pulang-balik ke tempat kerja, harga pemukiman, dan pemandangan di suatu daerah. Contoh-contoh di atas tersebut menggunakan open government data.

Secara ekonomi, open data juga bermakna penting. Beberapa studi telah memperkirakan nilai ekonomis open data, yaitu belasan milyar Euro per tahun di Uni Eropa saja. Produk dan perusahaan-perusahaan baru menggunakan

kembali (re-use) open data. Sebuah situs web Denmark, husetsweb.dk membantu masyarakat mencari cara memperbaiki efisiensi energi rumah mereka, termasuk dalam membuat perencanaan keuangan dan mencari tukang untuk merenovasi rumah. Informasi tersebut berdasarkan penggunaan kembali informasi dari survey dan dari subsidi pemerintah serta pendaftaran usaha dagang. Google Translate menggunakan banyak dokumen Uni Eropa yang muncul dalam bahasa-bahasa Eropa untuk melatih algoritma penerjemahan, sehingga memperbaiki kualitas layanannya.

Open data juga menambah nilai bagi pemerintah sendiri, misalnya, meningkatkan efisiensi pemerintah. Kementerian Pendidikan Belanda mempublikasikan semua data terkait pendidikan secara online untuk digunakan kembali. Sejak itu, jumlah pertanyaan yang diajukan menurun, demikian pula dengan beban kerja dan biaya. Hal itu juga memudahkan pegawai negeri terkait menjawab pertanyaan karena mereka bisa mengetahui jelas di mana menemukan data terkait. Open data juga membuat pemerintah lebih efektif, sehingga pada akhirnya juga mengurangi biaya. Lembaga pemerintah Belanda yang mengurus warisan budaya aktif merilis data dan berkolaborasi dengan masyarakat pemerhati sejarah dan kelompok-kelompok seperti Wikimedia Foundation untuk menjalankan tugas mereka secara lebih efektif. Hasilnya bukan saja peningkatan kualitas data, namun juga lembaga yang lebih ramping.

Meskipun ada banyak contoh bagaimana open data memberi nilai sosial dan ekonomis, kita belum tahu apa saja yang akan bisa diciptakan dari open data. Perpaduan baru data-data dapat menciptakan pengetahuan dan wawasan baru yang membawa pada bidang penerapan yang baru pula. Kita bisa melihat contoh di masa lalu, misalnya ketika Dr. Snow menemukan kaitan antara polusi air minum dan penyakit kolera di London pada abad 19. Beliau memadukan data mengenai kematian akibat kolera dengan lokasi sumur air. Temuan tersebut berakibat pada dibangunnya sistem pembuangan limbah di London yang hasilnya sangat memperbaiki kesehatan masyarakat. Kita bisa melihat perkembangan semacam itu terjadi lagi dengan terjadinya aliran wawasan baru dari perpaduan berbagai rangkaian open data.

Potensi yang belum tergali tersebut bisa muncul jika kita menjadikan data pemerintah, yang merupakan data publik, menjadi open data. Hal ini hanya bisa terjadi jika tidak ada batasan (legal, financial, maupun teknologi) pada data yang telah dibuka tersebut untuk digunakan lagi oleh pihak lain. Setiap pembatasan atas penggunaan kembali data publik akan mempersulit kemungkinan untuk mewujudkan manfaat open data. Agar potensinya disadari dan tergali, data publik harus dijadikan open data.

1.3 Apa yang dimaksud dengan Open Data?

Buku pegangan ini membahas tentang open data. Namun, apa sebenarnya open data itu? Khususnya, apa yang membuat open data terbuka, dan data-data apa saja yang dimaksud?

1.3.1 Apa yang dimaksud terbuka?

Buku ini membahas open data –namun apakah sebenarnya yang disebut open data? Untuk tujuan itu, open data didefinisikan oleh pengertian berikut:

Definisi Open (terbuka):

Open data adalah data yang bisa bebas digunakan, digunakan kembali, dan disebarluaskan kembali oleh semua orang – hanya tergantung pada persyaratan sifat dan keterbagian yang dimiliki.

Definisi **Full Open (terbuka penuh)** memberi gambaran lebih rinci mengenai maksudnya. Jika dirangkum, pengertian terpenting adalah:

- **Ketersediaan dan Akses:** data harus tersedia utuh dan tidak memerlukan biaya reproduksi yang berlebihan, lebih disarankan jika data bisa diunduh dari internet. Data juga harus tersedia dalam bentuk yang mudah digunakan (*convenient*) dan dapat diubah (*modifiable*).
- **Penggunaan kembali dan penyebarluasan kembali data** harus dilakukan melalui syarat-syarat yang berlaku bagi penggunaan-kembali dan penyebarluasan-kembali, termasuk pencampuran dengan set data lain.
- **Partisipasi Universal:** setiap orang bisa menggunakan, menggunakan kembali, dan menyebarkan kembali – tidak boleh ada diskriminasi atas bidang usaha, orang, atau kelompok. Misalnya, batasan ‘non-komersial’ yang melarang penggunaan ‘komersial’, atau batasan penggunaan untuk tujuan tertentu (mis, hanya untuk pendidikan), tidak dibolehkan.

Jika Anda bertanya-tanya mengapa harus benar-benar jelas arti *open* dan mengapa definisi ini yang digunakan, jawabannya sederhana: *interoperability*.

Interoperability menunjukkan kemampuan berbagai sistem dan organisasi yang berbeda saling bekerjasama (*inter-operate*). Dalam hal ini, yang dimaksud adalah kemampuan bekerja bersama –atau berpadu- antara set data yang berbeda.

Interoperability penting karena memungkinkan berbagai komponen berbeda untuk bekerja bersama. Kemampuan untuk membentuk dan ‘menghubungkan bersama’ antar komponen adalah penting untuk membangun sistem yang besar dan kompleks. Tanpa *interoperability*, hal itu akan jadi nyaris mustahil. Ini terbukti pada mitos terkenal Menara Babel, di mana ke(tidak)mampuan berkomunikasi dan bekerjasama berujung pada gagalnya upaya pembangunan menara.

Kita menghadapi masalah serupa dengan data. Inti dari “perkampungan” data (atau kode) adalah bahwa sepotong bahan “terbuka” yang terkandung di dalamnya dapat saling bercampur bebas dengan bahan “terbuka”.

Interoperability ini adalah kunci mutlak untuk mewujudkan manfaat praktis dari “keterbukaan”: meningkatkan kemampuan untuk memadukan bermacam set data sehingga bisa mengembangkan produk dan jasa yang lebih baik (manfaat-manfaat tersebut akan dibahas lebih rinci pada bagian “mengapa” open data). Dengan memberi definisi jelas atas keterbukaan, ketika Anda menghadapi dua set data dari dua sumber, Anda akan mampu memadukannya. Dan itu memastikan bahwa Anda terhindar dari mengulang kegagalan “menara babel”: banyak set data namun hampir atau sama sekali tidak ada kemampuan untuk memadukannya menjadi sistem yang lebih besar tempat munculnya nilai-nilai nyata.

1.3.2 Data apa yang dimaksud?

Pembaca telah melihat contoh-contoh jenis data yang sudah atau bisa terbuka – dan masih akan ada contoh-contoh lain di bawah. Namun akan berguna jika kita jabarkan jenis-jenis data yang terbuka, bisa terbuka, dan –juga sama pentingnya- tidak bisa terbuka.

Yang terpenting ketika membuka data adalah berfokus pada data non-personal, yaitu data yang tidak mengandung informasi tentang orang tertentu.

Larangan yang sama berlaku pula bagi beberapa jenis data pemerintah dan keamanan nasional.

1.4 Bagaimana cara membuka data

Bagian ini adalah inti dari buku pegangan ini, karena mengetengahkan saran kongkrit dan rinci mengenai bagaimana pemilik data dapat membuka data. Kita akan menyajikan dasar-dasar sekaligus cara mengatasi jebakannya. Yang terakhir, kita akan membahas isu-isu lebih subtil yang mungkin timbul.

Ada tiga aturan kunci yang kami sarankan ketika Anda membuka data Anda:

- **Buatlah sederhana.** Mulailah dengan yang kecil, sederhana, dan cepat. Tidak semua set data harus dibuka segera secara bersamaan. Lakukan secara bertahap, dimulai dari satu set data, atau bahkan satu bagian dari set data yang besar. Walau tentu saja, semakin banyak set data yang Anda buka, akan semakin baik. Harap diingat bahwa ini menyangkut inovasi. Bergerak secepat yang Anda bisa itu baik karena memudahkan Anda membangun momentum dan belajar dari pengalaman. Selain itu, inovasi juga mengandung kemungkinan gagal maupun berhasil, serta tidak setiap set data bisa berguna.
- **Pelibatan sejak awal dan sering.** Libatkan calon pengguna dan calon pengguna-kembali (re-user) data tersebut seawal dan sesering mungkin. Calon pengguna tersebut bisa warga, usahawan, atau developer. Langkah ini akan memastikan bahwa ketika mereka menggunakan layanan Anda lagi, akan lebih relevan. Perlu diingat bahwa banyak data tidak akan langsung menjangkau pengguna, namun melalui ‘perantara informasi’ (*‘info-mediaries’*). Perantara informasi adalah orang-orang yang mengambil data dan mengubah-olah atau me-remix data tersebut sebelum disajikan. Misalnya, banyak pembaca yang tidak mau membaca data yang berisi banyak koordinat GPS, namun lebih suka membacanya dalam bentuk peta. Oleh karena itu, jalin komunikasi dengan perantara informasi lebih dulu. Mereka akan menggunakan-kembali dan mengubah bahan itu ke dalam bentuk yang sesuai kebutuhan pengguna akhir.
- Atasi masalah kekhawatiran dan kesalahpahaman yang lazim terjadi. Ini penting, terutama jika Anda bekerja dengan atau dalam lembaga besar seperti pemerintah. Ketika membuka data, Anda akan menemui banyak pertanyaan dan kekhawatiran. Penting bagi Anda untuk (a) mengidentifikasi pertanyaan dan masalah-masalah terpenting, dan (b) mengatasinya pada tahap seawal mungkin.

Ada empat langkah untuk membuat data terbuka. Masing-masing akan dijelaskan rinci di bawah. Langkah-langkah berikut ini diurutkan menurut yang paling baik. Beberapa langkah tersebut bisa dilakukan bersamaan.

1. Memilih set data. Pilih set data yang akan dibuka. Ingat, Anda bisa (dan mungkin perlu) kembali ke langkah ini jika menemui masalah pada tahap selanjutnya.
2. Terapkan ijin terbuka (open license).
 - (a) Tentukan Hak Kekayaan Intelektual yang ada pada data tersebut.

- (b) Terapkan jenis perijinan ‘terbuka’ yang tepat yang mencakup ijin atas hak-hak dan memenuhi definisi keterbukaan seperti yang dibahas di bagian “Apa yang dimaksud Open Data” di atas
 - (c) NB: Jika Anda tidak bisa melakukan ini, kembalilah ke langkah 1 dan coba set data yang lain.
3. Sediakan data tersebut bagi publik dalam bentuk lengkap (*bulk*) dan format yang berguna. Anda juga bisa menggunakan cara alternatif, misalnya melalui API.
 4. Pastikan bahwa data bisa/mudah ditemukan – posting di web dan buat katalog untuk memudahkan pencarian data tersebut.

1.4.1 Memilih set data

Memilih set data yang akan anda buka adalah langkah pertama. Namun tetap perlu diingat bahwa proses pembukaan data mengandung pengulangan, sehingga Anda bisa kembali ke langkah ini lagi ketika menemui masalah pada tahap-tahap berikutnya. Jika Anda sudah tahu pasti set data yang akan Anda buka, Anda bisa langsung melangkah ke bagian selanjutnya.

Dalam banyak kasus, terutama di lembaga-lembaga besar, memilih set data untuk difokuskan adalah masalah sekaligus tantangan besar. Bagaimana cara mengatasinya?

Cara yang cepat adalah membuat daftar dari berbagai set data yang bisa dijadikan awal. Pada tahap-tahap selanjutnya, Anda masih bisa kembali ke daftar ini untuk melihat set data mana yang cocok.

Tidak ada ketentuan khusus tentang bagaimana membuat daftar yang komprehensif. Ketika akan mempublikasikan suatu set data, yang perlu diingat adalah apakah langkah tersebut mungkin/*feasible* untuk dilakukan (apakah akan dibuka atau tidak). Lihat kembali bagian sebelumnya.

Bertanya pada masyarakat

Kami sarankan Anda bertanya pada masyarakat sejak awal. Masyarakat yang dimaksud di sini adalah orang-orang yang akan mengakses dan menggunakan data, karena merekalah yang akan mengerti data mana yang berharga.

1. Buat daftar singkat sejumlah set data yang akan Anda mintai umpan baliknya dari masyarakat. Daftar tersebut tidak harus sama dengan harapan Anda, yang penting adalah mendapat gambaran atas permintaannya. Daftar bisa saja didasarkan pada katalog open data dari negara lain.
2. Buat permohonan komentar.
3. Publikasikan permohonan Anda melalui webpage. Pastikan bahwa permohonan tersebut bisa diakses publik melalui URL-nya, sehingga ketika dibagikan di social media, permohonan itu bisa mudah ditemukan.
4. Mudahkan cara untuk mengirim tanggapan. Hindari keharusan registrasi karena itu akan menyurutkan minat orang untuk memberi tanggapan.
5. Edarkan permohonan itu melalui milis, forum, dan orang-orang terkait. Jangan lupa memberi tautan ke webpage yang mengandung permohonan tersebut.
6. Buat dan jalankan acara konsultasi. Pastikan Anda menjalankannya pada saat yang nyaman bagi pekerja bisnis, pencari data, dan pejabat terkait, agar mereka bisa mengikutinya.

7. Minta politisi untuk berbicara atas nama organisasi Anda. Open Data adalah bagian dari meningkatkan akses ke informasi pemerintah.

Pertimbangan biaya

Berapa banyak uang yang dibelanjakan lembaga untuk mengumpulkan dan memelihara data mereka? Jika mereka membelanjakan banyak uang untuk suatu set data, maka sangat mungkin pihak-pihak lain tertarik mengaksesnya. Hal ini rentan terhadap masalah pemboncengan. Pertanyaan yang perlu Anda jawab adalah, "Mengapa orang lain bisa mendapatkan informasi mahal secara cuma-cuma?". Jawabannya adalah bahwa pembelanjaan itu diserap oleh sektor publik untuk melakukan fungsi tersebut. Setelah data terkumpul, biaya pengirimannya ke pihak ketiga nyaris tidak ada. Oleh karena itu, mereka juga tidak perlu membayar.

Kemudahan merilis data

Terkadang akan lebih berguna jika kita melihat data mana yang lebih mudah sampai ke tangan publik, alih-alih menentukan data mana yang paling berharga. Rilis yang kecil dan mudah bisa menjadi pemicu bagi perubahan perilaku yang lebih besar dalam organisasi.

Namun Anda perlu berhati-hati dengan pendekatan ini. Mungkin saja rilis yang kecil itu sangat kecil nilainya hingga tidak ada yang bisa dibangun karenanya. Jika ini terjadi, kepercayaan pada seluruh proyek akan terkikis.

Amati kawan

Open data adalah gerakan yang tumbuh. Ada kemungkinan banyak orang di daerah Anda yang memahami apa yang terjadi di daerah lain. Buatlah daftar berisi apa saja yang dilakukan lembaga-lembaga tersebut.

1.4.2 Menerapkan Lisensi Terbuka (Keterbukaan Legal/Legal Openness)

Banyak daerah yang menerapkan HAKI pada data, sehingga ada larangan bagi pihak ketiga untuk menggunakan dan menggunakan-kembali data tanpa ijin jelas/tertulis. Bahkan di daerah-daerah yang belum tentu memiliki hak tersebut, penerapan ijin masih penting untuk kejelasan. Jadi jika Anda berencana membuka data, Anda sebaiknya juga menerapkan ijin pada data tersebut. Ijin apa yang bisa Anda pakai? Kami menyarankan bahwa untuk data 'terbuka', Anda dapat menggunakan salah satu dari ijin yang sesuai dengan Definisi Terbuka dan ditandai sebagai cocok untuk data. Pada tautan berikut ada daftar ijin (beserta instruksi penggunaannya) tersebut:

- <http://opendefinition.org/licenses/>

Instruksi sepanjang satu halaman untuk menerapkan ijin pada open data dapat Anda lihat pada situs Open Data Commons:

- <http://opendatacommons.org/guide/>

1.4.3 Mengupayakan Keberadaan Data (Technical Openness)

Menerapkan **Open Data** perlu pertimbangan teknis dan legal. Secara teknis, data yang dibuka tersebut perlu disediakan dalam bentuk yang lengkap dan format yang dapat dibaca mesin.

Tersedia. Harga data yang dibuka tersebut tidak boleh lebih mahal daripada biaya reproduksi. Sebaiknya data bisa diunduh gratis dari internet. Penentuan harga semacam ini terjadi karena lembaga Anda tidak perlu mengeluarkan biaya ketika menyediakan data untuk dibuka dan digunakan.

Lengkap. Data harus disediakan dalam bentuk lengkap. Jika Anda memiliki register yang dikumpulkan menurut peraturan internal, seluruh register harus bisa diunduh. Suatu web API atau layanan sejenis bisa digunakan, namun tidak bisa menggantikan akses ke data lengkap (*bulk access*).

Dalam format terbuka dan dapat dibaca mesin. Ketika membuka data publik, seyogianya data tersebut tidak dibatasi oleh hak paten. Terlebih penting, pastikan bahwa Anda menyediakan data dalam format yang bisa dibaca mesin (*machine-readable format*) agar data mudah dipakai kembali. Sebagai gambaran, data statistik yang dipublikasikan dalam bentuk PDF (Portable Document Format) memang memiliki tampilan yang bagus dan dapat dibaca manusia. Namun komputer akan sulit menggunakan data dengan format tersebut, sehingga juga menyulitkan pihak lain untuk menggunakannya kembali.

Ada beberapa kebijakan yang akan sangat bermanfaat:

- Buatlah sederhana,
- Bergerak cepat
- Pragmatis.

Penyediaan data sebaiknya disegerakan. Lebih baik memberi data mentah sekarang daripada menyediakan data yang sempurna namun dalam enam bulan mendatang.

Ada beberapa cara penyediaan data. Pada era internet ini, cara yang umum tentunya adalah publikasi daring (online). Ada banyak variasi dari model publikasi ini. Pada dasarnya, lembaga menempatkan datanya di situs web serta menempatkan katalog pusat yang mengarahkan pengunjung ke sumber yang tepat. Ada beberapa alternatif dari cara ini. Jika konektivitas terbatas atau ukuran data sangat besar, distribusi melalui format-format lain bisa terjamin. Bagian ini juga akan membahas alternatif-alternatif yang bisa menekan harga.

Metode Online

Lewat situs web Anda

Sistem yang paling mudah bagi tim web content Anda adalah menyediakan file-file untuk diunduh dari halaman web. Karena Anda menyediakan akses ke dokumen diskusi, penyediaan file-file data melalui cara ini akan lebih mudah. Kesulitan dari pendekatan ini adalah bagi orang luar yang ingin mencari informasi yang telah diperbarui. Pilihan ini juga menjadi beban bagi orang yang hendak membuat alat/*tool* dengan data Anda.

Melalui situs pihak ketiga

Banyak penyimpanan yang menjadi simpul penghubung (*hub*) data pada bidang tertentu. Misalnya, pachube.com dirancang untuk menghubungkan orang-orang dengan sensor dengan mereka yang ingin mengakses data darinya. Situs seperti Infochimps.com dan Talis.com menjadi tempat bagi badan-badan sektor publik menyimpan datanya secara cuma-cuma.

Situs pihak ketiga bisa sangat berguna karena mampu mengumpulkan pihak-pihak yang berkepentingan dan data-data terkait. Jika data anda menjadi bagian dari platform ini, ia menjadi bagian perpaduan kepentingan yang positif. Platform data besar menyediakan infrastruktur yang dapat menyokong permintaan tersebut. Platform seperti itu memberi informasi analitis dan penggunaan. Bagi badan sektor publik, platform semacam itu biasanya gratis. Hanya ada dua biaya yang ditimbulkan. Pertama adalah kebebasan. Lembaga Anda harus bisa mendapat kendali atas yang lain. Inilah yang seringkali sulit secara politis, legal, dan operasional. Biaya kedua adalah keterbukaan. Pastikan bahwa platform data Anda skeptis terhadap siapa yang bisa mengaksesnya. Pembuat perangkat lunak dan ilmuwan bisa menggunakan berbagai sistem operasi, dari telepon pintar hingga komputer super. Semua cara itu harus bisa dipakai untuk mengakses data.

Melalui FTP servers

Cara yang kurang menarik dalam penyediaan akses ke file adalah melalui File Transfer Protocol (FTP). Cara ini mungkin cocok jika pengguna data Anda memiliki teknis khusus, misalnya para pembuat perangkat lunak atau ilmuwan. Sistem FTP bekerja dalam ranah HTTP, namun dirancang khusus untuk menyokong transfer file. FTP sudah tidak populer. Alih-alih menyediakan website, melihat melalui server FTP lebih mirip seperti mencari-cari di folder-folder komputer. Sehingga meskipun cocok dengan tujuan open data, kapasitas bagi perusahaan pembuat perangkat lunak untuk melakukan penyesuaian (*customization*) tidak terlalu dibutuhkan.

Dirilis sebagai torrents

BitTorrent adalah sistem yang mulai akrab bagi pembuat kebijakan karena asosiasinya dengan pelanggaran hak cipta. BitTorrent menggunakan file-file yang disebut *torrents*, yang bekerja dengan cara membagi biaya pendistribusian file ke orang-orang yang mengakses file-file tersebut. File-file tidak akan mengalami beban berlebihan karena penyediaan/penawaran meningkat bersamaan dengan permintaan yang meningkat. Inilah sebab mengapa sistem ini sangat berhasil dalam hal berbagi file-file film. Cara ini memang sangat efisien untuk mendistribusikan data dalam volume besar.

Sebagai API

Data bisa dipublikasikan melalui [Application Programming Interface](#) (API). Antar-muka ini sudah menjadi populer, karena memungkinkan programmer memilih porsi data tertentu saja, sehingga tidak harus menyediakan semua data dalam bentuk lengkap dan file besar. API terhubung dengan database yang diperbarui secara *real-time*. Ini berarti penyediaan informasi melalui API bisa memastikan bahwa data selalu up to date.

Sebaiknya semua inisiatif open data memprioritaskan publikasi data mentah dalam bentuk lengkap (bulk).

Penyediaan API akan membutuhkan biaya sebagai berikut:

1. Harga. API membutuhkan lebih banyak upaya pengembangan dan pemeliharaan daripada sekedar menyediakan file-file.
2. Ekspektasi. Untuk mendukung komunitas pengguna di balik sistem, memberi kepastian adalah penting. Jika ada yang berjalan tidak semestinya, Anda terpaksa mengeluarkan biaya untuk memperbaikinya.

Akses ke data bentuk lengkap (*bulk data*) memastikan bahwa:

1. tidak ada ketergantungan pada penyedia asli data, sehingga jika ada restrukturisasi atau siklus anggaran yang memengaruhi situasi, data masih tetap tersedia, tidak terpengaruh.
2. siapa saja bisa mendapatkan salinan dan mendistribusikannya kembali. Hal ini bisa mengurangi biaya distribusi dari lembaga sumber, sehingga tidak ada satu pun titik kegagalan.
3. pihak lain bisa membangun jasa mereka menggunakan data tersebut karena ada kepastian bahwa data tidak akan diambil/dicuri dari mereka.

Menyediakan data dalam bentuk lengkap memungkinkan pihak lain menggunakan data di luar tujuan aslinya. Misalnya, data bisa diubah ke format baru, dihubungkan dengan sumberdaya lain, atau dijadikan suatu versi dan disimpan di beberapa tempat. Meskipun versi terbaru data tersebut tersedia melalui API, data mentah sebaiknya tetap disediakan dalam bentuk lengkap secara teratur.

Misalnya, kantor statistik [Eurostat](#) memiliki fasilitas unduh data lengkap/*bulk* yang menawarkan lebih dari 4000 file data. Dilakukan update dua kali sehari, dan ditawarkan data dalam bentuk [Tab-separated values \(TSV\)](#), serta dokumentasi mengenai fasilitas unduh dan file data.

Contoh lain adalah Katalog Data [District of Columbia](#), yang memungkinkan data diunduh dalam format CSV dan XLS selain fasilitas *live feeds*.

1.4.4 Memastikan data bisa ditemukan pihak luar

[Open data](#) tidak akan berarti tanpa pengguna. Anda harus memastikan bahwa orang-orang bisa mencari bahan sumber. Bagian ini akan membahas berbagai pendekatan untuk itu.

Yang terpenting adalah menyediakan ruang netral yang dapat mengantisipasi politik antar-lembaga dan siklus anggaran di masa mendatang. Batas yurisdiksi, baik sektoral maupun geografis, bisa menyulitkan kerjasama. Padahal sebenarnya ada manfaat signifikan dari kerjasama, yaitu agar pihak luar bisa menemukan data dengan mudah, sehingga alat-alat baru yang berguna bisa dibuat.

Alat-alat yang ada

Ada sejumlah alat yang ada di web dan dirancang untuk membuat data mudah ditemukan. Salah satu yang menonjol adalah [DataHub](#) dan merupakan katalog serta penyimpanan data bagi berbagai set data dari seluruh dunia. Situs ini memudahkan individu dan organisasi dalam mempublikasikan bahan dan bagi pencari data yang ingin memperoleh bahan yang mereka perlukan.

Selain itu masih ada puluhan katalog khusus untuk berbagai sektor dan tempat. Banyak komunitas ilmiah yang telah menciptakan sistem katalog bagi bidangnya, karena data-data tersebut seringkali dibutuhkan untuk publikasi.

Bagi pemerintah

Umumnya, lembaga pengarah (lead agency) membuat katalog bagi data pemerintah. Ketika membuat katalog, cobalah membuat struktur yang memudahkan departemen dalam memperbarui data-datanya.

Tahan desakan untuk membangun perangkat lunak untuk membuat katalog dari tahap yang sangat awal. Ada solusi perangkat lunak yang gratis dan open source (seperti CKAN) yang juga telah diadopsi banyak pemerintah. Dengan demikian, mungkin tidak perlu lagi mengeluarkan uang pada platform lain.

Ada beberapa hal yang sering terlewatkan pada katalog open data. Program Anda bisa mempertimbangkan hal-hal berikut:

- memberi jalan bagi sektor swasta dan masyarakat untuk menambahkan data mereka. Akan lebih berguna jika Anda memanfaatkan katalog sebagai katalog wilayah, alih-alih katalog milik pemerintah daerah.
- memfasilitasi peningkatan data dengan memungkinkan turunan set data dimasukkan katalog. Misalnya, ada yang ingin membuat kode geografis (*geocode*) dari alamatnya dan ingin membagi hasil tersebut dengan semua orang. Jika Anda hanya membolehkan satu versi set data, peningkatan ini tidak akan terlihat.
- bersikap toleran jika melihat data Anda tampil di tempat lain. Itu artinya konten tersebut bisa disalin oleh masyarakat yang berminat. Jika Anda mempublikasikan data pemantauan ketinggian air sungai misalnya, data Anda mungkin saja bisa tampil pada katalog ahli hidrologi.
- pastikan bahwa akses setara. Jangan membuat tingkatan akses untuk pejabat atau peneliti tetap karena akan menyurutkan minat masyarakat untuk berpartisipasi.

Bagi masyarakat sipil

Bersedia membuat katalog tambahan untuk data tidak resmi.

Pemerintah jarang berhubungan dengan sumber tidak resmi atau pihak yang tidak berwenang. Pejabat (pemerintah) sering mengeluarkan banyak biaya untuk memastikan bahwa tidak akan terjadi hal-hal yang secara politis memalukan atau kerugian lain karena mengandalkan data yang salah.

Terlebih, pemerintah tidak mau mendukung kegiatan yang menautkan informasi mereka dengan informasi dari sektor bisnis. Pemerintah memang berhak bersikap skeptis terhadap motivasi laba. Oleh karena itu, perlu katalog independen/terpisah untuk kelompok komunitas, bisnis, dan lain-lain.

1.5 Setelah membuka data, lalu apa?

Kita telah membahas cara agar informasi pemerintah terbuka dengan legal dan secara teknis dapat digunakan kembali. Langkah selanjutnya adalah mendorong pihak-pihak lain memanfaatkan data tersebut.

Bagian berikut ini membahas hal-hal tambahan yang dapat dilakukan untuk mendorong penggunaan-kembali data.

1.5.1 Umumkan pada dunia!

Pertama dan terutama, pastikan bahwa Anda mengumumkan kampanye untuk mempromosikan open data di area tanggung jawab Anda.

Jika Anda membuka sejumlah set data, sebaiknya Anda meluangkan waktu juga untuk memastikan bahwa orang-orang lain tahu (atau setidaknya bisa mencari tahu) bahwa Anda membuka data. Selain *press release*, pengumuman di situs, dan sebagainya, Anda bisa:

- menghubungi organisasi atau orang-orang menonjol yang bekerja atau tertarik pada area kerja tersebut

- menghubungi milis atau kelompok jaringan sosial terkait
- langsung menghubungi calon pengguna yang setahu Anda mungkin tertarik pada data tersebut.

Kenali audiens

Seperti halnya komunikasi publik pada umumnya, berhubungan dengan komunitas data juga perlu penetapan sasaran. Seperti halnya semua kelompok pemangku kepentingan, pesan yang benar bisa saja sia-sia jika diarahkan ke area yang salah.

Komunitas digital cenderung sangat bersedia berbagi informasi baru, bahkan sangat cepat mengkonsumsinya.

Menulislah seolah pesan Anda hanya akan dibaca sekilas, alih-alih diteliti mendalam. Anggota komunitas teknologi tidak seperti publik umum yang menggunakan MS Windows. Artinya, Anda sebaiknya tidak menyimpan dokumen dalam format MS Office yang dapat dibaca offline, karena 2 sebab berikut:

- Pertama, dokumen-dokumen tersebut tidak akan mudah diakses. Alih-alih terbaca di layar, dokumen Anda akan terbaca sebagai salinan yang buruk.
- Kedua, lembaga Anda seolah menyampaikan pesan tersirat bahwa Anda tidak bersedia mendekati developer (pengembang perangkat lunak).

Sebaliknya, Anda menunjukkan kesan bahwa Anda mengharapkan komunitas teknologi yang mendatangi Anda.

Publikasikan bahan Anda pada situs pihak ketiga

Banyak blog yang membahas topik tertentu dan mengundang banyak pembaca yang berminat pada area tersebut. Anda pun bisa menambahkan artikel atau tulisan mengenai inisiatif Anda pada situs mereka. Kerjasama ini saling menguntungkan. Anda akan mengundang lebih banyak pihak yang berminat, sedangkan pemilik blog menerima sumbangan post di blog mereka.

Menjadikan komunikasi Anda lebih ramah social-media

Mengharapkan pejabat meluangkan banyak waktu untuk berinteraksi di social media tidaklah realistis. Namun ada beberapa hal yang bisa Anda lakukan agar konten bisa mudah dibagi antar pengguna teknis. Beberapa tips yang bisa Anda coba:

Sediakan halaman-halaman unik untuk setiap konten. Ketika pesan dibagi dengan pihak lain, penerima (*referral*) akan segera mencari konten terkait.

Jangan membuat orang mengunduh *press release* Anda. *Press release* itu bagus, karena mengandung pesan singkat tentang suatu hal tertentu. Namun jika Anda meminta pengunjung mengunduh konten itu dan membukanya di luar *web browser*, maka hanya sedikit orang yang akan membacanya. Mesin pencari kurang bisa mengindex konten. Pengunjung kurang suka mengklik untuk mengunduh.

Pertimbangkan untuk menggunakan **Open license untuk konten Anda.** Selain member kepastian pada orang yang ingin berbagi konten, Anda menyiratkan pesan bahwa lembaga Anda memahami prinsip keterbukaan. Langkah ini akan jauh lebih mengesankan bagi pendukung open data daripada segala kalimat pada *press release*.

Social media

Bagi lembaga yang keuangannya terbatas, tidaklah efisien menghabiskan berjam-jam di situs social media. Cara paling efisien untuk membuat suara Anda terdengar adalah memanfaatkan forum-forum tersebut untuk menyampaikan pesan bahwa post-post Anda di blog mudah dibagi. Artinya, sebelum melangkah ke pembahasan selanjutnya, pastikan Anda telah membaca yang terakhir. Dengan mengingat itu, ada beberapa saran:

Forum-forum diskusi Twitter menjadi pilihan untuk menyebarluaskan informasi dengan cepat. Segala yang bertagar #opendata akan segera dilihat oleh ribuan pengguna twitter. LinkedIn memiliki seleksi kelompok besar yang menasar khusus open data.

Walau Facebook sangat bagus bagi sasaran pembaca umum, namun komunitas open data tidak menaruh perhatian besar pada media ini.

Link aggregators. Kirimkan konten Anda ke *newswires* untuk panggilan komputer (*geeks*). Saat ini, Reddit dan Hacker News adalah dua newswires terbesar pada bidang ini.

Dalam lingkup lebih kecil, Slashdot dan Digg juga merupakan alat berguna pada bidang ini.

Situs-situs tersebut memiliki kecenderungan untuk menyetir lalu lintas pengguna ke bahan-bahan yang menarik, selain sangat berfokus pada area topik tersebut.

1.5.2 Mengumpulkan kawan: acara tatap-muka

Acara tatap-muka bisa menjadi cara yang sangat efektif untuk mendorong pihak lain menggunakan data Anda.

Alasan untuk menyelenggarakan acara tatap-muka di antaranya adalah:

- mengetahui lebih banyak calon pengguna-kembali data Anda
- mengetahui permintaan berbagai jenis set data
- mengetahui bagaimana orang ingin menggunakan-kembali data Anda
- memungkinkan calon pengguna-kembali mengetahui lebih banyak tentang data yang Anda miliki
- memungkinkan para calon pengguna bertemu sehingga bisa membuka kemungkinan-kemungkinan lain, seperti kolaborasi
- memaparkan data Anda ke cakupan publik yang lebih luas, (misalnya dari post di blog ke liputan media yang mungkin terjadi karena adanya acara tatap-muka tersebut)

Ada banyak cara untuk mengadakan acara dan banyak pula jenis acara tatap-muka, tergantung pada tujuan yang ingin Anda raih. Seperti halnya model konferensi tradisional yang melibatkan perbincangan formal yang telah disiapkan, presentasi, dan demonstrasi, ada berbagai acara yang dimotori peserta, di mana hadirin bisa:

- Memandu atau menentukan agenda acara
- Mengenalkan diri, bicara mengenai apa yang diminati dan sedang dikerjakan, secara sementara
- Memberi presentasi pendek secara spontan mengenai apa yang sedang mereka kerjakan
- Sesi-sesi penuntun mengenai sesuatu yang mereka minati

Ada banyak dokumentasi online mengenai cara menyelenggarakan acara-acara tersebut. Anda dapat mencarinya dengan kata kunci seperti: 'unconference', 'barcamp', 'meetup', 'speedgeek', 'lightning talk', dan sebagainya. Anda juga bisa menghubungi orang yang biasa menyelenggarakan acara semacam itu untuk masukan dan bantuan untuk

menjalankan acara Anda. Bisa juga bermitra dengan organisasi lain (misalnya organisasi masyarakat warga, organisasi berita, atau organisasi pendidikan) untuk memperluas basis peserta dan cakupan Anda.

1.5.3 Buat acara khusus: lomba, hadiah, purwarupa, *hackdays*

Pada lomba-lomba ini, sejumlah set data dirilis dan programmer diberi jangka waktu pendek (sekitar 48 jam hingga beberapa minggu) untuk membuat aplikasi menggunakan data tersebut. Hadiah akan diberikan pada pemenang lomba yang membuat aplikasi terbaik. Lomba-lomba semacam ini telah diadakan di banyak negara, termasuk di Inggris, Amerika Serikat, Norwegia, Australia, Spanyol, Denmark dan Finlandia.

Contoh-contoh perlombaan

Kompetisi pertama di dunia adalah **Show us a better way** yang diadakan oleh gugus kerja pemerintah Inggris, yaitu “The Power of Information Taskforce” yang dikepalai oleh Cabinet Office Minister Tom Watson pada Maret 2008. Lomba ini mengajukan pertanyaan “Apa yang bisa Anda buat dari informasi publik?” dan terbuka untuk dijawab oleh programmer di seluruh dunia, dengan imbalan hadiah £80,000 bagi lima aplikasi terbaik.

Apps for Democracy, salah satu lomba pertama di Amerika Serikat, diluncurkan pada Oktober 2008 oleh Vivek Kundra, yang pada saat itu menjabat Chief Technology Officer (CTO) dari pemerintah District of Columbia (DC). Kundra mengembangkan katalog DC data yang menjadi pionir, yaitu <http://data.octo.dc.gov/>. Katalog data tersebut berisi beberapa set data seperti data kejahatan yang dimasukkan secara *real-time*, skor ujian sekolah, dan indikator kemiskinan. Lomba itu menjadi saat pembentukan katalog data daerah yang paling lengkap di dunia. Tantangannya adalah membuat katalog tersebut berguna bagi warga, pengunjung, organisasi bisnis, dan lembaga-lembaga pemerintah Washington, DC.

Solusi kreatifnya adalah membuat kontes Apps for Democracy. Penyelenggara meminta peserta membuat aplikasi menggunakan data dari katalog data yang baru saja diluncurkan, dan mengirimkan aplikasinya secara online. Kontes tersebut memberi hadiah-hadiah kecil selain beberapa hadiah besar, dan berbagai kategori seperti pemenang favorit. Kompetisi tersebut dibuka selama 30 hari dan menghabiskan biaya pemerintah DC sebanyak \$50,000. Hasilnya adalah 47 aplikasi iPhone, Facebook dan web, senilai sumbangan kira-kira lebih besar dari \$2,600,000 bagi perekonomian daerah.

The Abre Datos (Open Data) Challenge 2010. Lomba yang diadakan di Spanyol, April 2010, ini mengundang pembuat perangkat lunak untuk menciptakan aplikasi open source yang memanfaatkan data publik hanya dalam waktu 48 jam. Lomba ini diikuti 29 tim peserta. Peserta yang membuat aplikasi ponsel untuk mengakses informasi lalu lintas di Basque Country, untuk mengakses data bus dan halte bus di Madrid, masing-masing meraih juara satu dan dua dengan hadiah C3,000 dan C2,000.

Nettskap 2.0. Pada April 2010, Kementerian Pemerintahan Norwegia (*the Norwegian Ministry for Government Administration*) mengadakan lomba “Nettskap 2.0”. Para pembuat perangkat lunak/developers – baik perusahaan, badan publik, maupun perorangan – ditantang untuk mewujudkan ide-ide proyek berbasis-web pada area pengembangan jasa, proses kerja yang efisien, dan peningkatan partisipasi demokratis. Peserta sangat didorong menggunakan data pemerintah. Meskipun tenggat waktu penyerahan aplikasi masih sebulan lagi, pada 9 Mei,

Menteri Rigmor Aasrud menyatakan bahwa tanggapan atas lomba tersebut “luar biasa”. Secara total diterima 137 aplikasi, tidak kurang dari 90 di antaranya dibangun dari penggunaan-kembali data pemerintah. Hadiah total sejumlah NOK 2.5 juta dibagikan pada 17 pemenang; sedangkan jumlah total yang diajukan untuk 137 aplikasi tersebut adalah NOK 28.4 juta.

Mashup Australia. The Australian Government 2.0 Taskforce mengundang warga untuk menunjukkan mengapa akses terbuka ke informasi pemerintah Australia akan berdampak positif bagi pembangunan ekonomi dan social negara. Lomba itu berlangsung dari 7 Oktober hingga 13 November 2009. The Taskforce merilis beberapa data dengan ijin terbuka/open license dan dalam sejumlah format yang dapat digunakan kembali. Sebanyak 82 aplikasi yang masuk ke lomba tersebut adalah bukti bahwa aplikasi baru dan inovatif bisa muncul dari data pemerintah yang dibuka ke publik.

Konferensi, Barcamps, Hackdays

Salah satu cara yang lebih efektif yang dilakukan CSO untuk menunjukkan pada pemerintah nilai penting dari membuka data mereka adalah dengan menunjukkan berbagai cara bagaimana informasi dikelola untuk mencapai manfaat sosial dan ekonomi. CSO yang mempromosikan penggunaan-kembali telah berguna di negara-negara di mana ada kemajuan dalam kebijakan dan hukum untuk menjamin bahwa data dibuka secara teknis dan legal. Kegiatan umum yang dilakukan sebagai bagian dari inisiatif open data biasanya berupa konferensi open data pemerintah, “unconferences”, lokakarya, dan “hackdays”. Kegiatan-kegiatan tersebut seringkali diselenggarakan oleh kelompok pengguna dengan data yang telah dipublikasikan secara proaktif atau didapatkan menggunakan akses ke permohonan informasi. Pada kasus lain, masyarakat sipil bekerja dengan pejabat publik yang progresif untuk melindungi rilis baru set data yang dapat digunakan oleh programmer untuk menciptakan aplikasi-aplikasi inovatif.