

Pengembangan Sistem Portal Satu Data Indonesia Pada Kantor Staf Presiden Menggunakan *Comprehensive Kerbal Archive Network* (CKAN)

Bayu Wicaksono¹, Denny Sagita Rusdianto², Adam Hendra Brata³

Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹baywica@gmail.com, ²denny.sagita@ub.ac.id, ³adam@ub.ac.id

Abstrak

Data adalah sumber daya nasional yang berharga dan merupakan aset strategis bagi Pemerintah dan Masyarakat Indonesia. Mengelola data sebagai sebuah aset dan membuat data tersebut tersedia, dapat ditemukan dan dapat digunakan untuk mendorong keterbukaan, efisiensi, dan efektivitas pemerintahan. Data sejauh ini kurang diperhatikan oleh instansi. Sehingga perlu adanya penelitian untuk melakukan integrasi data antar lebih dari dua platform yang berbeda dengan menggunakan *comprehensive kerbal archive network* (CKAN). Untuk dapat melakukan integrasi data, terdapat metode yang penggunaannya didasari ketersediaan sumber *metadata* dan melalui *application programming interface* (API) CKAN. Memanen data dibagi menjadi tiga tahap yaitu tahap pertama mengumpulkan sumber *metadata* yang akan diambil melalui API CKAN. Tahap selanjutnya yaitu mengambil *metadata* dari sumber terkait dan menyimpannya ke dalam *database*. Tahap terakhir yaitu memasukkan *metadata* ke dalam *dataset*. Penelitian ini memiliki tujuan untuk melakukan integrasi antar sistem CKAN. Sistem ini telah berhasil diuji dengan menggunakan metode *white box testing* pada pengujian unit dan pengujian integrasi serta *black box testing* pada pengujian validasi. Pengujian sinkronisasi data dilakukan dengan presentase keberhasilan diatas 90%, dimana pengujian tersebut dilakukan dengan menggunakan 2 sampel pada portal data Bandung dan portal data Nasional. Portal data Bandung menunjukkan data yang berhasil diambil sebesar 100% dan portal data Nasional sebesar 99.87%.

Kata kunci: *data, metadata, CKAN, API, database, integrasi.*

Abstract

Data is a valuable national resources and strategic asset for Government and society in Indonesia. Manage data as an assets and make it available, can be found and used for encourage government's openness, efficiency and effectivity. So far, data become less attention from instance, so it needs some research to do data integration between more than two different platform using comprehensive kerbal archive network (CKAN). To be able to do data integration, there's some method the usage is based on the availability of metadata source and through application programming interface (API) CKAN. Data Harvesting are divided into three steps, the first step is collecting metadata source to be taken through API CKAN. The second step is take the metadata from related source and save it into database. The last step is input metadata into dataset. The purpose of this research is to do integration between CKAN system. This system has been successfully tested using white box testing method on unit testing and integration testing, also black box testing on validation testing. Testing data synchronization is done with percentage of success above 90%, where the test is done by using 2 samples on Bandung data portal and National data portal. Bandung data portal shows successful data taken as much as 100% and the National data portal is 99.87%.

Keywords: *data, metadata, CKAN, API, database, integration.*

1. PENDAHULUAN

Teknologi informasi pada era globalisasi saat ini telah berkembang pesat. Hal tersebut memberikan dampak kepada seluruh bidang

kehidupan manusia. Salah satunya adalah kegiatan pemerintahan di Indonesia, berbagai kegiatan pemerintah yang telah memanfaatkan teknologi komputer dengan tujuan untuk lebih efisien. Sebelum teknologi komputer berkembang pada umumnya banyak perusahaan

dan instansi pemerintahan masih menggunakan sistem manual yang bekerja secara sederhana. Tetapi seiring perkembangan jaman dan pemikiran manusia semakin maju, sebuah badan usaha bahkan instansi pemerintah dapat melakukan pemasukkan data dan penyebaran informasi untuk diolah dengan bantuan komputer (Sutriyani, 2013). Instansi pemerintah yang memanfaatkan teknologi komputer untuk dapat membantu kebijakan pemerintah berdasarkan teknologi, termasuk Kantor Staf Presiden.

Banyak sekali data yang berada pada Kementerian, Lembaga Pemerintah maupun Pemerintah daerah (K/L/D) tetapi tidak adanya platform yang memiliki interoperabilitas membuat data itu sulit ditemukan karena tidak tersedia atau tidak mudah diakses, padahal kebutuhan akan data sangatlah tinggi terutama dari akademisi, komunitas IT maupun masyarakat pada umumnya. Pada hasil wawancara dengan Bapak Robertus Theodore selaku manager Satu Data Indonesia, terdapat pengembangan yang ingin dilakukan yaitu K/L/D (Kementerian/Lembaga/Satuan Kerja Perangkat daerah) yang mulai menggunakan portal data sendiri membuat banyak sekali *dataset* yang perlu di satukan dalam portal nasional. Satu Data Indonesia adalah sebuah inisiatif Kantor Staf Presiden untuk meningkatkan interoperabilitas dan pemanfaatan data pemerintah melalui portal *data.go.id*. Portal ini berisi data kementerian, lembaga pemerintahan, pemerintah daerah dan semua instansi terkait lainnya untuk menghasilkan data yang berhubungan dengan Indonesia (datagoid, 2016). Data adalah sesuatu yang belum memiliki nilai bagi penerimanya dan masih harus dilakukan pengolahan. Data bisa berwujud suatu keadaan, gambar, suara, huruf, angka ataupun simbol-simbol lainnya. Dengan banyaknya data tersebut perlu dilakukan data *harvesting* terhadap portal K/L/D yang akan membuatnya terhubung langsung ke dalam portal nasional. Sehingga saat terdapat perbaruan data pada portal K/L/D portal nasional juga akan diperbarui secara langsung.

CKAN adalah sebuah sistem manajemen data yang membuat data dapat diakses - dengan menyediakan alat untuk merampingkan penerbitan, berbagi, menemukan dan menggunakan data (CKAN, 2016). CKAN ditujukan untuk penerbit data (pemerintah nasional dan daerah, perusahaan dan organisasi) yang ingin membuat datanya terbuka dan

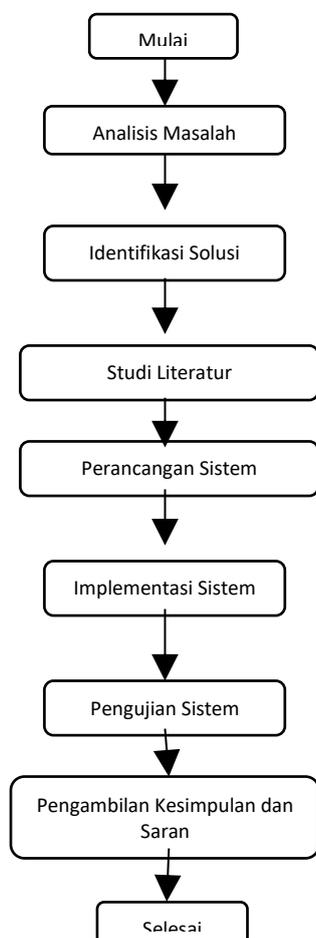
tersedia. Digunakan oleh pemerintah nasional dan daerah, lembaga penelitian, dan organisasi lain yang mengumpulkan banyak data. Dengan satu data Indonesia diharapkan data yang diterbitkan memiliki integritas tinggi, dimutakhirkan, dan dapat diakses luas secara cuma-cuma dalam format data yang mudah untuk digunakan kembali atau dibagai pakai oleh pengguna data. Dengan adanya *data harvesting* (pemanen data) ini diharapkan portal satu data Indonesia ini bisa menjadi rujukan untuk mencari data K/L/D bagi masyarakat umum dan masyarakat juga mengetahui portal K/L/D yang datanya terintegrasi dengan portal nasional.

Pemanen data adalah proses di mana sebuah skrip dijalankan juga dikenal sebagai bot untuk secara otomatis mengekstrak sejumlah data dari situs *web* dan menggunakannya (Dataharvesting, 2014). Sebagai cara mudah untuk mengumpulkan data, teknik ini sering digunakan pada website-website diluar sana untuk mengambil data pada website lain. Sedangkan pilihan penulis setelah meneliti menggunakan dengan cara mengakses metadata, hal ini dapat dilakukan dengan cara memanfaatkan API (*Application Programming Interface*) CKAN. Dalam Penelitian ini titik akhir penulis dapat memanen data (*data harvesting*) pada portal yang menggunakan sistem yang sama yaitu *comprehensive kerbal archive network* (CKAN) dalam kasus ini di Indonesia sudah cukup banyak pemerintah daerah, Kementerian maupun Lembaga pemerintahan yang mulai ataupun akan mengembangkan portalnya dengan yang sistem yang sama seperti portal Nasional. Sebagai contoh Pemerintah Provinsi Jakarta dan Pemerintah Kota Bandung sudah menggunakan portal yang sama dengan portal nasional dan akan menyusul Pemerintah daerah lainnya untuk mengembangkan portal sendiri. Dengan begitu perlu adanya pemanen data (*data harvesting*) agar tidak adanya masukan data berulang-ulang ke dalam portal yang berbeda dan dapat terjadi integrasi antar portal yang mempermudah kerja K/L/D (Kementerian, Lembaga Pemerintahan dan Pemerintah Daerah).

2. METODOLOGI

Pada bab ini akan diuraikan langkah-langkah penelitian yang akan digunakan dari Pengembangan sistem portal Satu Data Indonesia yang terdiri dari beberapa tahapan yaitu analisis domain masalah, identifikasi solusi, studi literatur, pengembangan perangkat

lunak dan pengambilan keputusan serta saran. Diagram alir metodologi penelitian dapat dilihat pada Gambar 2



Gambar 1. Diagram Alir Tahapan Penelitian

Pada penelitian pada Satu Data Indonesia ini penulis berkesempatan untuk mewawancarai dan melakukan observasi serta menggali langsung terhadap permasalahan yang terdapat dalam Satu Data Indonesia sesuai dengan Gambar 3.1 khususnya pada poin analisis masalah. Penulis melakukan wawancara kepada Robertus Theodore selaku manager Satu Data Indonesia sekaligus Tenaga Ahli Kantor Staf Presiden, penulis juga melakukan observasi selama 3 bulan pada satu data Indonesia. Dalam wawancara tersebut membahas bagaimana masalah-masalah pada Satu Data Indonesia terkait portal yang akan dikembangkan. Permasalahan utama adalah bagaimana portal Satu Data Indonesia dapat terintegrasi dengan portal Kementerian, Lembaga Pemerintahan dan Lembaga Daerah (K/L/D) melalui sistem yang sama yaitu *comprehensive kerbal archive network (CKAN)* dimana sistem itu sekarang sudah ada beberapa K/L/D yang melakukan

implementasi dan selanjutnya ingin melakukan integrasi ke dalam portal Nasional.

Setelah wawancara tersebut penulis melakukan observasi selama 3 bulan untuk mengetahui proses bisnis Satu Data Indonesia secara langsung sesuai dengan gambar 3.1 khususnya pada poin studi literatur, selama observasi itu dilakukan penulis mempelajari *Data Management System* yang digunakan dalam Portal Satu Data Indonesia yaitu *comprehensive kerbal archive network (CKAN)* dimana *system* tersebut menggunakan bahasa *python* dalam pengembangan backend dan *PostgreSQL* untuk *database engine*. *CKAN* sendiri sudah digunakan oleh pemerintah dari Negara lain untuk mendukung keterbukaan data pemerintah beberapa Negara yang menggunakan *CKAN* untuk pengembangan portalnya adalah *United Kingdom, Amerika, Australia* dan masih banyak lagi, yang membuat penulis melanjutkan untuk melakukan pengembangan lebih lanjut menggunakan *CKAN*. Dalam pengujian sistem peneliti melakukan pengujian sinkronisasi data untuk memastikan data yang masuk sama dengan data yang berada dalam portal sumber.

Diharapkan dari hasil penelitian yang penulis lakukan dapat membantu integrasi data pada portal yang menggunakan sistem yang sama yaitu *comprehensive kerbal archive network (CKAN)* dalam kasus ini di Indonesia sudah ada beberapa pemerintah daerah, Kementerian maupun Lembaga pemerintahan yang mulai ataupun akan mengembangkan portalnya dengan yang sistem yang sama seperti portal Nasional. Sebagai contoh Pemerintah Provinsi Jakarta dan Pemerintah Kota Bandung sudah menggunakan portal yang sama dengan portal nasional dan akan menyusul Pemerintah daerah lainnya untuk mengembangkan portal sendiri. Dengan begitu perlu adanya pemanen data (*data harvesting*) agar tidak adanya memasukkan data berulang-ulang ke dalam portal yang berbeda dan dapat terjadi integrasi antar portal yang mempermudah kerja K/L/D (Kementerian, Lembaga Pemerintahan dan Pemerintah Daerah).

3. PENGUJIAN DAN ANALISIS

Tahap pengujian dilakukan setelah melakukan implementasi. Pengujian ini bertujuan untuk memeriksa apakah hasil implementasi sudah sesuai dengan analisis kebutuhan dan perancangan sistem atau tidak. Tahap pengujian yang akan dilakukan adalah

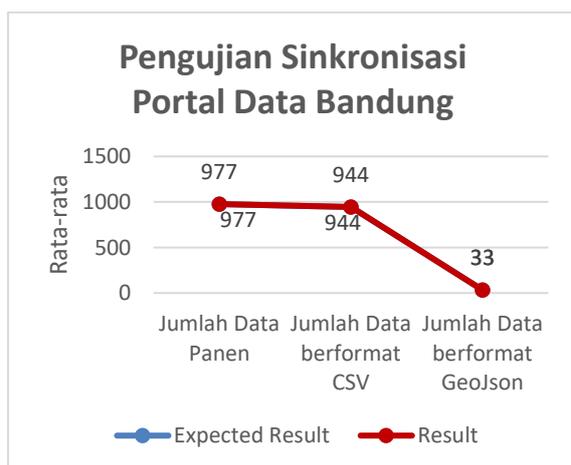
pengujian sinkronisasi data, unit, pengujian integrasi, dan pengujian validasi.

3.1 Pengujian Sinkronisasi Data

Pengujian sinkronisasi data ini digunakan untuk menguji tingkat keberhasilan data yang berhasil dipanen melalui sistem yang berbeda. Pada pengujian ini dilakukan analisis perbandingan terhadap data yang berada pada portal yang memanen data dan portal sumber panen data tersebut dimana total dari kumpulan data yang terdapat dalam portal sumber dibandingkan dengan total kumpulan data dari portal yang memanen. Portal data Bandung memiliki beberapa data yang berada dalam sistem, kumpulan data tersebut berformat CSV dan GeoJson. Berikut ini penulis melakukan pengujian pertama sinkronisasi data perbandingan antar portal Data Bandung yang penulis panen datanya.

Tabel 1. Pengujian Sinkronisasi Portal Data Bandung

No	Test Case	Expected Result	Result
1	Jumlah Data Panen	977	977
2	Jumlah Data Berformat CSV	944	944
3	Jumlah Data Berformat GeoJson	33	33



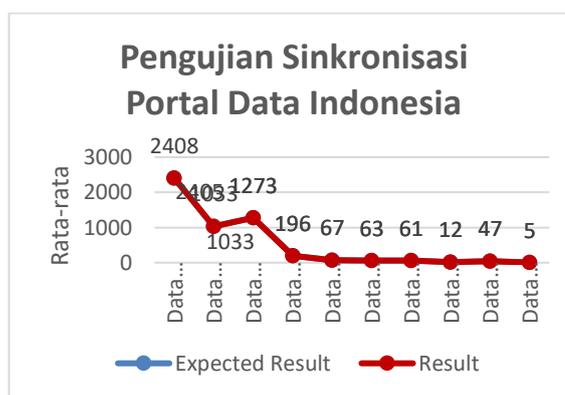
Gambar 2. Grafik Pengujian Sinkronisasi Bandung

Jumlah data yang penulis panen ujikan ini dapat dilihat pada gambar 3. Pada pengujian ini terdapat total kumpulan data sebesar 977 dengan kumpulan data yang berformat CSV sebesar 944 dan berformat GeoJson sebesar 33. Maka dari analisis tersebut nilai presentase keberhasilan panen data yang diterima sebesar 100 % yang

didapat dari hasil matematik jumlah data yang seharusnya dipanen dan hasil data yang berhasil dipanen. Penulis juga melakukan pengujian kedua dengan menggunakan portal Satu Data Indonesia sebagai *sample* dalam penelitian berikut hasilnya.

Tabel 2. Pengujian Sinkronisasi Portal Data Indonesia

No	Test Case	Expected Result	Result
1	Jumlah Data Panen	2408	2405
2	Jumlah Data Berformat CSV	1033	1033
3	Jumlah Data Berformat XLS	1273	1273
4	Jumlah Data Berformat xlsx	196	196
5	Jumlah Data Beformat JSON	67	67
6	Jumlah Data Berformat KML	63	63
7	Jumlah Data Berformat SOAP	61	61
8	Jumlah Data Berformat API	12	12
9	Jumlah Data Berformat HTML	47	47
10	Jumlah Data Berformat KMZ	5	5



Gambar 3. Grafik Pengujian Sinkronisasi Portal Nasional

Jumlah data yang penulis ujikan ini dapat dilihat gambar 4, pada pengujian kedua ini semakin sedikit *loss package* saat melakukan integrasi data antar sistem ini dari 2408 kumpulan data yang akan dipanen hanya sebesar 2405 kumpulan data yang berhasil dipanen yang 3 diantaranya dikarenakan data tersebut bersifat privat. Maka dari analisis tersebut nilai presentase keberhasilan panen data yang

diterima sebesar 99.87 % yang didapat dari hasil matematik jumlah data yang seharusnya dipanen dan hasil data yang berhasil dipanen.

3.2 Pengujian Unit

Pengujian unit ini digunakan untuk menguji unit (misalnya: komponen, klas, atau objek) dari perancangan yang telah dibuat. Pada pengujian unit ini menggunakan metode white box testing dengan jenis pengujian basis path. Pengujian unit akan dilakukan pada tiga sampel uji yaitu pada klas harvester, package, dan user.

Tabel 3. Pengujian Unit klas package untuk operasi new()

No Jalur	Prosedur Uji	Expected Result	Result
1	Memanggil Operasi new() dengan status cek session(user), session(auth_user), dan session(logged_in)	Menyimpan kumpulan data sebagai draft	Menyimpan kumpulan data sebagai draft
2	Memanggil Operasi new() dengan status cek session(user), session(auth_user), dan session(logged_in) dengan kondisi aktor tidak memenuhi syarat membuat kumpulan data	Menampilkan halaman dengan pesan 'Tidak punya izin untuk membuat paket'	Menampilkan halaman dengan pesan 'Tidak punya izin untuk membuat paket'
3	Memanggil Operasi new() dengan status cek session(user), session(auth_user), dan session(logged_in) dengan kondisi tag terdapat error	Menampilkan pesan error panjang tag maksimal 100	Menampilkan pesan error panjang tag maksimal 100
4	Memanggil Operasi new() dengan status cek session(user), session(auth_user), dan session(logged_in)	Menampilkan pesan 'Nama URL tersebut sudah digunakan'	Menampilkan pesan 'Nama URL tersebut sudah digunakan'

n) dengan kondisi url data sudah terdapat dalam database

3.3 Pengujian Integrasi

Pengujian ini berfokus pada desain dan konstruksi dari arsitektur perangkat lunak. Pengujian integrasi digunakan untuk menguji klas-klas yang saling berhubungan. Pengujian integrasi dilakukan dengan menggunakan white box testing dengan teknik basis path testing. Terdapat dua sampel pengujian integrasi yang dilakukan. Pengujian yang pertama akan dilakukan pada operasi read() pada klas tag sebagai operasi yang memanggil operasi search() pada klas Package. Pengujian yang kedua akan dilakukan pada operasi search_url() pada klas harv_organization sebagai operasi yang memanggil operasi bulk_process() dan read() pada klas organization dan operasi read pada klas group

Tabel 4. Pengujian Integrasi klas Tag pada operasi read()

No Jalur	Prosedur Uji	Expected Result	Result
1	Memanggil Operasi read() dengan inialisasi cek model, session(user), dan session(auth_user) dengan memilih dataset berdasarkan tag	Menampilkan halaman dataset berdasarkan tag	Menampilkan halaman dataset berdasarkan tag
2	Memanggil Operasi read() dengan inialisasi cek model, session(user), dan session(auth_user) dengan memilih halaman tag	Menampilkan halaman tag	Menampilkan halaman tag
3	Memanggil Operasi read() dengan inialisasi cek model, session(user), dan session(auth_user)	Menampilkan halaman 404 tag tidak ditemukan	Menampilkan halaman 404 tag tidak ditemukan

er) dengan tag tidak ditemukan

3.4 Pengujian Validasi

Pengujian validasi digunakan untuk mengetahui apakah sistem yang telah dibangun sudah sesuai dengan seluruh kebutuhan yang telah ditetapkan berdasarkan hasil proses analisis kebutuhan atau tidak. Pengujian validasi dilakukan dengan cara memeriksa apakah sistem sudah berjalan dengan baik dan tidak ada kesalahan yang terjadi. Pada pengujian validasi ini digunakan metode *black box testing*. Dengan *black box testing* memungkinkan pembuat perangkat lunak untuk memperoleh kondisi yang terjadi untuk suatu masukan yang akan menjalankan semua kebutuhan.

Tabel 5. Pengujian Validasi Berhasil Tambah Sumber Harvest

Nama Kasus Uji	Pengujian Tambah Sumber Harvest Berhasil
Prosedur	<ol style="list-style-type: none"> 1. Membuka halaman sumber <i>harvest</i> 2. Menekan tombol tambah Sumber <i>Harvest</i> 3. Mengisi Form <i>URL</i>, Judul, nama, Deskripsi, tipe sumber, frekuensi perbaruan, konfigurasi, dan organisasi 4. Menekan tombol "<i>simpan</i>"
Hasil yang diharapkan	Sistem akan menampilkan sumber <i>harvest</i> yang telah dibuat
Hasil	Sistem menampilkan sumber <i>harvest</i> yang telah dibuat
Status	Valid

4. KESIMPULAN

Berdasarkan hasil perancangan yang telah dilakukan menghasilkan rancangan klas-klas untuk melakukan data harvesting dengan cara *metadata harvesting* dimana keseluruhan data pada portal K/L/D memiliki API *metadata* yang membuat proses *data harvesting* dapat dilakukan. Pada awalnya sistem akan melakukan *data harvesting* dengan validasi config berupa inputan dari *sysadmin*, dari config itu diperoleh *api key* dan info portal lainnya. Setelah itu akan dilakukan pengumpulan terhadap keseluruhan *metadata* portal sumber, tahap selanjutnya yaitu tahap pengambilan *metadata* dan menyimpannya dalam *database*. Tahap terakhir tahap impor data dan membuat kumpulan data pada sistem.

Berdasarkan pengujian yang telah dilakukan, pengujian ini menggunakan metode white box testing untuk pengujian unit dan pengujian integrasi, dan metode black box testing untuk pengujian validasi. Dari seluruh hasil pengujian yang telah dilakukan menyatakan hasil yang valid atau tidak terdapat kesalahan dalam system. Sedangkan pengujian sinkronisasi dilakukan untuk melakukan pengecekan terhadap data yang masuk pada sistem dari portal sumber, pada sampel pertama melakukan panen data dari Portal Data Bandung. Dari 977 kumpulan data yang akan diambil, data yang masuk dalam sistem 977 kumpulan data dikarenakan tidak terdapat kesalahan pada sistem saat melakukan pengambilan data dengan presentase keberhasilan 100%. Sedangkan pada sampel kedua dari 2408 kumpulan data, data yang berhasil masuk pada sistem berjumlah 2405 dengan presentase keberhasilan 99.87%. Hal ini membuktikan bahwa sistem ini dapat digunakan untuk Integrasi antar sistem CKAN pada portal Satu Data Indonesia.

DAFTAR PUSTAKA

- Bapennas. 2017. Paparkan 10 Prioritas nasional Tersedia di <http://ppid.bapenas.go.id/menteri-bambang-paparkan-10-prioritas-nasional-dan-30-program-prioritas> [Diakses 25 Mei 2017].
- Booch, Grady. 1999. *The Unified Modeling Language Reference Manual*, Addison-Wesley.
- Ckan. 2016. *About CKAN* Tersedia di <http://ckan.org/developers/about-ckan/> [Diakses 23 Januari 2017].
- Datagoid. 2016. Pedoman Inisiatif Satu Data <http://data.go.id/toolkit/> [Diakses 23 Januari 2017].
- Dataharvesting, 2014. What You Need to Know About Data Harvesting and How to Prevent it <http://blog.caspio.com/tech-tips/what-you-need-to-know-about-data-harvesting-and-how-to-prevent-it> [Diakses 5 Mei 2017]
- Henderi. 2008. *Unified Modelling Language*. Tangerang: Raharja Enrichment Centre (REC).
- Kspgoid. 2016. Tentang Kantor Staf Presiden Tersedia di <http://ksp.go.id/tentang-kantor-staf-presiden/> [Diakses 23 Januari 2017].

- Opendatahandbook. 2016. Buku Pegangan Data Terbuka. Open Knowledge International
- Opendatalabjakarta. 2016. *Mempersiapkan Data Terbuka*. Open Data Lab Jakarta.
- Pressman, R.S. 2010. *Software Engineering : a practitioner's approach*, McGraw-Hill, New York.
- UK Data Archive. 2012. *Preservation policy*. UK Data Archive, University of Essex.
- Unit Kerja Bidang Pengawasan dan Pengendalian Pembangunan (Ukp4). 2014. *Cetak Biru Satu Data*.
- Whitten L,Jeffery, Bentley D,Lonnie,Dittman C. 2004. *Metode Desain dan Analisis Sistem*. Terjemahan oleh Tim Penerjemah ANDI. 2004. ANDI:Yogyakarta.