

PENGUNAAN FORENSIK METADATA UNTUK MENDETEKSI PLAGIARISME PADA DOKUMEN MICROSOFT WORD

Nugroho Budhisantosa
Fakultas Ilmu Komputer Universitas Esa Unggul Jakarta
Jln. Arjuna Utara Tol Tomang-Kebon Jeruk Jakarta 11560
nugroho.budhisantosa@esaunggul.ac.id

Abstract

Plagiarism is the act of taking essays or opinions from others people and make themselves as essays or opinions maker. Now day, in the era of information technology where assignment reporting formats of students have now been made in the formats of Microsoft Word electronic documents, plagiarism activities becomes a very easy things to by only need asking assignment reports from his/her colleagues following by editing process to change the name of author and the replace the font format and ended by sending the documents to the lecturers. Information technology provides not only convenience to plagiarism but also lecturer to detect it. This paper describes how the knowledge of computer forensics can be performed on Microsoft Word's metadata for detection of plagiarism activity as quickly. Using this method, plagiarism activities can be detected by comparing Authors name versus Last save by information, dissimilarity of the two is the first clue of the existence of plagiarism activities while the total sum of editing time and the Revision number will show the file editing history information where time editing recorded in minutes and the number of revisions that few can strengthen their alleged of plagiarism activities.

Keywords : forensic, metadata, plagiarism

Abstrak

Plagiarisme atau sering disebut plagiat adalah perbuatan pengambilan atau penjiplakan karangan, pendapat, dan dari orang lain dan menjadikannya seolah karangan dan pendapat sendiri. Pada era teknologi informasi dimana format-format pelaporan tugas mahasiswa kini telah dibuat didalam format-format dokumen elektronik Microsoft Word, kegiatan plagiarisme menjadi sangat mudah dilakukan yaitu dengan meminta berkas laporan dari kolega dan melakukan proses editing yang umumnya dilakukan dengan mengganti nama pengarang dan melakukan penggantian huruf untuk kemudian mengirimkan dokumen hasil jiplakan tersebut kepada dosen yang bersangkutan. Teknologi informasi selain memberikan kemudahan pada plagiat juga memberikan kemudahan yang sama untuk para pendidik untuk mendeteksinya. Tulisan ini bertujuan membahas bagaimana pengetahuan forensik komputer yang dilakukan pada metadata berkas dokumen Microsoft Word dengan metode mengamati kewajaran informasi dari *Authors*, *Last saved by*, *Total editing time* dan *Revision number* yang ada di dalam metadata. Melalui pengamatan ini, adanya aktivitas *plagiarism* dapat dideteksi secara cepat tanpa perlu membuka berkas dan membaca kandungan isinya terlebih dahulu. Berkas dokumen asli akan menampakkan kesesuaian antara informasi nama *Authors* dan informasi *Last saved by* sedangkan *Total editing time* dan *Revision number* akan menampakkan informasi riwayat penyuntingan berkas dimana waktu penyuntingan yang tercatat dalam hitungan menit dan jumlah revisi yang sedikit dapat menguatkan adanya dugaan aktivitas *plagiarism*.

Kata kunci : forensik, *metadata*, plagiarisme

Pendahuluan

Plagiarisme atau sering disebut penjiplakan adalah perbuatan pengambilan atau penjiplakan karangan, pendapat, dari orang lain dan menjadikannya seolah karangan dan pendapat sendiri.

Tugas-tugas mata kuliah yang mensyaratkan dibuatnya pelaporan adalah hal umum di kalangan dunia pendidikan demikian juga yang terdapat pada sistem pendidikan di perguruan tinggi. Seiring kemajuan teknologi informasi, bentuk pelaporan tradisional yang berbasis kertas sekarang ini mulai ditinggalkan oleh para dosen dan mahasiswa, digantikan oleh dokumen-dokumen berkas elektronik yang lebih mudah didalam penanganannya. Kemajuan teknologi ini ternyata juga memberikan kemudahan bagi sebagian mahasiswa untuk melakukan perbuatan tidak terpuji, yaitu dengan melakukan plagiarisme dengan cara yang sangat mudah dengan cara melakukan modifikasi pada dokumen elektronik karya intelektual kepunyaan mahasiswa lain yang berhasil didapatkannya dan mengakuinya sebagai karyanya setelah melakukan perubahan pada nama pengarang dan kadang dikombinasikan dengan perubahan format jenis huruf. Plagiarisme adalah perbuatan yang tercela secara akademis karena perbuatan ini merupakan perbuatan pencurian kekayaan intelektual dari seseorang yang seharusnya perlu dijaga agar tidak terjadi di dalam dunia akademis.

Dosen sebagai pendidik memainkan peran terdepan dalam menjaga integritas kekayaan intelektual. Tindakan tegas berupa pemberian sanksi ringan berupa tindakan skorsing sampai sanksi berat seperti pencabutan ijazah sekolah/universitas dapatlah dilakukan terhadap pelanggaran kekayaan intelektual ini.

Aplikasi pengolah kata Microsoft Word adalah aplikasi umum yang hampir selalu ada pada komputer yang menggunakan sistem operasi Microsoft Windows sehingga aplikasi ini secara umum banyak digunakan mahasiswa untuk menyusun baik laporan-laporan tugas perkuliahan, juga di dalam penulisan tugas akhir atau skripsi mahasiswa.

Sama seperti aplikasi profesional lainnya, Microsoft sebagai pengembang aplikasi juga menambahkan informasi tambahan berupa

informasi metadata pada setiap berkas yang dibuat.

Metadata

Berbicara tentang pembuktian keaslian dokumen elektronik yang berisi karya intelektual seseorang dapat dilakukan melalui informasi *metadata* yang terkandung di dalam dokumen elektronik tersebut.

Situs Wikipedia menyebutkan “*Metadata* adalah informasi terstruktur yang mendeskripsikan, menjelaskan, menemukan, atau setidaknya menjadikan suatu informasi mudah untuk ditemukan kembali, digunakan, atau dikelola. Metadata sering disebut sebagai data tentang data atau informasi tentang informasi.” yang oleh National Information Standards Organization (NISO) pada publikasi onlinenya yang dapat diakses melalui <http://www.niso.org/publications/press/UnderstandingMetadata.pdf> menggolongkan *metadata* ke dalam tiga jenis utama yaitu:

1. *Metadata* deskriptif yaitu *metadata* yang menggambarkan suatu sumberdaya dalam bentuk penemuan dan identifikasi. *Metadata* jenis ini bisa meliputi elemen semisal judul, abstrak, pengarang, dan kata kunci.
2. *Metadata* struktural menunjukkan bagaimana kumpulan objek disusun secara bersama-sama menjadi satu, semisal bagaimana halaman-halaman ditata untuk membentuk suatu bab.
3. *Metadata* administratif menyediakan informasi untuk membantu mengelola sumberdaya, semisal terkait kapan dan bagaimana suatu informasi diciptakan, tipe dokumen dan informasi teknis lainnya, serta siapa yang bisa mengaksesnya.

Metadata bukanlah informasi yang dibuat oleh pengguna melainkan adalah informasi yang dibuat oleh sistem tanpa sepengetahuan pengguna, *metadata* juga bukanlah informasi yang mudah dilihat secara kasat mata. *Metadata* adalah informasi yang tersembunyi yang memerlukan teknik tersendiri untuk dapat membaca informasi yang terkandung di dalamnya.

Letak Metadata Pada Berkas Elektronik

Dalam suatu berkas elektronik, informasi *metadata* umumnya dapat ditemui pada awal *header* dari berkas, namun pada beberapa format berkas, informasi *metadata* dapat terletak pada bagian akhir *header* atau area lainnya di dalam *header*.

Header sendiri adalah istilah yang mengacu pada data tambahan yang ditempatkan pada blok data yang akan disimpan atau di transmisikan.

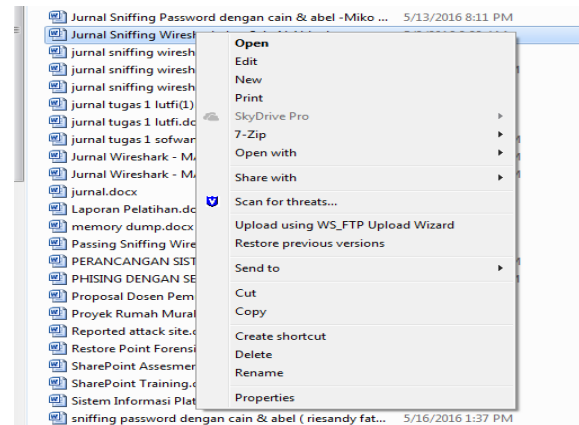
Format penulisan *metadata* pada kebanyakan berkas biasanya akan ditulis mengikuti format penulisan *header*-nya. Pada penulisan *header* berbasis teks, informasi *metadata* biasanya akan dituliskan didalam bentuk teks yang dengan mudah dapat dibaca oleh mata manusia, dan *header* dengan format penulisan binari biasanya akan menuliskan *metadata*-nya dalam bentuk binari pula.

Membaca Metadata

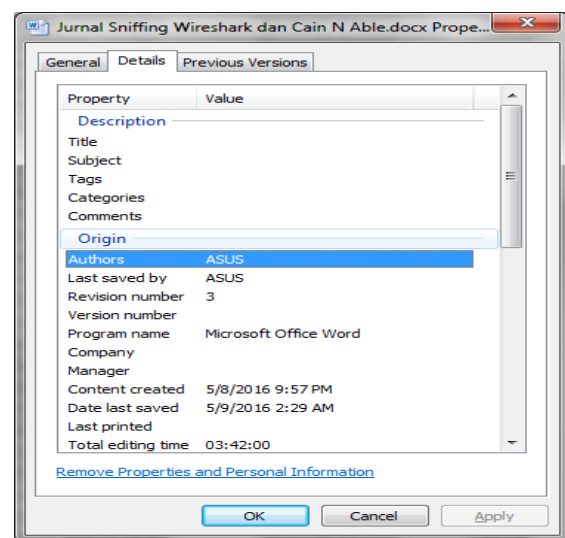
Membaca *header* berkas dari berkas Microsoft Word dalam bentuk aslinya tidaklah mudah untuk dilakukan karena dituliskan didalam format binari, namun sistem operasi Windows adalah sistem operasi ramah berbasis tampilan grafis yang memberikan kemudahan bagi penggunaannya untuk melakukan pembacaan *metadata* melalui *tool properties*. Melalui *tool properties* ini, format penulisan binari dari *header* berkas diterjemahkan kedalam format grafis yang dapat dibaca dengan mudah oleh pengguna.

Untuk melakukan pembacaan *metadata* pada berkas dokumen Microsoft Word, langkah pertama yang perlu dilakukan yaitu adalah dengan melakukan klik kanan pada berkas yang ingin diketahui informasi *metadata*-nya dan memilih opsi *properties* untuk kemudian dilanjutkan dengan memilih panel *Details* dari jendela *Properties* yang ditampilkan.

Pada jendela *Properties – Details* inilah informasi *metadata* berkas elektronik ditampilkan oleh sistem operasi Microsoft Windows di dalam bentuk tampilan grafis yang mudah dipahami mata awam.



Gambar 1
Klik Kanan untuk Opsi *Properties*



Gambar 2
Panel *Details* yang Menampilkan Informasi *Metadata*.

Dalam sistem operasi keluarga Linux terdapat aplikasi *Hexdump* yang merupakan aplikasi pembaca informasi *metadata* berbasis baris perintah.

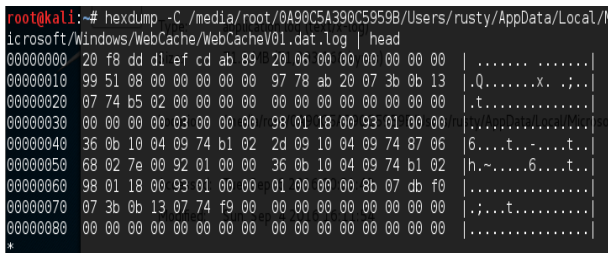
Kelebihan aplikasi *Hexdump* yang berjalan di dalam sistem operasi Linux dan *tool properties* yang ada di dalam sistem operasi Microsoft Windows adalah sifatnya yang gratis dan dapat digunakan dengan cepat untuk mendokumentasikan temuan. Pada aplikasi *Hexdump*, pemeriksaan *metadata* dapat dilakukan langsung melalui terminal dengan masuk kedalam folder berkas dan mengetikkan perintah:

```
#hexdump -C nama_file | head
```

Jika tampilan yang diinginkan adalah tampilan dalam bentuk berkas dengan nama berkas bukti01.txt maka perintah berikut dapat dijalankan:

```
#hexdump -C nama_file | head > bukti01.txt
```

Tampilan bentuk asli *header* dari berkas dokumen elektronik yang berisi informasi metadata dari sistem operasi Linux dapat dilihat pada Gambar 3.



Gambar 3
Tampilan Header

Selain menggunakan pembacaan *offline* menggunakan dua sistem operasi utama di atas, pembacaan *metadata* secara online juga dilakukan melalui beberapa situs yang menyediakan fasilitas pembacaan hanya saja proses pembacaan akan memerlukan koneksi Internet dan berjalan tidak secepat pembacaan secara *offline* serta informasi yang diberikan tidak mendetail. Salah satu situs yang dapat digunakan untuk keperluan pembacaan metadata adalah situs <http://www.extractmetadata.com/>

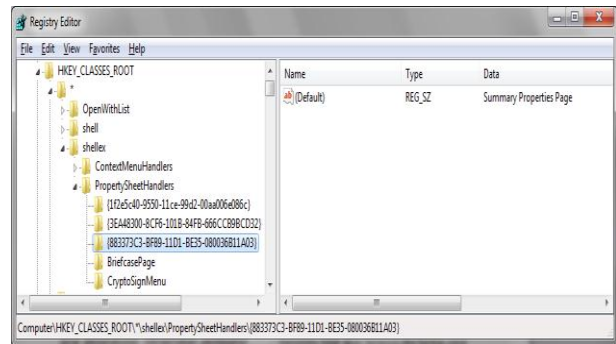
```
Result
Language: Bahasa
Mimetype: application/CDFV2-corrup
Creator: MISS EUIS
Unknown date: 2016-07-04T10:07:00Z
Character count: 11966
Last saved by: MISS EUIS
Page count: 13
Word count: 2099
Creation date: 2016-06-04T11:45:00Z
Editing cycles: 76
Mimetype: application/vnd.ms-files
Created by software: Microsoft Office Word
Template: Normal
Line count: 99
Paragraph count: 28
Mimetype: application/zip
Embedded filename: [Content_Types].xml
Embedded filename: _rels/.rels
Embedded filename: theme/theme/themeManager.xml
Embedded filename: theme/theme/theme1.xml
Embedded filename: theme/theme/_rels/themeManager.xml.rels
```

Gambar 4
Tampilan Grafis Informasi Metadata yang disediakan oleh <http://www.extractmetadata.com/>

Masalah Umum Pada Panel Details

Secara *default* panel *Details* akan ditampilkan oleh sistem operasi Microsoft Windows, namun pada beberapa kasus dapat saja panel ini tidak muncul dari jendela *Properties*. Hal ini dapat terjadi karena hilangnya *key* yang menampilkan panel *Details* di dalam *registry* sistem operasi. Hal ini dapat diatasi dengan menyalin *key* register berikut kedalam *registry* sistem operasi Microsoft Windows

```
[HKEY_CLASSES_ROOT\*\shellex\PropertySheetHandlers\{883373C3-BF89-11D1-BE35-080036B11A03}] @="Summary Properties Page"
```



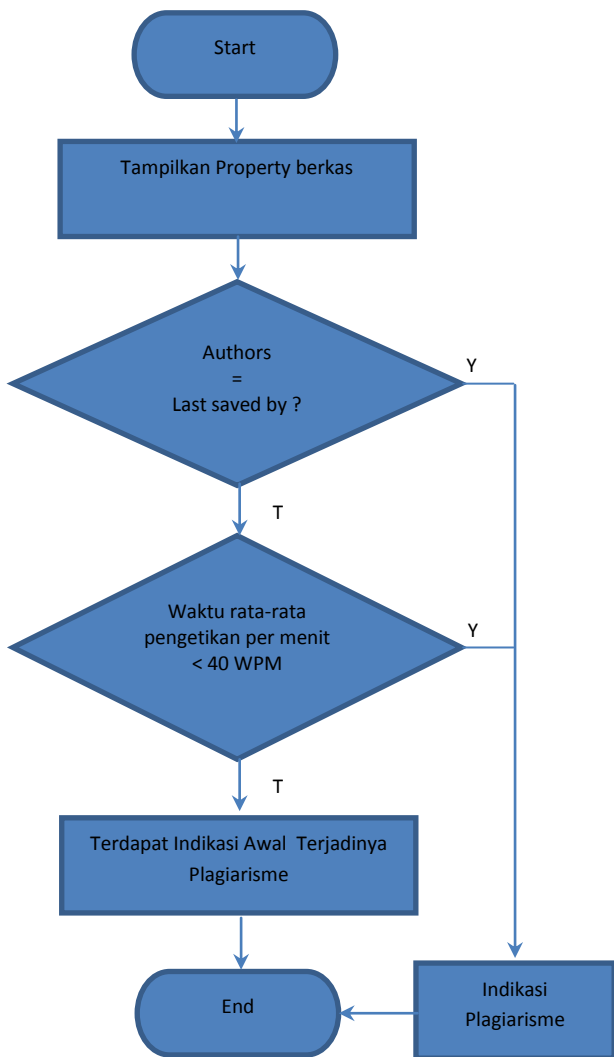
Gambar 5
Tampilan Key Registry dari Panel Details

Kecepatan Pengetikan Halaman

Brandon Raziano didalam satu studinya mengenai kecepatan rata-rata pengetikan yang dilakukan oleh pengguna komputer pada tahun 1998 menunjukkan bahwa kebanyakan pengguna komputer dapat mengetik sebanyak 33 kata per menit atau 33 *word per minute* (WPM) sedangkan pengguna dalam kelompok pemilik kemampuan khusus dapat mengetikkan kata hingga 40 WPM, waktu tersebut akan melambat jika pengkomposisian halaman diperlukan. Informasi ini diperoleh melalui https://en.wikipedia.org/wiki/Words_per_minute

Hal ini berarti kebanyakan orang dapat menyelesaikan pengetikan 1 halaman folio yang berisi sekitar 390 kata dalam waktu 11,8 menit, sedangkan pengguna mahir dapat melakukannya dalam waktu 9,75 menit.

Proses Analisis Plagiarisme



Gambar 6

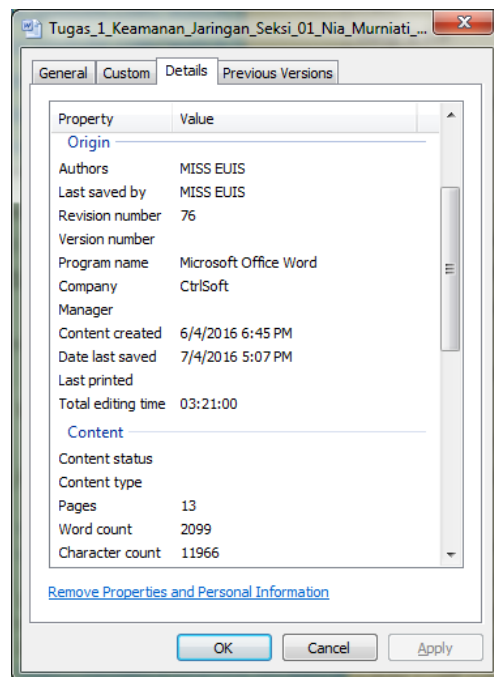
Diagram Alir Pemeriksaan Indikasi Plagiarism

Adanya aktivitas plagiarisme pada dokumen tugas perorangan mahasiswa yang dibuat dalam berkas Microsoft Word di dalam sistem operasi Microsoft Windows dapat dideteksi dengan mengamati kewajaran informasi-informasi yang terdapat pada metadata dari berkas dokumen Microsoft Word. Analisis plagiarisme pada berkas Microsoft World dapat dilakukan mengikuti diagram alir atau *flow chart* pada Gambar 6 di atas.

Petunjuk awal adanya plagiarisme pada dokumen dapat terlihat dari informasi *Authors* dan informasi *Last saved by* serta informasi *Total editing time* dan informasi *Word count*. Dokumen tugas perorangan yang dikerjakan oleh pengguna yang sama akan memiliki

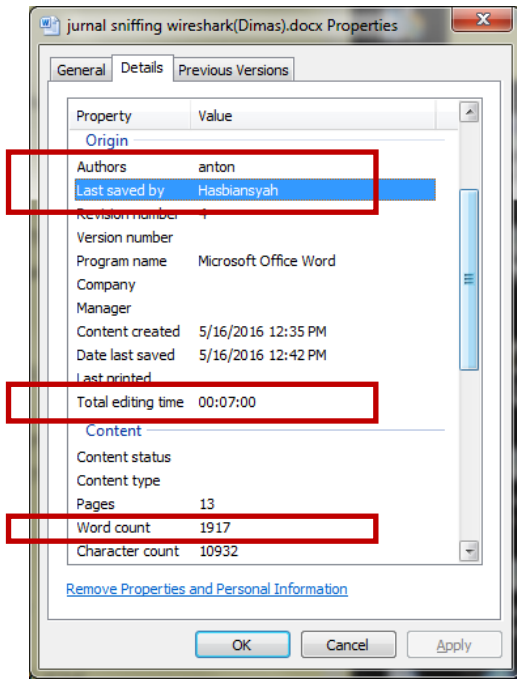
kesamaan nama pada informasi *Authors* dan informasi *Last saved by*, sedangkan informasi *Total editing time* dapat menginformasikan kecepatan pengetikan kata per menit atau Word Per Minute – WPM. Dokumen yang diperoleh dengan cara membagi jumlah total kata – *Word count* yang ada di dalam dokumen dengan *Total editing time*.

Pada Gambar 7 terlihat terdapatnya kesamaan antara informasi *Authors* dan informasi *Last saved by* mengindikasikan dokumen dibuat dan diselesaikan oleh mahasiswa yang sama sehingga dugaan adanya aktivitas plagiarism untuk sementara dapat dikesampingkan. Indikasi ini diperkuat melalui informasi yang dinerikan oleh jumlah *Total editing time* sebanyak 3 jam 21 menit atau 201menit yang jika digunakan sebagai angka pembagi dari total kata pada berkas yaitu sebanyak 2099 kata akan menghasilkan kecepatan pengetikan 10.4 kata per menit atau 10.4 WPM.



Gambar 7

Tampilan Metadata Berkas Non Plagiarisme



Gambar 8

Tampilan Metadata Berkas yang Patut diduga Mengandung Konten Plagiarisme

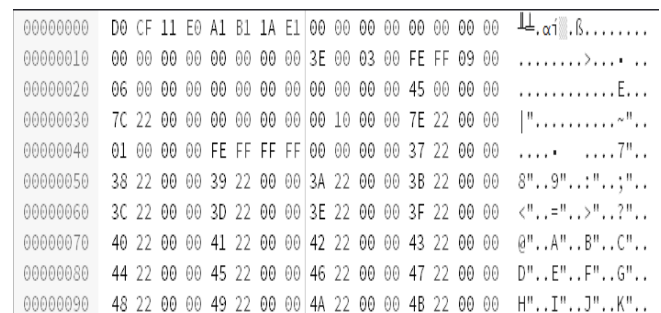
Gambar 8 adalah informasi *metadata* dari berkas dokumen Microsoft Word lainnya yang dibuat oleh mahasiswa yang menginformasikan adanya perbedaan antara informasi *Authors* dan informasi *Last saved by*. Hal ini menunjukkan bahwa dokumen dibuat dan diselesaikan oleh dua mahasiswa yang berbeda. Dalam temuan kasus seperti ini maka patut diduga telah terjadi aktivitas plagiarisme pada berkas yang memerlukan pemeriksaan lanjutan yaitu dengan menganalisa informasi *Word count* dan *total editing time* dimana kata sebanyak 1917 yang memerlukan waktu pengetikan hanya dalam waktu 7 menit saja, dalam hal ini jumlah pembagian antara *Word count* dan *total editing time* menghasilkan kecepatan pengetikan sebanyak 273.9 WPM.

Rekayasa Manipulatif Pada Berkas Metadata

Modus rekayasa manipulasi *metadata* selama ini banyak dilakukan dengan melalui penyuntingan langsung pada *header* berkas baik dengan melakukan teknik penyuntingan manual satu-persatu maupun dengan cara menempelkan informasi *header* dari *header* berkas lain yang memiliki kemiripan. Penyuntingan yang dilakukan dengan tepat pada *header* berkas

tidaklah akan merusak berkas tersebut karena informasi *metadata* sebenarnya hanyalah informasi tambahan atau keterangan yang ditambahkan pada berkas yang tidak akan merusak integritas berkas sekalipun informasi tersebut dihilangkan.

Seperti telah disebutkan sebelumnya di atas, Informasi *metadata* berkas Microsoft Word yang tersimpan di dalam *header* berkas tidaklah mudah dibaca secara langsung dan dilakukan penyuntingan secara langsung. Gambar 9 di bawah memperlihatkan contoh tampilan *header* berkas Microsoft Word yang “sulit dipahami” karena ditulis di dalam format binari.



Gambar 9

Contoh Tampilan Header berkas Microsoft Word

Beberapa *tools* seperti sejenis *MetadataTouch™ 7.0* dan *GPX editor* memang tersedia dan siap digunakan untuk memanipulasi isi informasi *metadata* yang akan sulit dilakukan secara manual menggunakan *tools WinHex* dan *HexEditor* standar, namun walaupun dilakukan dengan bantuan *tools* yang tepat sekalipun, melakukan tindakan manipulasi *metadata* yang menghasilkan hasil yang sempurna tidaklah mudah untuk dilakukan, terlebih jika manipulasi dilakukan pada penanda waktu dari informasi *metadata* karena forensik penanda waktu – *time stamp forensics* yang dilakukan langsung pada berkas dapat dengan mudah dilakukan untuk mengungkap adanya rekayasa manipulatif pada penanda waktu.

Lebih dari itu rekayasa manipulatif informasi *metadata* tidak akan berguna jika berkas yang telah mengalami rekayasa dibaca dan diperiksa dengan teliti oleh dosen karena dosen yang bersangkutan akan segera

menemukan adanya kesamaan isi dengan dokumen lainnya yang dibaca.

Kesimpulan

Pendeteksian awal adanya indikasi aktivitas plagiarisme dapat dilakukan dengan cepat melalui analisis informasi yang diberikan oleh *metadata* dari berkas dokumen.

Rekayasa manipulatif isi informasi *metadata* yang dilakukan memang mungkin dapat lolos dari proses *screening* awal analisis *metadata*, tetapi proses pemeriksaan melalui pembacaan dokumen yang dilakukan oleh dosen secara teliti untuk keperluan pemberian nilai tentu akan dapat dengan mudah menemukan adanya kesamaan isi antar dokumen.

Daftar Pustaka

Situs Internet:

[https://en.wikipedia.org/wiki/Header_\(computing\)](https://en.wikipedia.org/wiki/Header_(computing)), diakses 2 September 2016

<http://www.scanmyessay.com>, diakses 2 September 2016

http://www.tutorialspoint.com/unix_commands/hexdump.htm, diakses 2 September 2016

<https://id.wikipedia.org/wiki/Metadata>, diakses 2 September 2016

Understanding Metadata, National Information Standards Organization, NISO Press, 2004