Effecting and Discovering the Different Views

HTML Code	Web Views	
	Japanese View	
	を熱望しているのに、経理部は応えてくれない!と嘆く経営者のなんと多いことでしょう!	
sigh how is=20 many! <p>* " Monthly balancing of accounts " =being=20 completed with some timing? It is very important in age of =speed.=20 </p>	「月次決算」がどのタイミングで出来上が多のか!は、スピードの時代に非常に重要です。 大企業では基本的しまり月末が終わったも翌月85歳月以内」には前月の決算が構む固定っていま す。大企業はその事業の拠点も、取引の絶対教行任務においたもかからず、5歳集日で月次決算が 出来る!というのは、会社システムと言語地理の水準に伝っています。 会社システムは全社をネッチリークで訪んであり、会社処理は「発生時点」で処理されているわけです。 その為には各事業所に少なくとと会社処理ができ人材を構成しています。 通知決算の時代になって他のこと、会社システムの減一と発生時点の処理が重要になってきていま す。会社処理じる疑定はなお記念を大け必経理的に送って本社で一元入力をするような形態では15歳 業日」を達成するのは極めて困難です。	
<font at="" big="" business<br="" color="3D#fffff)*">basic "=when end of=20 the month ends, balancing of accounts of front month become #firm in=20 general within next month 5 business days ". 	月次決算のスピードアップは大企業だから出来る!ということではありません。もとより取引の絶対量 が大企業より少ない会社ではより早くできるはまです。ところが中堅・中小企業で「5営業日」を実現して	
	"English View"	
	sigh how is many!	
Big =business=20 base of the business and the absolute number of transactions =preponderantly is many monthly can do	Eig business base of the business and the absolute number of transactions preponderantly is many nonthly can do balancing of accounts in 5 business days in spite. In other words, it depends on the lavel of the accounting system and accounting processing	
balancing of accounts = in 5=20 business days in spite.! In other words, it depends on	That for the talent who can designate accounting	
the =level of=20 the accounting system and accounting processing. 	"English View on Gray background"	
<font=20 color="3D#ffffff">* The accounting system has tied the whole = company=20 with network, accounting processing</font=20>	• Monthly belancing of accounts: "Being completed with combitining" it is very important in age of speed. • At hig buckness bore, when not of the month and, isalancing of accounts of front month become time in general when next month 5 buckness days. "The month accounts of the start of the month spectra activity of the spectra act	
	The accounting system has tied the whole company with network, accounting processing at is the clear that it is processed if occurrence spirit in time.	

EXHIBIT

Houdini Tech on the Internet¹ (HTI): Friends see one thing – foes another

By: David Israel-Rosen, Oded Maimon, and James Cook LinkCo, Inc.; Glencoe, Illinois 60022, USA

Abstract: In this proposal we present results of research that clearly shows that the Internet is being used, with what we call HTI, to present one type of information to one audience defined as friends, and another type (confused, partial, misleading, misdirected) to another audience, called foes. We show examples employed by one foreign company to deceive specific non-national audiences (especially USA) and we go on to expose the methods and the ways being used. The methods presented here can be used by any organization, such as a terrorist group and alike.

Introduction:

We have come across intentional measures that manage different views for different communities of viewers. The measures themselves are neutral; they can be used for good or for bad purposes. For example, adjusting what is seen to viewers from different cultures is a good purpose; however these schemes, termed here HTI (Houdini Technology for Internet), as we found out, are also used for bad/criminal purposes, and in an organized and systematic way. Examples include, hiding messages, keeping information (such as evidence or messages) from appearing in one jurisdiction, but appearing in another (such as patent usage hiding, where you sell in friendly home market, while hiding in the foe market, e.g., where the patent comes from). Organizations can use HTI methods to manage the flow of information in such a way that they thwart detection of fraud and, potentially, command and control messages.

In this proposal we show some of the methods by which internet pages can be made covert to one audience and overt to another audience. We start with a brief survey of five types of the HTI measures that are used.

Survey of Measures:



¹ This is the updated version of the documents which includes the finding of Google translation engine hacking of March 27 and April 18 of 2008.

Here we present the methods themselves, which comprise the HTI. We only present methods that we found and deduced from actual field examples. It is not a comprehensive list yet, and may never be.

1. Color concealment (background and text in the same color): When the foe views the web page (after having submitted it to a translation engine) both the text and its background are the same color, thus making text invisible. One way that the designer can effect this is to install special instructions (probably in the style sheet) to foul the translation presentation, so as to change the color of parts of the text (e.g., white text on white background).

Immediately below is an actual example taken from the internet (at the time of this research was at: <u>http://glovia.fujitsu.com/jp/event/kansai/02sfl113.html</u>). The "Japanese View" is easily visible text in the middle of the frame. The translated "English View" has that text translated, but it is presented in the identical color (white) as the background, hence is invisible. If you highlight the "white text" area (using Control A in Internet Explorer), you would be able to see the text we have presented in the "Exposed View".

Japanese "friend" View	English "foe" View	Exposed View
共催: FUJITSU ファミリ会	共 催: FUJITSUファミリ会	共 催: FUJITSU ファミリ会
猛暑の今年7月、東京国際フォーラムにて フォーラム2002」がいよいよ満を持して大阪(ト時代に向けた富士通の先進的取り組みや. ます。		* The July this year of severe heat, at Tok comment, ~ the Fujitsu solution forum 2002 * It introduces the place where the advanc which is directed to broadband * Internet a
本フォーラムでは、GLOVIAソリューション関;	With this forum, we exhibit GLOVIA solutior	With this forum, we exhibit GLOVIA solutior
GLOVIA-C Ritz	GLOVIA-C Mate:	GLOVIA-C 開建セミン

In this case, the perpetrator testified that no promotion of a subject had occurred, yet, on the internet, in Japanese, they had delivered essential information which the translation engine translated as: "The July this year of severe heat, at Tokyo international forum it received favorable comment..." in direct contradiction to their testimony. However, English readers could not see this translated statement since the text's background was the same color as the text.

There are a three courses to thwart this ploy: a) Read the page in Japanese, b) Have the translation engine "anonymously" fetch the page, and c) search for "font ... color=*same*" where "*same*" is a color the same as the background, in this case white.

2. Using image files to hide text: This is a way to put textual content in view without search engines being able to see the content. Therefore a foe cannot find the file, and a translation

engine cannot translate the file. (The search engine can translate encoded text, but not images).

If you examine the figure below, you can make out five occurrences of the word "EXCEL". By putting these occurrences into a graphics figure the perpetrator prevents their detection by all search engines. In fact, five detected occurrences would, generally, cause the search engine to score the page much higher thusly increasing the likelihood of the page being accessed. Therefore, this is a simple and effective method for operating in a low profile.

The Hidden "EXCEL"			
データ活用 含計残高試算表をはじめキャッシュ・プロー計算書や 経営分析グラブをタイレクトにEXCEL出力。EXCEL 機能技利用した簡単スピーティな二次加工をご支援。 ~管理会計:情報を的確に把握~ 経営分析 予算情報を表が細か(部門別・物定料目別/月 別に管理。当切予算はたとれ)見声子算も含めた	ー>理未務2×0++++>>>=用できます。 <u> 社員情報の活用</u> さいい情報をEXCELに取り出せば、新規採用時 の登録もスムーズに。 さらに、社員名薄や住所録の作成に活用して事 務効率を大幅アップ。 <u> EXCELシート入力</u> 動意の入力シートをSXCELに駆開し、各部署に 配布可能、収集データは一括登録でき、総務部 門の入力負荷を軽減します。		

Using this device in a friendly domain, you could be given information or instructions detrimental to foes. The information could be the "keys" to a web location or the instructions to perform an operation. This is a simple, well known, and well understood device, which is, nonetheless, very general purpose and effective for going undetected (by eluding search engines and defying translation engines).

There is no known effective way to harvest all these occurrences, character recognition programs will work in the simple instances (such as the upper right panel above), but not where there is a distracting background (such as the lower left panel above). There is a trick, though, and that is to search on the content of the URL as some will foolishly put clues to the hidden information there, such as "excel seminar.htm" or "seminar date.jpg".

3. **Hiding text in cursors**: This device puts messages into the cursor tag hidden to: a) translation engines, b) search engines, c) casual viewers, and d) those who don't place their cursor in the specific place. One can also see how this device used with the text hiding (#1, above) can effectively thwart the unintended viewer, that is, the "foe".

Below, there is a long message in the original "friend view" which is scrambled by the translation engine effectively hiding the message from the viewer. By going into the actual, original HTML code, we extracted the Japanese message and then had just that message translated. That's how we were able to expose the contents of the hidden cursor text.

Japanese "friend" View	English "foe" View	Exposed View	
会 場:マイドームおおさか 日 時:2002年11月13日(水) 入均():入場照料(セミナーは事前 会場:東京国際フォーラム(有楽町)日時 入場料:無料(セミナーは事前登録制)3	会場:マイトームああさか 日時:2002年11月13日(水)- 入場、入場際目(セミナーは事前) ま 77,7777777777777777777777777777777777	会場:マイドームあおさか 日時:2002年11月13日(水)- 入地 入場開発1(セミナーは電前日 Meeting place: Tokyo international forum (主 2nd (gold) 9:30 - 17:30 admission charges: 3	
共 催: FUJITSUファミリ会 間	共 催: FUJITSU ファミリ会 開記	共 催: FUJITSU ファミリ会 関西	

To thwart this device, it is necessary to view the page in the original language, upgrade translation engines to translate cursor tags (which has not been done to date), or examine the HTML code for the cursor tags and extract them for separate examination.

4. **Blocked or Dead Links**: This device is simple and crude, but effective. What is done is to give the appearance of having a link, but when you click on it, nothing happens. This is most often an innocent file management error, but, as shown below, can seem quite intentional. Usually, we can't show dead links (simply, there's nothing to show, it's just that the link doesn't work).

5.

Suspicious Dead Links		In this instance, there over 8,000 pages devoted to "GLOVIA" and here there are
>GLOVIAトップ	GLOVIA-C	links for eight GLOVIA pages (GLOVIA-BP, GLOVIA-C, GLOVIA/SUMMIT, glovia.com,
>GLOVIA-BPトップ ▶特長	slovia.com slovia.e GLOVIA/Hyperion -GLOVIA/SCP >GLOVIA関連製品 -MyOFFICE), but, when it comes to GLOVIA-Product and GLOVIA-Related product[s], they are uniquely DEAD.
└会計情報 └給与情報		
>新著情報		The "discovery" we wanted was to show that the technology of their product infringed or
>CyberSeminar		our technology, so we believe the produc description was hidden to thwart us.
>GLOVIA-BPIC関する資料 諸求/お問い合わせ		Note: if the background were not left white
≥GLOVIA製品		we would never have noticed this dead link!

Thwarting and detecting this device automatically is easy in one instance and impossible (or seemingly so) in another. In the case where the link is there but it doesn't work (usually because the target URL is not at its specified location), web crawlers regularly report these as "errors". In the instance shown above, where the link is "suspiciously" absent, there is no known automated method that draws attention to the link's deviant property.

6. Miscellaneous Devices:

- a. Using aliases, misspellings, and/or abbreviations in order to prevent search engines from finding the real meaning and reference. Instead of "disclosure" we found a phonetic substitution used, "disc rose". This confuses the translation machine, but doe not confuse the person who reads it phonetically (e.g. interchanging "r" and "l" phonetically is prevalent in Japan).
- b. **Outputting different pages to different viewers** (perhaps based on the URL of the calling page) is a device we have experienced. In this instance, a Japanese page existed in Japan, but was unavailable in USA.
- c. **Interchanging fonts and languages** to confound translation engines and search engines. For example, by interchanging different but legible fonts (UTF, Shift JIS, and EUC), translation machines become confused and output either strange characters or bands of question marks where legible text in the original language existed.
- d. **Gaming Search Engines** is a commonplace marketing strategy with the aim of getting a higher ranking to appear earlier in the list of "hits". However, when hiding, you want the opposite, a lower ranking so that the "hit" will appear later in the list and probably go unnoticed. This is the goal of inserting repetitions into web documents (i.e.., salting) that may be ranked by search engines. Here is an example:

You can see from the translation below that the salting of the page with meaningless "truth" adds no value and is only intended to bury the page at the bottom of any list of inquiry "hits". If the "foe" ever got to the page, below is what the "foe" would see.

e. **Gaming Translation Engines.** You may find translation engines asking if you want to make a suggestion to improve a translation. In this instance, we found an accurate translation in November of 2006 which is as follows:

"市場規模(2~5年):数百億円~数千億円") as: "Market scale (2~5 year): Several 10,000,000,000 Yen ~ several 100,000,000,000 Yen. In dollars, that is about \$90 million to \$900 million

When we revisited this in April 2008, to our amazement Google translation posted an erroneous translation which, given the identical Japanese text, read as follows:

"Market size (two to five years): 100 million yen to several hundreds of billion yen."

Now, this translation erroneous translation is off by two orders of magnitude, not to mention that the "several" modifier, 数, has been ignored, as well. The first number is 100 i.e., 百 followed by, not a million, but a hundred million, 億. The result is that these three characters are "several hundred hundred million" or several ten billion, not the hundred million erroneously reported by Google translation in April. Today, Google, at our prodding, has corrected this and you can test it for yourself. We make no suggestion about who gamed Google's translation engine, but we do note that that is a discovery hacking avenue that those who want one view for friends and another for foes could exploit.

Conclusion:

As the internet becomes the information utility of our age, it becomes increasingly important for discovery in litigation and law enforcement. As we have demonstrated above, it is possible for hackers, organized (corporations or terrorists) or individuals, to exploit deficiencies in internet search and translation tools to present different views for different audiences, friends and foes. We bring these to your attention with the hope that countermeasures will be devised, even if manual as in our case, to protect you from being deceived as foes might want.