# Data Type Based Security Alert Dialogs

**Max-Emanuel Maurer**
max.maurer@ifi.lmu.de

**Alexander De Luca**
alexander.de.luca@ifi.lmu.de

**Heinrich Hussmann**
hussmann@ifi.lmu.de

**University of Munich**
**Media Informatics Group**
Amalienstr. 17
80333 München
Germany


*CHI 2011*, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

## Abstract
Making users aware of insecure situations and behavior while browsing the Internet is a highly discussed and still difficult task. Both, passive and active warnings have their own specific disadvantages. While active warnings interrupt the current task and annoy the user, passive approaches often fail since they go unnoticed. In this work, we present first results of a concept displaying data type based alert dialogs whenever a user enters critical information into an online form. Such contextual dialogs appear right in the users' field of view representing a hybrid approach between active and passive warnings. An initial user study was conducted that showed a significant improvement of security awareness by participants that used the tool.

## Keywords
Security awareness, web browsing, data type based

## ACM Classification Keywords
H5.2 [Information Interfaces and Presentation (e.g. HCI)] User Interfaces – Input devices and strategies, evaluation.

## General Terms
Experimentation, Human Factors, Security

## Introduction

Research throughout the last decade has shown that providing the right level of security awareness while browsing the web is not as easy as putting up an alert box. Although security is one of the most important aspects when doing critical tasks online, it is never the user's primary goal [7].

Raising the users' awareness is usually either done by interrupting the current task in a blocking way, or by offering non-blocking information to the user, which can be consumed when needed. Blocking windows stop the users' current task and force them to decide on an action [4]. In case the interruption occurs too often or by mistake, the user quickly gets habituated to the kind of warning and tries to dismiss it without further notice [1]. Non-blocking alternatives are usually simply overlooked by the users being busy with the primary task [8]. Bearing this in mind, the user should be alerted with the right kind of dialog at the right time.

In this work, we present the concept and evaluation of data type based security warnings. The basic idea is that certain data sent to the Internet are more sensible than others. Raising security awareness is important when entering such crucial information, like credit card numbers. Our approach moves away from warnings for any unencrypted, self-certified or otherwise technically insecure form. Instead it notifies the users when they are about to transmit highly sensitive information. This way, the appearance of warnings is reduced to situations incorporating critical data, thus reducing habituation effects due to unnecessary dialogs.

In related work, a lot of effort is spent on security awareness for SSL certificates and encryption. Such warnings are not well understood by users, which makes them unlikely to be heeded [6]. Although encryption is important when transmitting sensitive data, it does not guarantee that the encrypted website is ingenuous. Other work focuses on attacks independent of encryption – like phishing. It showed that often, countermeasures do not hold when phishing websites have a professional look [3]. Again, data type based warnings have the potential to overcome these weaknesses by moving the user's attention from a technical term "encryption" to something they can understand like "you are about to submit your credit card information, which could be a security threat".

We performed an initial user study with the prototype. The results show that the participants were able to identify significantly more phishing websites using the plug-in than using a standard web browser. We used the phishing scenario to be able to measure quantitative values for the success of our concept although preventing phishing fraud is not the only concern the concept is intended for.

## Threat Model

Transmitting personal data to the Internet can expose it to many potential frauds. Phishing websites intentionally try to gather data by mimicking other websites using a number of different attacks [4]. We assume attacks that could be spotted by a user that carefully checks the native browser warnings. On unencrypted data connections, man-in-the-middle attacks can be used. This can be prevented by checking the encryption status before submitting. We thus assume an attacker that is able to apply any of the previous mentioned attacks.

## Concept and Prototype

To avoid habituation effects, warning messages should appear as infrequent as possible. We thus propose to skip warning messages in the absence of sensitive data. We try to achieve this by taking the data type of the input into account. Analyzing the users' input additionally allows for displaying warnings when critical data is entered.

The prototype is implemented as a Firefox plug-in that monitors user input. As soon as the input of a critical data type is detected, a warning message appears right next to the input field, informing the user that she is about to transmit critical data over the Internet (see figure 1). In its current state, the prototype checks for three different data types: *credit card numbers*, *passwords* and *bank transactions numbers* (or TANs) used to authenticate bank transactions in Europe.

In case one of these three data types is identified, a warning dialog informs the user about it and whether or not the data will be submitted over an encrypted channel. Figure 1 shows an example for a credit card number dialog on an SSL-secured connection with the domain name and the encryption status of the site. The dialog offers two possible input options: dismissing it using a little x icon at the top right or pressing the "Trust this!" button in the lower right. The user interface is a first version that was implemented for the primary study. Its design is based upon preliminary analyses for such a user interface but was neither evaluated nor optimized yet. This will be future work.

When the user "trusts" the input, the URL and the data type are stored. In future use, no more warnings for this particular type of data will appear on that very same website. This minimizes habituation and can also be used to identify malicious websites in case the dialog reappears on a website looking like a trusted one.

The warning dialog is placed just below the current form element and the user can continue typing. This way, the user can decide ad-hoc whether to firstly finish the input and react later or to take action immediately. The dialog can also be dragged to uncover parts of the form that may have been blocked. We call this a "semi-blocking" dialog. "Semi-blocking" refers to the fact that the dialog does not fully interrupt the users interaction – typing is still possible – but it is displayed right in the users view, partially covering a currently important portion of the screen.


**Figure 1.** Data type based confirmation dialog on an encrypted website that informs the user that important information (credit card number) will be transmitted over a secure channel. [staged for printing purposes]

Summed up, analyzing user input on forms has several advantages: The number of unnecessary warnings is reduced. Warnings appear in the users' focus at the place where that data is entered. This makes it very likely that the dialog will be noticed. For each website and data type, the warning message appears only once. This can implicitly indicate malicious websites.

## First Evaluation

We tested if the prototype raised security awareness for phishing attacks in a lab study. During the study the participants were asked to perform tasks for their grandmother. The scenario was thus partially based on the study design used by Wu et al. [8]. The purpose of the study was not revealed until the debriefing and instead presented as being about "Internet behavior".

*User Study Design*

The study was carried out using a mixed-model design, dividing subjects into two groups. The between-subject variable was *plug-in* (yes or no). That is, we had one control and one experimental group. Within the groups, a repeated measures design was used with the independent variable *data type* (credit card, password and TAN). The dependent variable was *achieved security*: correctly identified phishing websites and the number of false positives (genuine sites accidentally nominated as a phishing website).

The participants performed six tasks on six websites. For each data type, two websites were used, one of them being a (simulated) phishing site. The two tasks for each data type were similar but slightly adapted with respect to the scenario. We used two shopping websites for the credit card number scenario, two banking websites for TAN and an online community and a webmail provider for password. All brands used were well known to the participants.

During the study, all network traffic was diverted to a local server hosting copies of the original sites and the modified phishing websites. We used common phishing attacks to create those sites – e.g. similar name or IP address attack. It was impossible for participants to notice the fact that the network traffic was diverted – e.g. URLs still were the same. Every website seemed to come from its original location.

To minimize learning effects, a new bookmark set was used for each user of both groups. For each data type one of the two websites was randomly assigned to be the phishing one. Twelve study settings were derived using a 6x6 Latin square two times inverting the phishing sites in the second set. In total, 24 (2 groups x 12 sets) participants were required.

*Procedure*

At the beginning of the experiment, the participants were told that they should do some online transactions for their "naïve" grandmother who had to go to hospital. The grandmother's account data was written down in her "secret book". We used this scenario to avoid participants getting unmotivated because of role-playing someone else, whilst still protecting their privacy as it would be a problem with real data. This approach is similar to what Wu et al. did in 2006 [8].

The participants received the "handwritten" tasks one by one. The URLs were available as bookmarks in the browser. This way, we could make sure that all participants would visit the site. They were told to "think aloud" during the tasks and whenever they

mentioned any concern on the currently displayed web page, the experimenter told them that they would be allowed to skip this task if they thought it could have bad consequences for their grandmother. Aborting a task (and only this) was counted as having detected a phishing website. At the end of the study participants were debriefed and filled out a final questionnaire.

*Hypotheses*

For our experiment we stated one main hypothesis:
**H1** Participants of the plug-in group will recognize more fraudulent web sites than users in the control group.

*Participants*

We recruited 24 participants for the study, mostly students. They were randomly assigned to the two groups (12 per group). Participants of the 'plug-in group' had an average age of 24 years, 3 being female. Participants of the control group had an average age of 23 years with 4 being female.

*Results*

The results are based on identified phishing websites and false positives. Moreover, the 24 questionnaires filled in by the participants provided valuable qualitative feedback and information on user satisfaction.

PHISHING

Participants of the plug-in group were able to discover 20 of the 36 phishing websites (55.6%). Control group users only identified 5 of 36 (13.9%). A two-way mixed ANOVA revealed a significant main effect of *plug-in* ($F_{1,22}$ = 11.73, $p$<.05). This shows that the plug-in did help people to discover more phishing websites and confirms H1.

No significant main effect for *data type* was found ($F_{2,44}$=0.77, $p$>.05). There was an interaction effect of plug-in x data type ($F_{2,44}$ = 6.27, $p$= .004).Looking at this in detail, plug-in in combination with the data types credit card and password showed no significant changes in recognition ($F_{1,22}$=0.20, $p$>.05). In contrast, the difference between TAN and passwords was significant ($F_{1,22}$=6.50, $p$<.05). This difference is based on the fact that for transaction numbers an equal number of four phishing websites were found for both groups.

FALSE POSITIVES

Since the plug-in was designed to increase security awareness, a potential problem was accidentally aborted tasks on genuine websites. Whilst only phishing sites were skipped by the participants of the control group, two plug-in group subjects refused entering credentials on a real website. Thus, two of the 36 genuine websites were mistakenly reported as fraud. This shows that there is some danger to create false positives. The probability for false positives is rather small and can be further reduced as described below.

QUALITATIVE DATA

After being debriefed, the plug-in group was asked how helpful the concept was using a Likert scale from 1-'not helpful at all' to 5-'very helpful'. This was rated with a median of 4. An explanation for this could potentially be found in their answers on what they disliked about the plug-in. Some people were irritated by the high number of notifications in the study. In a real scenario, the number of those dialogs would decrease quickly as they would not reappear on already trusted websites.

*Discussion*

We were able to show that the concept of displaying warnings together with certain data types increases the users' awareness at the right time. The concept is not able to distinguish genuine from malicious websites. Especially when using the plug-in for the first time, many warnings appear. Participants who mentioned this effect recommended delivering such a plug-in with a predefined set of positive records for well-known websites. Participants were still able to identify phishing websites correctly despite dialogs being shown with every of the six websites. In our opinion the small number of only two false positives shows that the plugin did not influence people to stop using standard websites. Sadly we have no data on why people aborted the task in those situations. Details about how the number of appearing dialogs will decrease will be tested in a future field study.

## Future Work

To get more details on these promising ideas, the future work will look as follows. Firstly, the interface of the dialog will be evaluated and discussed using a focus group and related work on warning dialog design. The big "trust this" button suggests accepting this dialog without paying attention to its contents and the information about encryption needs to be made better understandable by inexperienced users.

As a second step, the plugin will be rolled out for a field study to test other variables. For such a concept being successful, we must know how often the dialog will appear throughout browser usage and how fast the number of appearing dialogs will decline using a dynamic white list.

## Conclusions

We presented a new concept for raising security awareness by displaying dialogs only for certain data types right in the user's focus. Participants of a lab study discovered significantly more phishing websites using a browser equipped with this concept. No significant number of false positives was created using the plug-in. We showed that displaying special kinds of warning messages in a browser, depending on what data the user enters, can help to raise security awareness at the right time to the right extent.

## REFERENCES


[1] Amer T.S., Maris J.B. Signal words and signal icons in application control and information technology exception messages–hazard matching and habituation effects. In Journal of Information Systems, 21, 2006.

[2] Close T. Web security experience, indicators and trust: scope and use cases. W3C Working Group Note. 2008.

[3] Dhamija R., Tygar J.D., Hearst M. Why phishing works. In Proc. CHI 2006. Montréal, Québec, Canada.

[4] Egelman S., Cranor L.F., und Hong J. "You've been warned": an empirical study of the effectiveness of web browser phishing warnings. In Proc. CHI 2008. Florence, Italy.

[5] Miller R.C., Wu M. Fighting phishing at the user interface. In Security and Usability: Designing Secure Systems That People Can Use, (2005).

[6] Sunshine J., Egelman S., Almuhimedi H., Atri N., und Cranor L.F., "Crying Wolf": An empirical study of SSL warning effectiveness. In Proc. Usenix 2008.

[7] Whitten A., Tygar J.D. Why johnny can't encrypt: A usability evaluation of PGP 5.0. In Proc. Usenix 1999.

[8] Wu M., Miller R.C., Garfinkel S.L. Do security toolbars actually prevent phishing attacks? In Proc. CHI 2006. Montréal, Québec, Canada.