**Phone Number and Email Address Extractor**

Say you have the boring task of finding every phone number and email

address in a long web page or document. If you manually scroll through

the page, you might end up searching for a long time. But if you had a program

that could search the text in your clipboard for phone numbers and

email addresses, you could simply press ctrl-A to select all the text, press

ctrl-C to copy it to the clipboard, and then run your program. It could

replace the text on the clipboard with just the phone numbers and email

addresses it finds.

Whenever you’re tackling a new project, it can be tempting to dive right

into writing code. But more often than not, it’s best to take a step back and

consider the bigger picture. I recommend first drawing up a high-level plan

for what your program needs to do. Don’t think about the actual code yet—

you can worry about that later. Right now, stick to broad strokes.

For example, your phone and email address extractor will need to do

the following:

Get the text off the clipboard.

Find all phone numbers and email addresses in the text.

Paste them onto the clipboard.

Now you can start thinking about how this might work in code. The

code will need to do the following:

Use the pyperclip module to copy and paste strings.

Create two regexes, one for matching phone numbers and the other for

matching email addresses.

Find all matches, not just the first match, of both regexes.

Neatly format the matched strings into a single string to paste.

Display some kind of message if no matches were found in the text.

This list is like a road map for the project. As you write the code, you

can focus on each of these steps separately. Each step is manageable

and expressed in terms of things you already know how to do in Python.

***Ideas for Similar Programs***

Identifying patterns of text (and possibly substituting them with the sub()

method) has many different potential applications.

Find website URLs that begin with *http://* or *https://*.

Clean up dates in different date formats (such as 3/14/2015, 03-14-2015,

and 2015/3/14) by replacing them with dates in a single, standard format.

Remove sensitive information such as Social Security or credit card

numbers.

Find common typos such as multiple spaces between words, accidentally

accidentally repeated words, or multiple exclamation marks at the

end of sentences. Those are annoying!!