



# NCMEC PREDICTION LOCATOR USABILITY TESTING RESULTS

Prototype v8 – 6/25/2017

Presented by: Shari Benko, UX Designer

# TEST STRUCTURE

- Users shared their screens and were given a URL to v8 of the working prototype.
- **2 questions & 8 tasks** were given to **5 participants (2 inexperienced NCMEC analysts and 3 experienced)**
- SUS survey given to each participant at the end of the tests

Report ID	Urgency	Predicted Location	Confidence	ESP	IP Organization(s)	IP Countries	Phone	Needs Review
2	Critical	Netherlands	83%	Facebook	DigitalOcean, UK2 - Ltd	Germany, Netherlands	91 172	Review/Change Approve
2	Critical	Egypt	76%	Facebook	TE Data, Earthlink Iraq	Egypt, Iraq		Review/Change Approve
2		United Arab Emirates	88%	Facebook	Umsida, Emirates Telecommunications Corporation	United Kingdom, United Arab Emirates		Review/Change Approve
2		United Kingdom	42%	Facebook	Falco Networks, Digital Ocean	United Kingdom, United States, Germany	91 55	Review/Change Approve
2		Guatemala	99%	Facebook	Telgua	Guatemala, Colombia		Review/Change Approve
2		India	96%	Facebook	Jio	India, Pakistan	91 45	Review/Change Approve
2		Iraq	98%	Facebook	Alсандfiber	Iraq, Iran		Review/Change Approve
2	Critical	Iraq	75%	Facebook	SatGate-IQ-ISP, TIGO El Salvado	Iraq, El Salvador		Review/Change Approve
2		Iraq	83%	Facebook	Newence Telecom Ltd., Total Server Solutions L.L.C.	Iraq, United States		Review/Change Approve

**NCMEC**

**Location Prediction**

**Prototype v8**

**Usability Testing 6/14 - 6/20/2017**

# QUESTION & TASK ANALYSIS

Completion rates and times

# QUESTION #1

What is your first impression of this queue? What is its purpose?

- **Reasoning:** This is to gain qualitative feedback on first entrance into queue.
- **Prediction:** There is no predicted value for qualitative feedback.

## Results:

- Subject 1 (inexperienced analyst) was confused and overwhelmed at first
- Subjects 2-5 jumped to answering question #2

## Conclusions:

- The queue could seem overwhelming to inexperienced analysts, but wasn't troubling to users who were comfortable with their current workflow

# QUESTION #2

## Which reports would you start with first and why?

- **Reasoning:** To discover which items a user will tackle first.
- **Prediction:** Each analyst will have their own way of choosing which reports to start with first.

### Results:

- 3 out of 5 started with either high or low confidence reports
- 2 out of 5 started with Critical urgency, then moved to confidence levels

### Conclusions:

- Confidence was clear to the end users as the place to start urgency was a close 2nd

# TASK #1

## How would you forward a report to the predicted jurisdiction?

- **Reasoning:** To assess how intuitive the “Approve” button is to find and use.
- **Prediction:** User should indicate that “Approve” is the correct button to forward a report.

### Results:

- **Task Completion Rate: 80%**
- **Task Average Completion Time: 5.25 seconds (KLM = X)**

### Conclusions:

- The approve button was clear to most users as the way to forward a report from this queue

# TASK #2

Which reports do you feel that you need to review before forwarding to the predicted jurisdiction?

- **Reasoning:** To assess whether or not the confidence levels are what indicate to the user that the report needs to be reviewed, and to assess the content presented in the unexpanded view.
- **Prediction:** The user should indicate reports that need to be reviewed based on the confidence level, or content presented in the unexpanded view.

## Results:

- Users were presented with dummy information which made the 2<sup>nd</sup> prediction impossible to evaluate quantitatively
- 2 of the users concluded that the low confidence was the indication

## Conclusions:

- This task is inconclusive and no conclusions can be made

# TASK #3

## How would you save a report for review at a later time?

- **Reasoning:** To find out how quickly a user understands that they can flag a report for review and it drops the report to the bottom of the list.
- **Prediction:** The user should indicate or click on the flag, to flag a report for review at a later time.

### Results:

- **Task Completion Rate: 20%**
- **Task Average Completion Time: 33 seconds**

### Conclusions:

- The review flag button is not obvious enough to users

# TASK #4

Which reports do you think have an accurate enough prediction to forward without review?

- **Reasoning:** To assess how intuitive the wording “confidence” is to an analyst, and how much they trust the system’s predictions.
- **Prediction:** User should indicate that the high confidence reports are accurate enough to forward.

## Results:

- 2 out of 4 users indicated that high confidence
- 3 users indicated they would trust the system over time

## Conclusions:

- While completion rate and times couldn’t be calculated due to dummy information, the users indicated they would trust the system over time

# TASK #5

How would you forward multiple reports at the same time, that don't need review?

- **Reasoning:** To assess how quickly it is for a user to filter and approve multiple reports at once.
- **Prediction:** User will filter for items that indicate an accurate prediction, then check select all either from the action menu or from the table header, then select the button for "Approval All Selected."

## Results:

- **Task Completion Rate: 100%**
- **Task Average Completion Time: 15 seconds**
- **1 out of 5 users used filters to complete the task**

## Conclusions:

- All users completed the approve all task using different methods, but were successful indicating the quick actions is intuitive to new users
- Filters were not an obvious interactions to users

# TASK #6

How would you go about reviewing reports that do need review before forwarding?

- **Reasoning:** To assess whether or not the expand/close interactions are easy to find.
- **Prediction:** The user should click the “Review/Change” down arrows to expand the report to start a review.

## Results:

- **Task Completion Rate: 40%**
- **Task Average Completion Time: 5 seconds**
- **Subject 1 eventually clicked review after being prompted a 2<sup>nd</sup> time**
- **Subject 2 attempted to use unexpanded info to review**
- **Subject 5 clicked through to external report to review**

## Conclusions:

- Since 3 out of 5 eventually reached the expanded view – but 2 used info on screen or external, this functionality while somewhat intuitive needs training

# TASK #7

## How would you change the jurisdiction of a report?

- **Reasoning:** To discover how intuitive it is to change the location of a report.
- **Prediction:** The user will click the “Review/Change” expander, select the drop down box for change location, select a new location and interact with the modal window.

### Results:

- **Task Completion Rate: 100%**
- **Task Average Completion Time: 10.4 seconds**
- **All subjects understood the purpose of the feedback box**

### Conclusions:

- Changing a location of a report is easy for users to find
- The feedback modal window is easy for users to understand

# TASK #8

Your co-worker Casey Blake is out sick today and you have been given her reports to handle. How would you start working on his assigned reports?

- **Reasoning:** To assess how easy it is for a user to work on a co-workers assigned reports.
- **Prediction:** User should select the filter for Casey Blake

## Results:

- **Task Completion Rate: 80%**
- **Task Average Completion Time: 6.25 seconds**
- **Subject 5 looked for an export reports feature**

## Conclusions:

- **Most users had no trouble reaching a co-workers reports**

# OVERALL COMPLETION RATE AND TIME

**Average completion rate: 70%**

**Average completion time: 10.5 seconds**

**What rate is a good rate?** An analysis of almost 1200 usability tasks shows that the average task-completion rate is 78%. <https://measuringu.com/sus/>

# SYSTEM USABILITY SCALE SURVEY

Each test subject was given a link to a SUS survey to complete after the user testing session. The following cases were given to each subject along with a scale ranking and each users answers are then calculated into an overall score:

- I think that I would like to use NCMEC Prediction Locator frequently.
- I found NCMEC Prediction Locator unnecessarily complex.
- I thought NCMEC Prediction Locator was easy to use.
- I think that I would need the support of a technical person to be able to use NCMEC Prediction Locator.
- I found the various functions in NCMEC Prediction Locator were well integrated.
- I thought there was too much inconsistency in NCMEC Prediction Locator.
- I would imagine that most people would learn to use NCMEC Prediction Locator very quickly.
- I found NCMEC Prediction Locator very cumbersome (awkward) to use.
- I felt very confident using NCMEC Prediction Locator.
- I needed to learn a lot of things before I could get going with NCMEC Prediction Locator.

For more information on SUS scores, visit: <https://measuringu.com/sus/>

# SYSTEM USABILITY SCALE SURVEY RESULTS

SUS score: **81.3** (percentile: 91%, which corresponds to the following letter grade: B-)

Learnability: **84.4**

Usability: **80.5**

Standard deviation: **11.3**

Cronbach's alpha: **0.84** (above 0.7 is acceptable)

**What score is a good score?** A SUS score above a 68 would be considered above average and anything below 68 is below average. <https://measuringu.com/sus/>

# CONCLUSIONS AND RECOMMENDATIONS

What do we do based on this test?

# CONCLUSIONS & RECOMMENDATIONS

The queue could seem overwhelming to inexperienced analysts, but wasn't troubling to users who were comfortable with their current workflow

**Recommendations:** No changes needed

Confidence was clear to the end users as the place to start urgency was a close 2<sup>nd</sup>

**Recommendations:** No changes needed

The approve button was clear to most users as the way to forward a report from this queue

**Recommendations:** No changes needed

# CONCLUSIONS & RECOMMENDATIONS

The review flag button is not obvious enough to users

**Recommendations:** rethink the necessity of this interaction

Users indicated they would trust the system over time

**Recommendations:** no changes needed

All users completed the approve all task using different methods, but were successful indicating the quick actions is intuitive to new users

**Recommendations:** no changes needed

A user was misled by the check icon on the Approve button

**Recommendations:** remove the check icon from the approve buttons

# CONCLUSIONS & RECOMMENDATIONS

Filters were not an obvious interactions to users

**Recommendations:** No changes needed, this interaction will be learned through use of the system and training on the system

Change/Review expand interaction required more time to learn

**Recommendations:** No changes needed, this interaction will be learned through use of the system and training on the system

Changing a location of a report is easy for users to find

**Recommendations:** No changes needed

# CONCLUSIONS & RECOMMENDATIONS

The feedback modal window is easy for users to understand

**Recommendations:** No changes needed

Most users had no trouble reaching a co-workers reports

**Recommendations:** No changes needed

# FINAL RECOMMENDATIONS

- Rethink the necessity of the save and review later flag interaction

Discover whether or not this is an interaction that only a small fraction of people will use, or include a “Save for Later” interaction in the actions menu removing the flag option.

- Remove the check icon from the approve buttons

# ESTIMATED COMPLETION RATE AFTER CHANGES

This estimate is calculated on the following assumptions:

- The completion rate for the “save for review later” feature is at least 80% once changes have been made.
- With a small amount of training which will take place due to the extensive training conducted by NCMEC for analysts, using the filters and using the expanded view for review should both increase to at least 80%.

**Estimated completion rate: 86.6%**

**What rate is a good rate?** An analysis of almost 1200 usability tasks shows that the average task-completion rate is 78%. <https://measuringu.com/sus/>

