# Commanding at Scale – Email commands

May 2020

# This document outlines the plan for enabling Assistant-triggered emailing from Outlook Mobile for enterprise employees with their primary work account.

# 1. Problem and Hypothesis

## Problem: Core problem and hypothesis as stated in the master Github Issue/spec:

**Problem**: Email remains the most common method of communication in the modern workplace (72%). The number of emails an IW gets has increased constantly and is expected to continue to increase at a compound annual rate of 15%.

* An IW receives an average of 120 new emails every day and spends 2.6 hours per day.
* Keeping up with email is a source of stress. Getting to “zero inbox” is a “relief” for those few (and declining) who achieve it or considered “impossible” by those who cannot.
* With the rise of mobile devices, 60-70% of the email is read and triaged on mobile devices but more complex emails and tasks related to email are left for more robust devices and form factors such as PC. This is possibly a GUI-driven problem as the mobile interface is limited in terms of how much and how easy it can do and many of the advanced functionalities are disabled or hidden due to lack of real estate.
* Leaving the advanced email tasks for later without proper triaging and reminding tools leaves the burden of remembering mostly on the user and making them prone to missing important communications. They typically resolve with sending notes to self of responding with short notes and adding themselves to the “To” line.

**Hypothesis**: We believe that by enabling voice and touch capabilities for core user jobs related to the use of email we could save users time and make them more productive.

As the assistant embedded in OM, being able to do key communication tasks via voice and touch is a fundamental expectation users have and therefore we need to deliver to earn their trust and enable them to graduate to more complex tasks

**Value Prop**

**For Users**: The assistant saves me time and makes me more productive so I can focus on things that matter more while making me feel more organized and confident and that I'm on top of my communications

**For Outlook**: We want users to turn to OM as their email app of choice because the Assistant enables them to be more productive, accomplish more, and feel more organized/in control of their email communication.

### **JTBD: I can catch up and communicate with others more easily using voice and touch**

Customer Promises:

* I can easily compose quick emails or notes to others, including myself
* I can easily reply to or forward to others emails and include short messages/notes
  + Reply to message/thread
  + Reply with adding one or more recipient
  + Reply with removing one or more recipient
  + Reply with moving people to cc and bcc line and DLs or teams (stretch)
  + Forward message/thread with adding one or more people
  + Forward with adding people to cc and bcc line (stretch)
* I can easily compose, reply to emails and forward them to others with elaborate email body (will be covered under Dictation spec)
* I can easily refine my new email communications by:
  + Adding one or more people (including self)
  + Adding more recipients to cc and bcc line (stretch)
  + Altering the subject line
  + Add file attachment (with and without disambiguation)
  + Altering the message/body of the email with handoff to dictation when needed (will be covered under Dictation spec)

|  |  |  |
| --- | --- | --- |
| Crawl (Good) | Walk (Better) | Run (Best) |
| * Initiate new email composition using voice * Support voice invocation of a blank email compose screen * Support voice-initiated emails to a single recipient (e.g. Send an email to Jane”) * Support voice-initiated email to single recipient with disambig (e.g. Send email to Jane (actually John)) * Support single shot email to self (e.g. Send me a note to work on the budget) * Support single short email to a single recipient with a message body pre-filled * Support single short email to two recipients with a message body pre-filled * Support voice initiated emails to 2 people, 3 people * Refining email communications   + Adding one or more people (including self)   + Altering the subject line   + Add file attachment (with and without file disambiguation)   + *Add file as a link* * Reply to message/email thread * Forward message/email thread to one recipient * Compose elaborate new message with dictation & multi-modal/touch   \*Green signifies feature is in WIP from FY20 (Nelson’s spec)  \*\* Compose/Dictation to be covered in separate spec  Support pronouns  Signature for Assistant sent emails  Undo – capture  Daisy chain commands  Rank commands by nr of clicks saved  Recovery | * Refine new email by adding more recipients to cc or bcc line * Reply to message adding one or more recipient * Reply to message removing recipients * Forward message to more than one recipient * Reply with elaborate message (dictation & multi-modal/touch) * Forward with elaborate message (dictation & multi-modal/touch)   **Support Attachments:**   * Picture from device * Picture from OneDrive (of whiteboard)   **Support links to internal sites (Bing @Work indexed sites):**   * Support links to internal websites: “Send Jane link to timecard site” * Support links to sharepoints “Send Jane our team sharepoint link”   **Edit email settings such as:**   * Sensitivity * Assign policy * Importance * Signature * Flag for follow up * Encryption * Delay delivery * Categorize with color code | * Compose new message with people in cc and bcc line * Reply to message/thread adding recipients to cc and bcc (*“Reply and move Adam to bcc”)* * Forward with adding people to cc and bcc line   Reply & Forward to DLs & Groups   * Reply to message/thread adding DLs and Groups (*“reply adding my team”, reply adding Core Search + Assistant team)* * Forward to DLs and Groups (*“reply adding my team”, reply adding Core Search + Assistant team)*   **Support outputs for company sites & apps (Bing @Work indexed sites/apps):**   * “Send Jane the Jira roadmap” * “Send Jane the output of last month’s sales numbers” (Power BI) |

## Roadmap for Communication Commands (work in progress)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Rk* | *Product Backlog* | *Define* | *Design* | *API* | *Android* | *iOS* | *Teams* | *OWA* | *Other* |
| *1* | Single shot email commands—ship-learn-iterate | *20’Q3* | *20’Q3* | *20’Q4* | *20’Q4* | *20’Q4* | *--* | *--* | *--* |
| *2* | Refinement of email commands (refine | *20’Q4* | *20’Q4* | *21’Q1* | *21’Q2* | *21’Q2* | *--* | *--* | *--* |
| *3* | Dictation in email with manual transition to/from commanding | *20’Q4* | *20’Q4* | *21’Q1* | *21’Q3* | *21’Q3* | *--* | *--* | *--* |
| *4* | Refinement of email with attachments and links | *20’Q4* | *20’Q4* | *21’Q1* | *21’Q3* | *21’Q3* | *--* | *--* | *--* |
| *5* | Expand single shot email commands with Reply & Forward | *20’Q4* | *20’Q4* | *21’Q1* | *21’Q2* | *21’Q2* | *--* | *--* | *--* |
| *6* | Dictation with seamless transition to/from commanding | *20’Q4* | *20’Q4* | *21’-Q2* | *21’Q4* | *21’Q4* | *--* | *--* | *--* |
| *7* | Smart actions for email in L3 | *21’Q4* | *21’Q4* | *21’Q1* | *21’Q2* | *21’Q2* | *--* | *--* | *--* |
| *8* | Natural language recovery | *20’Q4* | *20’Q4* | *21’Q1* | *21’Q1* | *21’Q1* | *--* | *--* | *--* |
| *9* | Message commanding (Teams) | *21’Q1* | *21’Q1* | *21’Q2* | *21’Q3* | *21’Q3* | *21’Q3* | *--* | *--* |
| *10* | Message – add/remove people | *21’Q1* | *21’Q1* | *21’Q2* | *21’Q3* | *21’Q3* | *21’Q3* | *--* | *--* |
| *11* | Message – send file | *21’Q1* | *21’Q1* | *21’Q2* | *21’Q3* | *21’Q3* | *21’Q3* | *--* | *--* |

# 2. Goals & Constraints

## Goals (to be revised)

Support Crawl scenarios by FY21-Q1

Support Walk scenarios by FY21-Q2 (proposed)

## Constraints

* No email composition within Cortana canvas. Direct hand off to native UI only.
* No voice multi-turn – but note that these are commands that can be used in a chain, one following right after the other.
* No voice output

# 3. Scenario Walkthrough (PM art below, interaction model will follow Design’s explorations):

Invocation: Assumes invocation can be through the microphone placed at L3.

**Scenario 1: Refine with add recipient(s) including self**

Context: User has already started a new email via single shot. Before sending they decide to add more people to the “To” line.

User journey: User touches mic icon à invocates Cortini canvas à says “add Allison and Chad to this email (or add me to this email) à Assistant updates recipients in the “To:” line (no disambiguation needed) à User hits send button

**A screenshot of a cell phone

Description automatically generated**

**Sample commands/utterances:**

* Add Allison, Chad, and myself
* Add Allison, Chad, and Beryl
* Include Allison and Chad

Dependency: Multi person disambig. (Interaction, designs to iterate) [@Rachel Ng](mailto:Rachel.Ng@microsoft.com), need some design explorations for this. Also, update design for the invocation POR

Have intent clear:

* Remove for recipient
* Cancel for email all together

Disambig: Move everything to people picker, Q: how will it impact multi person disambig?

**Scenario 2: Refine subject line**

Context: User has already started a new email via single shot. Before sending they decide to change the subject line.

User journey: User touches mic icon à invocates Cortini canvas à says “change subject line to running late” à Assistant updates subject line à User hits send button

A screenshot of a cell phone

Description automatically generated

Additional scenario: **Add to subject line: Quick note about…. (scope it)**

**Run: Punctuation commands, exclamation, spacing, -- (double dash,**

**Sample commands/utterances:**

* Make topic, subject
* Change topic, subject

**Scenario 3: Refine with add document (no file name disambiguation needed) (Q2)**

Context: User has already started a new email via single shot. Before sending they decide to add a document.

User journey (no file disambiguation): User touches mic icon à invocates Cortini canvas à says “add sales report april deck” à Assistant adds the correct file à User hits send button

A screenshot of a cell phone

Description automatically generated

**Sample commands/utterances:**

* Add <file name>
* Attach <file name>
* Include <file name>

**Sample commands/utterances for links:**

* Add link <file name>
* Attach link <file name>
* Include link <file name>
* Make that a link

**Scenario 4: Refine with add document (with file name disambiguation) as attach or link (Q2)**

**Need to spec for Photos (different buttons) (use last photo taken/chose from library) (use similar language for LU work). Choosing photo from library also local. (attach vs. insert in canvas)**

Context: User has already started a new email via single shot. Before sending they decide to add a document.

User journey (with file disambiguation): User touches mic icon à invocates Cortini canvas à says “add sales report deck” à Assistant shows file search results for two most likely files intended based on invocation à User selects correct file/version à Assistant attaches file à User hits send button

(add the same for link)

A screenshot of a computer

Description automatically generated

**Sample commands/utterances for attach as file:**

* Add <file name>
* Attach <file name>
* Include <file name>

**Sample commands/utterances for links:**

* Add link <file name>
* Attach link <file name>
* Include link <file name>
* Make that a link

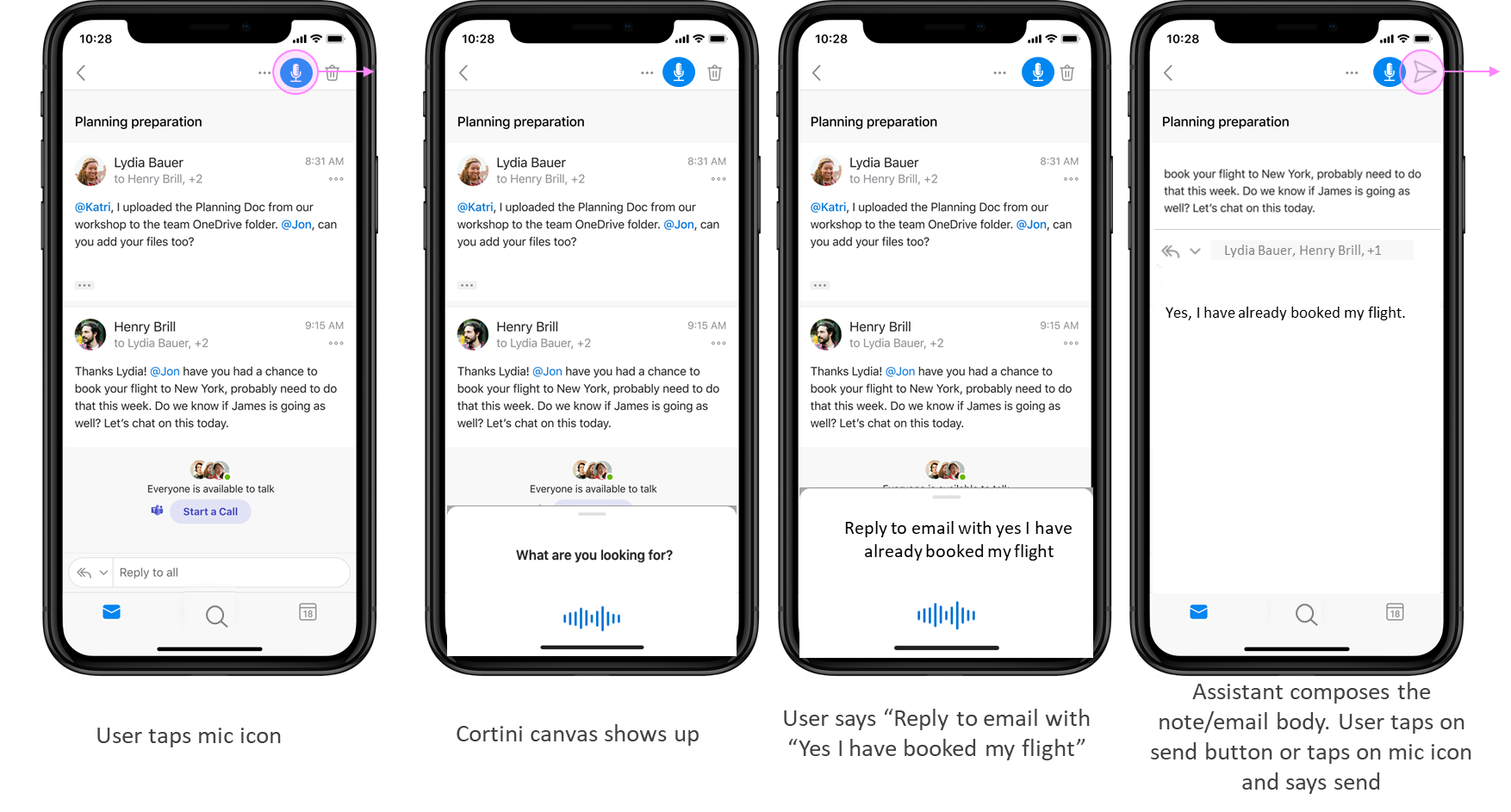
**Scenario 5: Reply to thread with singe (one) shot**

Context: User is on L2 on an email thread. They decide to reply with single shot. Q1 assumes “Reply all” (no removing, narrowing list of recipients)

User journey: User touches mic icon in L2 à invocates Cortini canvas à says “reply to email with yes I have already booked my flight” à In “Reply” mode, Assistant populates the email body message à User hits send button

Reply vs Reply all:

* How do we handle the two?
* Maybe phase Reply all when we have remove



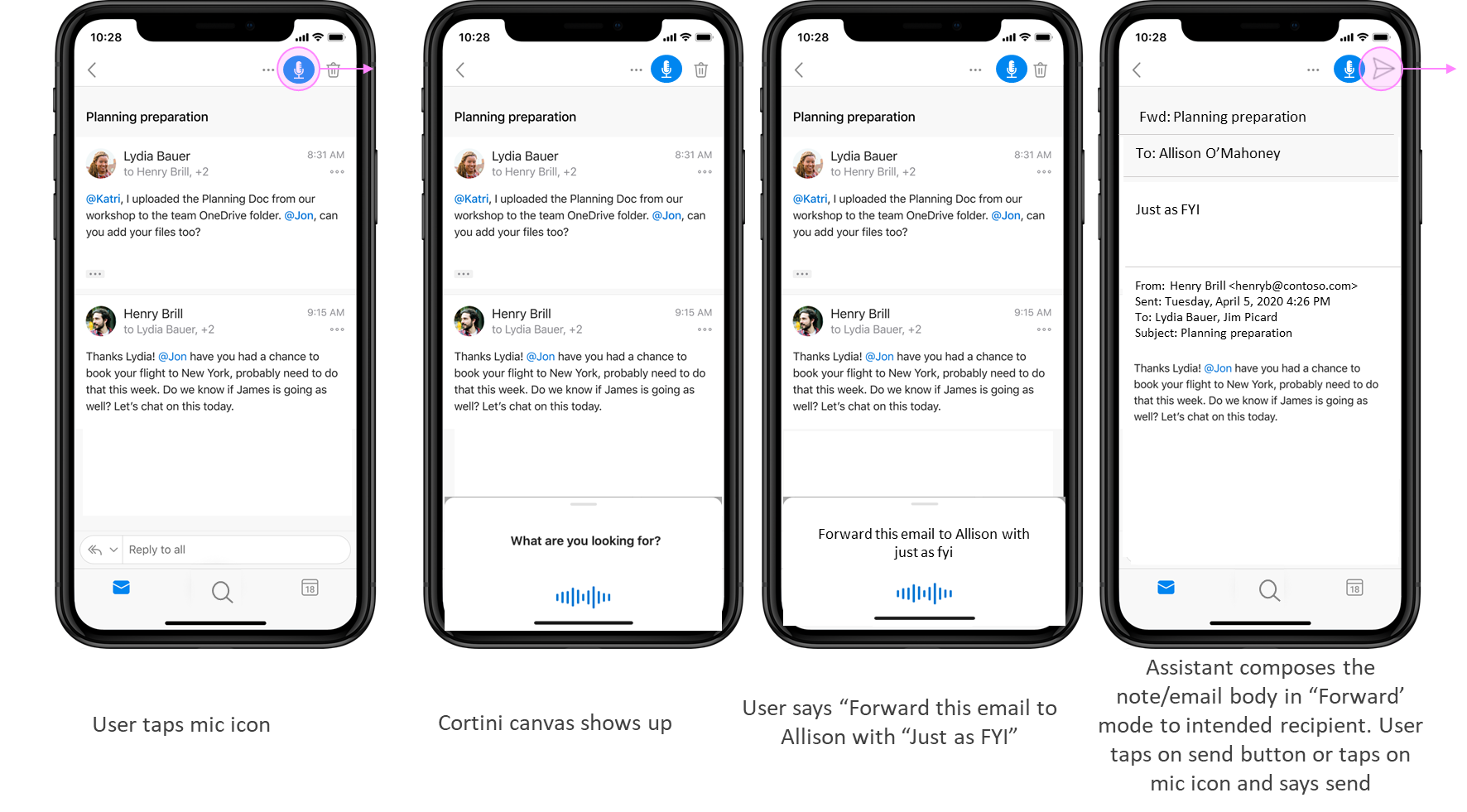
**Sample commands/utterances:**

* Reply
* Reply to email
* Reply to thread
* Reply to Lydia (will default to “Reply all”—user can edit recipients later)
* Reply to Lydia and Henry (will default to “Reply all”—user can edit recipients later)

**Scenario 6: Forward email/thread to single recipient with one shot**

Context: User is onL2 on an email thread and decided to forward it to another colleague via single shot.

User journey: User touches mic icon à invocates Cortini canvas à says “forward to Allison with just as fyi” à In “Forward” mode, Assistant populates the “To:” line and the email body message à User hits send button



**Sample commands/utterances:**

* Forward to <recipient name>
* Forward email <no recipient name> à defaults to FW mode, no recipient in “To:” line
* Send this to Jane (recipient not present in main thread à defaults to FW mode with Jane in “To:” line
* Share this with <recipient name>

# 4. Timeline

Targeting OM beta FY21-Q1 (to be revised)

# 5. Proposed Work & Support Needed

## 5.1 Discoverability and awareness of the skill

|  |  |
| --- | --- |
| Propose Work | Support Needed |
| Add email skills to hints to spur usage. | Need NL capabilities to be enabled, designed, and configured.  POC: |

## 5.2 Functional requirements

|  |  |  |
| --- | --- | --- |
| Use Case | Work Needed | Support Needed |
| Email <contact> <subject slot> <message slot> | 1. Ability to recognize the right intent  2. Ability to recognize the right name  3. Ability to recognize subject and message body within the command | -LU enable the email creation model and analyze the slots -3S return the actionable answer with recommended contact with confidence labels -Client work to hand off to native app for email creation and slot prefill |
| Disambiguation | Build the disambig flow and allow users to self correct & pick the right contact |  |

## 5.3 Accessibility requirements

|  |  |
| --- | --- |
| Proposed Work | Support Needed |
| Pass iOS test with VoiceOver enabled |  |
| Pass Android test with TalkBalk enabled |  |

# 6. Metrics and Telemetry

In order to determine success of the feature as outlined in section 1, we need to gather the following telemetry on usage and understand the user flow dropout rate:

* E2E Task completion rate for one   
  Funnel analysis with task cancellation rate, abandonment rate and success rate
* Steps/clicks/time saved on email task for both touch and speech-initiated emails to measure the efficiency vs. existing email creation flow
  + Start: Mic invocation
  + End: Email is sent
* Number of emails initiated using touch for voice enabled users
* Number of emails initiated using speech for voice enabled users
* Number of commands chained in one session
* Repeated use per user
* DAU/MAU ratio (goal 30% - to be revised)
* Retention curve D1-D2-D7-D30 by cohort

# 7. Open questions

* Work with LU team to enable supporting multiple recipients
* Work with LU team to understand how many ways can a user say “note to self”? Will LU handle all these variations ranging from “Me”, “”, to “remind me”.

# 8. Resources

[S+A Github issue 31](https://ghe-us.microsoft.com/Fast/Search-Assistant-Product-Design/issues/31)