Talking to Bots: Estimating the Environmental Impact of Human-Computer Interaction Methods

Peter Bastiaan den Boer
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
p.b.denboer@student.utwente.nl

ABSTRACT
Energy consumption in the ICT industry has been steadily increasing at an unsustainable rate, though the effect of innovative IT solutions, such as superapplications, virtual assistants and interactive chatbots, remain unknown. Therefore, based on previous studies in the fields of environmental informatics and human-computer interaction, this paper proposes a measuring method to determine the effect of these human-computer interaction methods on ICT power consumption through network package size analysis of response content types. The results imply that as much as 99.69% of energy could be conserved by preferring plain text conversational interaction over full-sizes web applications. However, reservations are made regarding the attainability of such change within the industry as not all interaction methods are suitable for each IT product or service. Yet, the results of this research can prove useful to governments and businesses wanting to adapt their IT products and services to favor a more environmentally sustainable solution. Furthermore, as there is little existing research on this topic, this paper provides a starting point for future research in this field. Similarly, it allows for the analysis of singular products and services through case studies applying the energy analysis method put forward in this paper.

Keywords
Cloud computing, energy efficient computing and networking, sustainable human-computer interaction, environmental informatics, ICT power consumption, green computing

1. INTRODUCTION
When performing any task on any kind of electronic device, energy is consumed. The same applies to the entire Information and Communication Technology (ICT) industry, however on a much larger scale. For the last few years, energy consumption in the ICT industry has been steadily increasing at an unsustainable rate [11]. It has been projected that ICT will have doubled its share of the world’s electricity consumption to 14.57% by 2020, which Vereecken also argues “[...] indicates that this growth scenario is not sustainable” [29]. Although different studies have produced varying estimates of the total energy consumption by the ICT sector [1, 6, 24, 29], it remains clear that there is a necessity to reduce the amount of energy consumed by ICT products and services in order to limit their environmental impact.

Data from various studies [17, 23] suggest that we should expect to see the largest increase in energy consumption from data centers and networking technologies. This is not surprising considering cloud computing’s recent attention [4]. Not only does a boost in cloud computing logically require additional storage capacity in data centers, subsequently increasing the data center’s energy consumption, but it also causes an inevitable upsurge of network traffic compared to local computing solutions.

All the while, companies such as Microsoft and Tencent, amongst others, are continuing the research and development of several innovative human-computer interaction (HCI) methods, specifically:

- Interactive chat bots: conversation-based interactive computer agents accepting plain text input
- Virtual assistants: personal digital assistants combining different data sources and programs to provide user-centered functionality
- Superapplications: platform as an application, incorporating third-party applications and services

These innovative HCI methods operate differently, though there are some commonalities to be found. For example, all of the above can communicate using either plain text or some form of markup language, such as the HyperText Markup Language (HTML). In fact, these HCI methods were designed so that they could be integrated effortlessly with one another as demonstrated by the Microsoft Bot Framework [20] as well as the launch of China Southern Airlines [31], Airbnb [12] and The New York Times [26] on the WeChat Open Platform [30]. This topic, including a more elaborate comparison of the HCI methods mentioned, is discussed further in Chapter 3.

The introduction of these HCI methods will undoubtedly influence the energy consumption of the ICT industry, specifically in data centers and due to an increase of overall network traffic. Sadly, there is little research regarding the environmental impact of these HCI methods.

Yet, it is vital that research efforts are made in this field. As of now, it remains unclear how these HCI methods compare in terms of energy efficiency and consequently their impact on global ICT energy consumption. This is a pressing issue as the amount of active (digital) virtual assistants is projected to increase by 462% and 544% for the consumer and enterprise markets respectively between 2015 and 2021 [27]. Moreover, the global chat bot market is expected to see revenue grow by 880% between 2015 and 2024 [25].

Therefore, the primary goal of this research is to provide business leaders and government agencies with valuable
insight in the energy consumption of certain HCI methods, so that it might positively influence their environmental policies. To achieve that goal, some questions need answering. As no studies found have so far discussed the possibility of an influence on ICT energy consumption by HCI methods, the first question to ask is whether such an influence even exists:

- To what extent is the causal relationship between the use of IT services and power consumption affected by HCI methods?

Once the answer to this question had been determined through the analysis of several literary resources (Chapter 2), the research study progressed to the next question:

- What commonalities and differences of the HCI methods discussed in this paper are of importance in assessing their energy consumption?

To accurately assess the energy consumption of HCI methods, the differences and similarities of each had to be explored in more detail. This is discussed in Chapter 3, in which multiple aspects of the HCI methods are brought up and investigated. This analysis is based on information from various sources, such as product websites and relevant literature, but also relies on professional experiences and market observation.

Subsequently, a method to measure HCI energy consumption is developed in Chapter 4 based on the average query size of interactive chat bots estimated through literature review. This method can be used to answer the following question:

- How do the HCI methods mentioned in this paper compare in terms of energy consumption per information request?

We will shortly discuss the implications of these results in Chapter 5 (Results and Discussion), after which the final conclusions of this research are drawn, while also acknowledging the limitations of this work and providing suggestions for future research endeavors in this field.

### 2. The Influence of HCI Methods on ICT Energy Consumption

Although it is evident that there is a causal relationship between the use of IT services (and therefore electronic equipment) and power consumption, most studies have only gone so far as to attribute power consumption values to specific equipment categories, such as PCs, TV sets and servers. Other studies, such as a study published by Van Heddegem [10], attribute ICT power consumption to slightly more abstract categories instead, which are:

- Data center equipment
- Networking technology
- End-user equipment

To determine whether certain human computer interaction methods could possibly have a measurable effect on ICT energy consumption, the effects of HCI methods on each of these categories must be considered carefully.

#### 2.1 Data center energy consumption

Let us start by looking at the energy efficiency of data centers. Although it is somewhat affected by the energy usage of the server units – and thus the software installed –, it is important to note that most of the energy consumption in data centers is actually a result of cooling and power distribution [8].

When it comes down to the power consumed by servers, there is an important distinction to be made: On the one hand, there is the hardware: facilitating the operation and network connectivity of different software packages and operating systems. As each type of server component has its own energy efficiency, the components installed subsequently determine whether a server is power efficient. Furthermore, the amount of heat generated by each of these components has a considerable influence on how much cooling is needed.

On the other hand, software and operating systems installed on machines also have a noteworthy influence on overall power consumption in data centers. Not only does scheduling of server resources have a considerable influence on power usage [24], but so does the efficiency and complexity of individual software packages. Van Heddegem mentions this as part of an argument centered around the energy efficiency of end-user operating systems such as Windows [10], however the same concept applies to server-side software packages.

Based on this information, it is possible to argue that most of the energy consumption in data centers results from hardware components as well as supporting software and systems. However, the effect of human-computer interaction methods on data center power usage remains undefined.

Regarding the size of data centers, Trotman and Zhang mention the following: “[…] it is the user behavior measured in number of queries that determine the size of the data centre, not the number of users” [28]. If applied to the topic of human computer interaction, one could argue that the type of HCI method does not determine the energy consumption in data centers. Instead, the number of queries generated by users through these interaction methods defines energy usage. Considering that the human-computer interaction methods are quite similar (as seen in Chapter 3) and therefore produce a similar number of queries, there should be no noticeable difference in data center power consumption between different HCI methods.

Thus, preferring the use of a certain HCI method over another will have a negligible effect on the overall power efficiency of a data center, provided the same hardware is used and the number of queries is similar.

#### 2.2 Networking technology power usage

It thus becomes evident that the number of queries generated by using certain HCI methods can have a considerable effect on the size of data centers and consequently their energy consumption. However, it should be obvious that the number of queries generated by HCI methods not only affects the size of data centers, it also determines the amount of network traffic generated as these queries need to be passed along the nodes in the network. It is therefore important to make a distinction between several kinds of networks.

While it might seem to end-users as if there is just one network, it is in fact comprised of numerous networks [3, 16]:

- Access network (e.g. 4G, Wi-Fi, Ethernet)
- Edge network (e.g. edge routers)
- Core network (e.g. core routers, optical cables)

All of these networks do not necessarily make use of the same equipment and therefore have energy consumption values attributed to them depending on the equipment used [3].

As we have determined that the number of queries will be similar between the different HCI methods, it becomes critical to understand how query size affects the amount of energy consumed by networking equipment.
Data transferred over the network is comprised of bits. It is possible to estimate the total value of energy consumed by each single bit travelling through a network by using the combined energy-per-bit values of equipment in access, edge and core networks [3]. Like other IT products or services, the human computer interaction methods discussed in this paper are not tied to implementing a single standard of cross-network communication. It is therefore possible that these HCI methods will yield diverging energy consumption values if they rely heavily on a certain communication standard and data type.

Therefore, evaluating query response size is essential to estimating the energy consumption of specific HCI methods.

2.3 End-user equipment
At this point, only the potential effect of end-user equipment on ICT energy consumption remains unknown. Yet, it could be argued that the energy consumption of end-user equipment is just as much affected by component and software efficiency as data center equipment. However, recent advancements in voice recognition and mixed reality technology have made it possible for manufacturers of end-user equipment to select from a more diverse set of components to create internet-connected devices. Consequently, the energy efficiency of HCI methods has become increasingly dependent on which device is used.

This is not completely accurate however, because of two key points. First, although the interfaces of certain devices might be disparate at times, the input data required to receive a response from the application is similar. This is demonstrated by virtual assistants, such as Microsoft Cortana [5] and the Google Assistant [19]. Users wanting to retrieve data through these virtual assistants can choose to either type or use voice-dictation to submit their query. Once the user’s query has been constructed, it is transmitted in the same fashion, no matter the device type, effectively generating a similar amount of network traffic. Second, as these HCI methods are expected to deliver the same information, albeit through different methods, the response will be similar provided the same content type is used.

To that end, this research does not consider potential energy savings because of data center and end-user equipment. Instead, as shown in Section 2.2, HCI power consumption is estimated by considering the query sizes of HCI methods.

3. COMPARISON OF HCI METHODS
Before estimating the amount of energy consumed by human-computer interaction methods such as interactive chat bots, virtual assistants and superapplications, there needs to be a clear understanding of the differences and commonalities between each of them. Therefore, it is necessary to have a detailed look at each of these HCI methods.

<table>
<thead>
<tr>
<th>Top-level HCI method</th>
<th>Integrated HCI methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual assistants</td>
<td>Interactive chat bots</td>
</tr>
<tr>
<td></td>
<td>Web applications</td>
</tr>
<tr>
<td>Superapplications</td>
<td>Interactive chat bots</td>
</tr>
<tr>
<td></td>
<td>Web applications</td>
</tr>
</tbody>
</table>

First, it is important to acknowledge that chat bots are generally not used independently, but are used in combination with a digital container, such as a standalone application, superapplication or virtual assistant (Table 1). The reverse is also valid as both superapplications and virtual assistants rely on the integration of singular services, such as interactive chat agents, to provide end-users with functionality. This is similar to the superapplication concept described by Han & Lee [9]. One product reflecting this technical symbiotic relationship is WeChat, which allows the integration of first- and third-party IT services [30]. In fact, without the integration of such services, it could be argued that this communication platform would be limited to providing a combination of arbitrary social media experiences and basic chat services to its customers, subsequently potentially losing its competitive advantage in the market. Therefore, the following statement could be made regarding the integration of these human-computer interaction methods:

- Virtual assistants and super applications require the integration of singular applications and services to provide sufficient functionality and generate business value.

Second, interactions through these HCI methods can generally be classified into two categories based on interaction type:
- **Graphical** (graphics-based interaction)
- **Conversational** (text-based interaction)

The main differentiator for this classification is whether additional graphical elements are rendered with or besides text. HyperText Markup Language (HTML), an essential building block to making webpages, is therefore a good example as it attempts to achieve certain graphical effects besides rendering textual elements to contextual convey information.

An alternative to this practice is conversational interaction. Instead of interacting graphically through framed markup to relay commands to an application, text or voice dictated input in a conversation-like environment would determine what action is taken. In this scenario response to the query is also shown as part of the conversation, either in text or markup. Although this does not eliminate the need to include markup language tags and additional content such as images and stylesheets, the average size of response data can be reduced as information is no longer required to be computationally tagged. Furthermore, even when markup is used as part of the response data, the document is often smaller in size with less linked content compared to full-scale web pages [7].

Based on this information, the following can now be stated for each of these methods:
- The classification of a human-computer interaction method as either conversational or graphical directly influences its response content type (Table 1).
- Both virtual assistants and superapplications are identical in having the ability to be classified as both conversational interaction and graphical interaction.
- Interactive chat bots are classified solely as a conversational HCI method; only interactive chat bots are classified as purely conversational.
- Web applications are classified solely as a graphical HCI method; only web applications are classified as purely graphical.
4. HCI POWER CONSUMPTION THROUGH QUERY ANALYSIS

In the previous chapter, we have concluded that an estimation for the power consumption of HCI methods can be determined by analyzing the size of the query response, which is directly influenced by the used content type. This method is accurate, provided the same hardware is used and the number of queries is similar. As discussed in Chapter 2, hardware efficiency can vary based on the components used and will therefore not be considered as part of this research.

Before estimating the energy consumption of each HCI method discussed in this paper, a base formula for calculating energy consumption per request needs to be defined:

\[ E_q = E_b \times n \times (\alpha_1 + \alpha_2) \]

In this formula, the total amount of energy required for each request is defined as the product of the energy-per-bit value \( E_b \), the natural number of queries \( n \) and the size of each query request \( \alpha_1 \) in addition to the size of each query response \( \alpha_2 \). However, as most HCI methods are quite similar (Chapter 3), the natural number of queries will be similar in most cases as well. Also, because the number of queries for individual HCI methods is highly dependent on the quality of implementation by software engineers, absolute results could prove to be inaccurate. Therefore, this research will measure the relative difference in energy consumption.

As such, the natural number of queries is considered equal to 1 for the purposes of this research. As mentioned in Section 2.3, the query request is assumed to be similar in size between different HCI implementations. Although measuring the query request size is not part of this research for that reason, it is still of value to estimating the total power consumption of HCI methods. Given these considerations, we can therefore rewrite the equation to:

\[ E_q = E_b \times \alpha_1 + \alpha_2 \]

This formula no longer considers the natural number of queries. Both the energy-per-bit value \( E_b \), the query request size \( \alpha_1 \) and the query response size \( \alpha_2 \) are still included in the calculation. The energy-per-bit value is computed in microjoules per bit; the query request and response sizes in bits; and the energy per query value in joules.

4.1 Energy consumption per bit

To obtain the energy-per-bit value, the energy consumption of core, access and edge networks were researched. It was found that the amount of energy consumed by a single bit travelling through core and edge networks combined is 2.7 \( \mu J \) [2, 3].

Because there exist several types of access networks, there is no single valid power consumption value for access networks. Although power consumption estimations can be found for individual access network types such as 4G/LTE and WiFi, these values change rapidly because of continuous cellular network planning and optimization [21]. The next fifth generation (5G) of mobile networks is expected to significantly improve on energy efficiency compared with current generation mobile networking solutions [15]. To future-proof the results of this research, the energy-per-bit value of access networks is therefore defined as a variable.
Consequently, the estimated amount of energy consumed for a single bit travelling through several types of networks, e.g., core, edge and access networks, needs to be defined as

\[
E_b = (2.7 + E_a)
\]

where \(E_b\) equals the amount of microjoules consumed per bit in an access network and \(E_a\) represents the average total microjoule-per-bit value.

### 4.2 Markup energy consumption model

Therefore, let us first determine the size of markup-styled content by analyzing the average web page size. Excluding linked resources, such as images and style sheets, is of utmost importance. Embedding these resources has become popular practice amongst web developers, resulting in an increase of the average web page size from 692 KiB to 2860 KiB over the last 6-7 years [13]. While datasets analyzed by Trotman & Zhang indicate an average size of 22.3 KB (21.8 KiB) HTML content per web page as of 2010 [28], data from the HTTP Archive have put this value closer to 25KB (23.8 KiB) HTML content per web page as of May 2017 [13]. As the latter estimation is based on more recent data and is not too different compared to the former, the more up-to-date estimation of 23.8 KiB = equal to 194,969.6 bits – HTML content per web page is used in calculating the estimated average energy consumption of web applications.

Based on the simplified formula described in the introduction of this chapter as well as the energy-per-bit value from Section 4.1, it is now possible to calculate the amount of energy consumed by each query with markup as its response content. This can be accomplished by solving for \(E_q\) in

\[
E_q = (2.7 + E_a) (\alpha + 194,969.6)
\]

where \(E_q\) equals the amount of microjoules per bit over an access network and \(\alpha\) is the variable query request size.

### 4.3 Plain text energy consumption model

To estimate the amount of energy consumed in conversational interaction methods, which can support both plain text and markup response content, we must derive the query response size of plain text content in bits. Although such an exact estimation does not yet exist, the average chatbot message size has already been defined at 9.72 words per message [18]. As most of the words found in English texts are between 5 and 8 characters long [22], the character count per sentence can now be defined as a value between 48.6 and 77.76 characters respectively. Additionally, this research considers a total of 12 additional punctuation elements, of which 8.72 spaces – the amount of spaces is calculated by subtracting 1 from the average amount of words – and 3 other punctuation elements, in order to base the size analysis on a well-formed sentence. The validity of these values is assumed, yet thought to be of relevance to the outcome of this analysis. Given these parameters, the average number of characters encountered in a message generated by chatbots is estimated between 60.6 and 89.76. Assuming UTF-8 encoding, after averaging character count, the query response size can be estimated at 601.44 bits.

Similar to Section 4.2.1, it is now possible to calculate the amount of energy consumed by each query with plain text as its response content. To find the amount of energy consumed per query, we want to solve for \(E_q\) in the equation

\[
E_q = (2.7 + E_a) (\alpha + 601.44)
\]

where \(E_q\) equals the amount of microjoules per bit transferred over an access network and \(\alpha\) is the query request size.

### 4.4 Comparing energy consumption models with respect to HCI methods

As we have seen in Section 4.2, the energy consumption of markup query response content can be defined as:

\[
E_q = (2.7 + E_a)(\alpha + 194,969.6)
\]

while a similar estimation was made in and Section 4.3 for plain text query response content by solving

\[
E_q = (2.7 + E_a)(\alpha + 601.44)
\]

where in both cases \(E_q\) equals the amount of microjoules per bit transferred over an access network and \(\alpha\) is the variable query request size. Although these two estimations could be directly compared, there is greater scientific value to be found elsewhere. For example, drafting a formula that computes the energy consumption of both markup as well as plain text response content allows for the easy measurement of energy consumption by the HCI methods discussed in Chapter 3. We can see that to that end, the following equation was derived

\[
E_q = \frac{n_1}{N} (2.7 + E_a)(\alpha_1 + 194,969.6) + \frac{n_2}{N} (2.7 + E_a)(\alpha_2 + 601.44)
\]

and simplified as

\[
E_q = (2.7 + E_a)\left(\frac{n_1}{N} (\alpha_1 + 194,969.6) + \frac{n_2}{N} (\alpha_2 + 601.44)\right)
\]

where \(E_q\) equals the amount of energy consumption per information query, \(n_1\) and \(n_2\) represent the number of queries with markup and plain text response types respectively, \(\alpha_1\) and \(\alpha_2\) are the average query request sizes for markup and plain text responses respectively, \(N\) represents the total amount of query responses, which should equal \(n_1\) added to \(n_2\).

When this model is used to estimate HCI method energy consumption through changing the preferred response content type, network energy usage does not need to be considered as it concerns a relative difference in power consumption. Furthermore, as the query request sizes (\(\alpha_1\) and \(\alpha_2\)) are thought to be similar in size, the percentage of energy saved compared to markup-based interaction (\(E_b\)) can be defined as

\[
E_x = \frac{194,969.6 - \frac{n_1}{N} (194,969.6) - \frac{n_2}{N} (601.44)}{194,969.6 + 601.44} \times 100
\]

where \(n_1\) and \(n_2\) represent the number of queries with markup and plain text response types respectively and \(N\) equals the total amount of query responses. Based on this equation, the potential energy savings in a 100:0 plain text-to-markup ratio could be as high as 99.69%. Other ratios and their associated energy savings are shown in Table 4.

| Table 4. Energy savings for plain text-to-markup content type ratios in query responses |
|---|---|---|
| Markup (%) | Plain text (%) | Energy savings (%) |
| 100% | 0% | 0% |
| 80% | 20% | 20.06% |
| 60% | 40% | 40.12% |
| 40% | 60% | 60.19% |
| 20% | 80% | 80.25% |
| 0% | 100% | 99.69% |
4.5 Estimation model considerations

Although these equations can be used to assist developers and future researchers to estimate the energy consumption of certain HCI method implementations based on the ratio in which specific content types are used, they cannot be used to determine the overall efficiency of specific HCI methods. Furthermore, it is worth noting that these equations assume that the markup response size is identical to the average web page size as of May 2017. In practice, however, the amount of markup in a query response can be significantly lower than that of a full-size web page. This is achieved through mobile-optimized web pages, which are 13.6% smaller in size than normal-sized web pages [14], custom smaller-sized HTML frames for chatbots or any other means of HTML content extraction [7]. Researchers and developers wanting to use these equations should therefore carefully contemplate about the expected response size of their queries and adjust the formulas accordingly to arrive at accurate and valid power consumption estimations for their IT products and services.

5. RESULTS AND DISCUSSION

Although the results from Section 4.4 indicate that there is a lot to gain in terms of energy savings from changing the ratio of content types in query responses of IT services, it remains unclear how exactly this would translate to products and services currently found on the market. As each IT solution is unique in some sense, the text-to-markup ratio of information requests is different for each case. Therefore, it needs to be concluded that without additional research it is impossible to determine the total amount of energy saved by the ICT industry if the results of this research were widely considered in IT solutions.

Yet, it does not seem that these energy savings will remain obscure much longer. With the expected upsurge of virtual assistants and global chatbots, as discussed in the Introduction (Chapter 1), it is likely that more companies, governmental agencies and researchers become increasingly aware of these human-computer interaction methods that have proven to be more energy-efficient than their predecessors.

Even though most markup-based web services will not be completely replaceable by conversational web services – and it is likely some will never be replaced –, this does not mean that this option should be left unexplored by researchers. In fact, as the results of this research indicate that the energy consumption of IT services predominantly making use of plain text interaction could be up to 99.69% lower than that of IT services exclusively using markup-based interaction, it could be stated that there is exciting potential to save a significant amount of kWh. As the strain on the climate created by the ICT sector is not set to decrease, these results could aid reducing the environmental impact of ICT.

6. CONCLUSION

The primary goal of this research was to provide business leaders and government agencies with valuable insight in the energy consumption of HCI methods to adapt policy to the most energy-efficient method if desired.

To achieve that goal, it was first established that the number and size of queries sent over a network represent the effect of HCI methods on ICT power consumption (Chapter 2), consequently answering the first research question.

Then, the differences between several HCI methods were explored in Chapter 3. With the intention of finding a result to the second research question, this comparison highlighted the differences that were thought to be affecting the energy consumption of IT services. It was found that any difference in the amount of network traffic can be explained as a consequence of using specific content types, of which HTML and plain text were explored, in information requests.

In Chapter 4, this research put forward a new method of estimating energy consumption of IT products and services of a specific HCI type. The equation found in this chapter should provide software engineers and researchers with adequate information to guide them to develop more energy-efficient IT solutions in future work.

6.1 Limitations

Although, this solution is adequate to give a reasonable estimation of the energy consumption by single IT services and products, it is by no means perfect.

First, it only considers the energy consumption as a result of network traffic. Therefore, any change in the amount of processing required in data centers is neglected. The same applies to client device energy consumption. Although this could be estimated by similar means, it was decided that it was better to be left out as the energy consumption of different hardware components was deemed to be too diverse.

Second, it assumes that the response size of markup-based messages will always be a full-sized web page. This is definitely not always the case. Instead, software engineers usually find themselves writing simplified markup messages in order to reduce complexity and improve transfer times. Therefore, any person wanting to make use of this research should be advised to take this into account and adapt the HCI content-type equation accordingly.

6.2 Future work

Based on the limitations put forward in Section 6.1, the following recommendations can now be made for future work.

First, any future work should focus on validating the claims made in this research as there is little comparative research in this field. Eventually, this will lead to the possibility of creating more accurate estimations for HCI-specific energy consumption values.

Second, future work in this field should also consider the impact of conversational interaction on data centers and customer hardware as this is currently left out of the research. As all hardware is different in terms of energy consumption, it is recommended to perform case studies for the most popular IT services and products.
7. REFERENCES


