



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

A large, stylized sunburst or fan-like graphic in a lighter shade of teal, positioned on the left side of the cover, partially overlapping the title text.

EXPLORING THE USE OF TWITTER DATA TO BETTER UNDERSTAND CHURCH ATTENDANCE

Anthony-Paul Cooper



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

EXPLORING THE USE OF TWITTER DATA TO BETTER UNDERSTAND CHURCH ATTENDANCE

Anthony-Paul Cooper

University of Turku

Faculty of Technology
Department of Computing
Computer Science
Doctoral Programme in Mathematics and Computer Sciences

Supervised by

Revd Professor Erkki Sutinen
University of Turku

Revd Dr Peter Phillips
Durham University

Reviewed by

Professor Love Ekenberg
Stockholm University

Professor Tim Hutchings
University of Nottingham

Opponent

Professor Lorenzo Cantoni
University of Italian Switzerland

The originality of this publication has been checked in accordance with the University of Turku quality assurance system using the Turnitin OriginalityCheck service.

ISBN 978-951-29-8656-9 (Print)
ISBN 978-951-29-8657-6 (PDF)
ISSN 2736-9390 (Print)
ISSN 2736-9684 (Online)
Painosalama, Turku, Finland 2021

I dedicate this thesis to loving memory of my late grandfather, Jack Holland. He was a constant source of love and support and took a keen interest in my research, reading and commenting on my publications and providing helpful and meticulous proofreading; including of this thesis. I am forever grateful for his encouragement and wisdom.

UNIVERSITY OF TURKU

Faculty of Technology

Department of Computing

Computer Science

ANTHONY-PAUL COOPER: Exploring the use of Twitter data to better understand church attendance

Doctoral Dissertation, 173 pp.

Doctoral Programme in Mathematics and Computer Sciences

September 2021

ABSTRACT

Over recent years, the predominant narrative of UK churchgoing within the academic literature has been one of decline, underpinned by the secularisation thesis which posits that as society modernises, religious belief becomes less plausible and important. Until recently, academic research into church growth in the UK has received comparably little attention, however, recent studies have begun to challenge the narrative of inevitable religious and church attendance decline and, accordingly, some of the assumptions associated with the secularisation thesis. These challenges have been met with robust academic debate and the matter is not yet settled. What is clear, however, is that both the proponents and challengers of the secularisation thesis acknowledge fundamental shortcomings in the available data on church attendance.

This doctoral thesis addresses this gap in knowledge by exploring whether Twitter data, a freely available source of social media data, can be used to better understand church attendance. It finds that it is possible to capture church-related Twitter data and interrogate that data using thematic qualitative analysis to understand the church-related topics which Twitter users discuss in their posts. Furthermore, this thesis finds that it is possible to automate that analysis using supervised machine learning (in this case, the Naïve Bayes classifier); thus, paving the way for researchers to quickly repeat and automate this process at scale. Having explored the topics covered in church-related tweets this thesis demonstrates, using an independent-samples t-test, that there is a relationship between the sentiment of church-related tweets and the presence of church growth in the location from which the tweets were posted. This paves the way for researchers to investigate the potential use of Twitter data in future research studies where church attendance data is missing or of poor quality. This thesis also experimentally demonstrates that it is possible to identify the location of churches, several of which were excluded from previous church attendance studies, using a spatial analysis of church-related tweets. This creates an exciting opportunity for Twitter data to be used in future church attendance studies to improve the targeting of surveys and data collection. Finally, having set the social media studies into the wider context of digital theology, this thesis concludes by exploring the response of the church globally to the recent COVID-19 pandemic. This thesis uses a qualitative survey to describe some of the recent technological innovations which churches have made and raises timely questions around the considerations which those innovations will necessitate for future studies of church attendance.

KEYWORDS: digital theology, sociology of religion, church attendance, church growth, secularisation thesis, social media, content analysis, machine learning, sentiment analysis, geospatial analysis.

TURUN YLIOPISTO

Teknillinen tiedekunta

Tietojenkäsittelytieteen laitos

Tietokone Tiede

ANTHONY-PAUL COOPER: Twitter-datan käyttömahdollisuudet

kirkossakäynnin tarkastelussa

Väitöskirja, 173 s.

Matematiikan ja tietojenkäsittelytieteen tohtoriohjelma

Syyskuu 2021

TIIVISTELMÄ

Brittien kirkossakäyntiä tarkastelevaa tutkimuskirjallisuutta on viime vuosina hallinnut kertomus taantumasta. Sen taustalla on ollut sekularisaatioteesi, jonka mukaan uskonnolliset menettävät uskottavuuttaan ja merkitystään yhteiskunnan modernisoituessa. Tieteellinen tutkimus kirkkojen kasvusta Yhdistyneessä kuningaskunnassa oli vielä taannoin saanut verraten vähän huomiota. Viime aikoina julkaistut tutkimukset ovat kuitenkin alkaneet kyseenalaistaa kertomusta uskonnollisuuden ja kirkossakäynnin väistämättömästä taantumasta sekä vastaavasti kritisoineet eräitä sekularisaatioteesin taustaoletuksia. Tämä on synnyttänyt kiivasta akateemista keskustelua, eikä asiasta toistaiseksi vallitse yksimielisyyttä. On kuitenkin selvää, että sekä sekularisaatioteesin kannattajat että kyseenalaistajat tunnustavat saatavilla olevan kirkossakäyntidatan perustavat puutteet.

Tämä väitöskirja pyrkii täyttämään kyseistä tietoaukkoa tutkimalla, voiko kirkossakäynnin nykytilasta saada tarkemman käsityksen hyödyntämällä Twitter-dataa, joka on vapaasti saatavilla oleva sosiaalisen median datan lähde. Tutkimuksessa todetaan, että kirkkoon liittyvän Twitter-datan kerääminen on mahdollista ja sitä voidaan tarkastella temaattisessa laadullisessa analyysissä, jotta saadaan käsitys niistä kirkkoon liittyvistä aiheista, joista Twitterin käyttäjät keskustelevat päivityksissään. Lisäksi väitöskirjassa todetaan, että kyseinen analyysi on mahdollista automatisoida valvotun koneoppimisen avulla (tässä tapauksessa naiivin Bayesin luokittimen avulla). Näin tutkijoille tarjoutuu mahdollisuus toistaa ja automatisoida kyseinen prosessi nopeasti suuremmassa mittakaavassa. Kirkkoon liittyvien twiittien aiheiden tarkastelun jälkeen tutkimuksessa osoitetaan riippumattomien otosten t-testillä, että kirkkoon liittyvien twiittien sävyn (sentiment) ja twiittien lähetyspaikassa tapahtuvan kirkon kasvun välillä vallitsee yhteys. Tällä perusteella tutkijat voivat selvittää Twitter-datan käyttömahdollisuuksia tulevissa tutkimuksissa, kun kirkossakäyntidata puuttuu tai on huonolaatuista. Väitöskirjassa osoitetaan myös kokeellisesti, että kirkkoon liittyvien twiittien spatiaalisella analyysillä voidaan paikantaa kirkkoja, joista monet ovat jääneet aiempien kirkossakäyntitutkimusten ulkopuolelle. Tulevien kirkossakäyntitutkimusten kannalta onkin kiinnostavaa, että Twitter-datan avulla kyselyt ja datankeräys on mahdollista kohdentaa paremmin. Väitöskirjassa sosiaalisen median tutkimus liitetään digitaalisen teologian laajempaan kontekstiin, ja tutkimuksen päätteeksi tarkastellaan kirkkojen maailmanlaajuisista reaktiosta viimeaikaiseen COVID-19-pandemiaan. Väitöskirjassa käytetään laadullista kyselyä eräiden uusien, kirkkojen tekemien teknologisten innovaatioiden kuvailuun ja herätetään tulevien tutkimusten tarpeisiin ajankohtaisia kysymyksiä siitä tarkastelusta, jota nuo innovaatiot edellyttävät.

ASIASANAT: digitaalinen teologia, uskontososiologia, kirkossa käyminen, kirkon kasvu, sekularisaatiotutkimus, sosiaalinen media, sisältöanalyysi, koneoppiminen, tunteiden analyysi, paikkatietoanalyysi.

Table of Contents

| | |
|---|-----------|
| Abbreviations | 8 |
| List of Thesis Publications | 9 |
| Author's Contribution to the Thesis Publications | 10 |
| Other Related Publications..... | 12 |
| 1 Introduction..... | 13 |
| 2 Literature Review..... | 16 |
| 2.1 Church Growth, Church Decline and the Secularisation Thesis | 16 |
| 2.2 Church Growth in London | 19 |
| 2.3 Twitter Research..... | 22 |
| 2.4 Digital Theology | 26 |
| 3 Research Design..... | 29 |
| 3.1 Research Questions..... | 29 |
| 3.2 Data Collection..... | 31 |
| 4 Component Studies: Summary and Results | 34 |
| 4.1 P1: Understanding London's Church Tweeters: A Content Analysis of Church-Related Tweets Posted from a Global City | 34 |
| 4.2 P2: Exploring the Use of Machine Learning to Automate the Qualitative Coding of Church-Related Tweets..... | 38 |
| 4.3 P3: Assessing the Possible Relationship Between the Sentiment of Church-Related Tweets and Church Growth | 43 |
| 4.4 P4: Using Geotagged Twitter Data to Uncover Hidden Church Populations..... | 48 |
| 4.5 P5: Faith Communities Online: Christian Churches' Reactions to the COVID-19 Outbreak | 54 |
| 5 Discussion | 58 |
| 5.1 Summary of Answers to Research Questions | 58 |
| 5.2 Key Insights and Relevance of Findings..... | 60 |
| 5.3 Key Limitations..... | 63 |
| 5.4 Opportunities and New Horizons for Future Research | 65 |

| | | |
|----------|---|-----------|
| 5.5 | Practical Recommendations for Church Leaders..... | 67 |
| 6 | Conclusion | 68 |
| | Acknowledgemets | 70 |
| | List of References..... | 72 |
| | Original Publications | 77 |

Abbreviations

| | |
|----------|--|
| ANOVA | Analysis of Variance |
| API | Application Programming Interface |
| COVID-19 | Coronavirus Disease (COVID-19) |
| DCA | Discussion about Church Attendance |
| DCSC | Discussion about Church Service Content |
| DT | Digital Theology |
| DTBB | Discussion about Theology, the Bible or Belief |
| GCD | General Church Discussion |
| HLSG | Discussion about Hillsong |
| ID3 | Iterative Dichotomiser 3 Algorithm |
| MISC | Miscellaneous |
| NCD | Non-Church Discussion |
| RQ | Research Question |
| SASRF | South African Science and Religion Forum |
| SD | Standard Deviation |
| SVM | Support-Vector Machine |
| UNK | Unknown Theme |
| uSa | Usual Sunday Attendance |
| WEKA | Waikato Environment for Knowledge Analysis |
| WGS84 | The World Geodetic System 1984 |

List of Thesis Publications

This doctoral dissertation summarises the following compilation of published research articles:

- P1 Cooper, A. P., Mann, J., Sutinen, E. & Phillips, P. (2021). “Understanding London’s church tweeters: A content analysis of church-related tweets posted from a global city”. *First Monday*, 26(9).
- P2 Cooper, A. P., Kolog, E. A. & Sutinen, E. (2019). “Exploring the use of machine learning to automate the qualitative coding of church-related tweets”. *Fieldwork in Religion*, 14(2), p. 140-159.
- P3 Cooper, A. P. (2017). “Assessing the possible relationship between the sentiment of church-related tweets and church growth”. *Studies in Religion/Sciences Religieuses*, 46(1), p. 37-49.
- P4 Cooper, A. P. (2018). “Using geotagged Twitter data to uncover hidden church populations”. In *The Desecularisation of the City: London’s Churches, 1980 to the Present*, edited by David Goodhew and Anthony-Paul Cooper. Abingdon, UK: Routledge, p. 134-147.
- P5 Cooper, A. P., Jormanainen, I., Shipepe, A. & Sutinen, E. (2021). “Faith communities online: Christian churches’ reactions to the COVID-19 outbreak”. *International Journal of Web Based Communities*, 17(2), p. 99-119.

Author's Contribution to the Thesis Publications

The author led the research design and drafting process for each of the thesis publications. Prior to commencing the studies described in publications P1 to P4, the author collected the required dataset using the approach described in the research design chapter of this dissertation. Prior to commencing the study described in P5 the author designed the research survey to be used by the contributing authors.

The precise role of the author in each of the thesis publications was as follows:

- P1 Having collected the required dataset, the author designed the study approach, conducted the primary data coding and wrote the majority of the first full draft of the publication. Joshua Mann conducted the secondary data coding and wrote some longer sections of text which were incorporated into the publication. Erkki Sutinen and Peter Phillips provided critique on the research design and structure of the publication and contributed to the writing of the publication, making tweaks and edits to the text using tracked changes and comments.
- P2 Having collected the required dataset, the author co-designed the study approach with Emmanuel Awuni Kolog. Emmanuel Awuni Kolog conducted the machine learning analysis. The author and Emmanuel Awuni Kolog interpreted the results and wrote the results and discussion sections together. The author wrote the majority of the first full draft of the remainder of the publication. Erkki Sutinen provided critique on the research design and structure of the publication and contributed to the writing of the publication, making tweaks and edits to the text using tracked changes and comments.
- P3 The author was the sole author for this publication and independently conducted the study and writing of the publication.

- P4 The author was the sole author for this publication and independently conducted the study and writing of the publication.
- P5 The author designed the research survey which was issued to church leaders in the UK, Finland and Namibia. The author led the UK survey component and wrote that section of the publication. Ilkka Jormanainen led the Finland component and wrote that section of the publication. Annastasia Shipepe led the Namibia component and wrote that section of the publication. The author oversaw the writing of the remaining sections of the publication, with the other authors making equal contributions. In addition to contributing to the writing of the publication, Erkki Sutinen provided critique on the research design and structure of the publication.

Other Related Publications

In addition to the thesis publications already set out, the author contributed to the following other relevant publications:

- P6 Cooper, A. P. & Goodhew, D. (2017). “‘Resacralising’ secular space: new churches in a northern city, 1980–2012”. *Journal of Contemporary Religion*, 32(3), p. 495-511.
- P7 Goodhew, D. & Cooper, A. P. [Eds.] (2018). *The desecularisation of the city: London’s churches 1980 to the present*. Abingdon, UK: Routledge.
- P8 Sutinen, E. & Cooper, A. P. (2020). “Interactive technologies, Missio Dei and grassroots activism”. In *Missio Dei in a Digital Age*, edited by Jonas Kurlberg and Peter M. Phillips. London, UK: SCM Press, p. 122-149.
- P9 Cooper, A. P., Laato, S., Nenonen, S., Pope, N., Tjiharuka, D. & Sutinen, E. (2021). “The reconfiguration of social, digital and physical presence: From online church to church online”. *HTS Teologiese Studies/Theological Studies*, 77(3).
- P10 Sutinen, E. & Cooper, A. P. (2021). *Digital theology: A computer science perspective*. Emerald, Bingley, UK.

1 Introduction

Over recent years, the predominant narrative of UK churchgoing within the academic literature has been one of decline (see, for example: Brown, 2001; Bruce, 2002), underpinned by the secularisation thesis which posits that as society modernises, religious belief becomes less plausible and important (Wallis & Bruce, 1991).

More recently, however, there has been an increase in academic research focused on pockets of church growth observed in the UK (see, for example: Goodhew, 2012a & P7; Goodhew & Cooper, 2018a). Such growth includes increasing numbers of congregations, increasing numbers of church members and increasing numbers of church service attendees. These studies do not claim to fully refute the secularisation thesis; indeed, in the introduction to *Church Growth in Britain, 1980 to the Present*, Goodhew remarks:

“Such developments are no cause for ecclesiastical triumphalism, given the extent of church decline elsewhere. But this narrative of *Church Growth in Britain* subverts the dominant narrative in much of academia and the media, a narrative which heavily influences many church leaders as well as wider society. The dominant narrative assumes that there has been wholesale church decline in recent decades and that such decline is the primary reality of British Christianity.” (Goodhew, 2012b, p. 3).

While recognising that church decline is taking place in parts of the UK, these studies highlight that church growth is *also* taking place. And that such church growth is not isolated or exceptional; it is taking place in a number of different places and across a number of different styles of worship. These studies have, through their assertions, therefore, begun to challenge the narrative of inevitable religious and church attendance decline and, accordingly, some of the assumptions associated with the secularisation thesis. These challenges have been met with robust academic debate and the matter is not yet settled. Resultant academic discussions on the empirical measurement of church growth and church decline have highlighted shortcomings in the currently available church attendance data.

These data shortcomings exist for a number of reasons. The results from the ten-yearly UK census which relate to religious affiliation and practice need to be

interpreted with caution; considered in isolation, the results can paint very different pictures – for example, in the 2011 British Census, only 47% of 18 to 34 year olds self-declared as having any religious affiliation, yet 67% of the same age group self-declared that they regularly or occasionally pray (Goodhew, 2012b, p4). This indicates that question wording design is a very important component of large survey studies and that question wording can influence self-declaration results, making interpretation of such data challenging (Goodhew, 2012b, p4). Furthermore, comprehensive UK-wide census studies of churchgoing are complex and expensive to conduct. For these reasons, the church attendance data reports produced by Peter Brierley (for example, the most recent: 2020) are probably the most reliable sources of data on UK church attendance. They are, however, largely based on returns from the church denominations themselves, which by their nature can be hard to verify. Where data points are missing, Brierley extrapolates from previous data which, again, may not always represent an accurate picture. Furthermore, the approach taken to contacting churches and denominations is very heavily dependent on Brierley’s own database of contact points. This database is not available on an open-source basis and, therefore, is considered black box in terms of auditability. The completeness of the database is also dependent on the effectiveness of the snowballing approach taken, whereby churches and denominations already recorded in the database are asked to provide details of new churches and denominations in their locality which should also be contacted as part of the survey; again, this is difficult to verify.

It is the issue of church attendance data quality challenges that this thesis seeks to begin to address. The conventional methods and approaches previously taken to gather data on church attendance are time consuming, resource-intensive and prone to error. The incomplete, inaccurate and contradictory data make it difficult to assess the reality of church attendance; an issue likely to persist into the future. The application of other, freely available and well-structured, datasets to complement traditional approaches and datasets might offer an opportunity to both reduce the burden of large-scale church attendance surveys and improve the reliability of their results. This thesis attempts to explore whether Twitter¹ data could be used in this way, by conducting a number of research studies using church-related tweets posted from within London, to assess the art of the possible using such data – to investigate

¹ Twitter is a popular microblogging social media platform, which enables users to post short text updates known as tweets. Tweets are visible to users across the platform and can optionally be accompanied by photographs and videos. When the thesis dataset reported on in this dissertation was collected tweets were limited to 140 characters. This has since been extended to 280 characters. See: <http://www.twitter.com/> (last accessed: 10 April 2021).

whether such data could be meaningfully qualitatively analysed, whether that analysis could be automated, whether a relationship could be detected between that data and offline activity and whether that data could be used to detect physical church locations. In addressing these questions, this thesis sets church-related social media research in the context of digital theology; an emerging field of academic enquiry, which is deeply rooted in practice. Considering church-related social media research in this way permits consideration of the wide range of other use cases which churches might be able to apply to the research methods and analyses described in this dissertation.

To explore the potential application of Twitter data to studies of church attendance, this thesis focuses on London. London was selected as the focus of the research as, at the time the research commenced, the recently published results of the London Church Census (Brierley, 2013) were arguably the most reliable recent statistics on church attendance within the UK. This, therefore, provided a suitable context within which to conduct the research and a useful baseline for interpreting results – it would be possible, for example, to compare the list of churches identified using Twitter data (this is described in more detail in the summary of **P4**) to the list of churches contacted as part of the London Church Census to understand whether any of the churches identified had been excluded from that church attendance study.

During the final year of the thesis research, the world was disrupted by the COVID-19 pandemic. The spread of the virus was rapid and unforeseen and forced countries around the world to very quickly impose restrictions to limit the spread of the virus and save lives. This meant that churches globally were required to innovate very quickly and work at pace to implement new technologies to maintain their services and offerings. In the wider context of digital theology, this represented a major shift in the digital posture of the church, which continues to evolve at the time of writing this dissertation. The final study of this thesis, therefore, focused specifically on the response of the church globally to the COVID-19 pandemic. This was explored using qualitative survey data collected from a convenience sample of churches in the UK, Finland and Namibia. Analysis of those survey responses offers a rapid and timely understanding of the actions taken by churches to permit their continued operation. As part of that study, thought was given to the considerations which those innovations will necessitate for future studies of church attendance.

2 Literature Review

2.1 Church Growth, Church Decline and the Secularisation Thesis

Whilst it is beyond the scope of this computer science dissertation to explore the secularisation thesis in great depth, it is important to briefly set the context for the research questions addressed in this dissertation. These research questions were set following a specific debate on the secularisation thesis between Steve Bruce and David Goodhew (Bruce, 2013a; Goodhew, 2013; Bruce, 2013b), during which the importance of addressing church attendance data quality issues was raised.

The prevalent narrative of church attendance in the UK in recent years has been overwhelmingly one of decline (see, for example: Brown, 2001; Bruce, 2002). Some interpretations of this trend have posited that declining attendance is not necessarily tied to proportionate declining belief, for example the proposed phenomenon of believing without belonging (Davie, 1994). Such interpretations have, however, tended to be met with challenge – Voas and Crockett (2005), for example, present an argument underpinned by data which they believe demonstrates that religious belief in the UK is declining at roughly the same rate as religious affiliation and religious attendance (p. 13). The secularisation thesis on the other hand has, for some time, been used to explain declining church attendance as a consequence of declining belief. Wallis and Bruce (1991) argued that the increasing modernisation of societies was driving a decrease in the plausibility and importance of religious belief. Such arguments have tended to drive a predominant narrative of a struggling church of fading relevance in an increasingly modern society.

More recently, however, a competing narrative of church growth has begun to emerge, challenging existing ideas around the presumed inevitability of continued decline in church attendance. An edited volume exploring church growth in Britain in the period 1980 to 2010 (Goodhew, 2012a) saw a team of 15 researchers present a compilation of essays exploring observed aspects and examples of church growth across a range of denominations and locations. In the introduction, the editor of that volume remarked that:

“This growth is large-scale; it is occurring across a wide geographical range; it is highly multi-cultural in its social reach; and it shows no sign of slowing down. The current consensus, by focusing almost exclusively on decline, is seriously mistaken.” (Goodhew, 2012b, p. 3).

It was argued within the volume that the findings, evidence of pockets of church growth, called into question the validity of the secularisation thesis narrative. This argument prompted a renewed discussion in the academic literature on the secularisation thesis and the indicators which can be used to argue for and against secularisation. It was that discussion which inspired and informed the research questions addressed in this thesis.

As part of that renewed discussion, Bruce (2013a) argued that church growth and increasing secularisation were not mutually exclusive phenomena; that the secularisation thesis did not require a complete elimination of religion, but rather a decrease in “the plausibility, popularity, and power of shared religious or spiritual beliefs” (Bruce, 2013a, p. 278). This argument was an expansion on a previous definition of secularisation by Bryan Wilson who stated, more simply, that secularisation was the decrease in “social significance of religion” (1982, p. 149).

Expanding on the observations of church growth more specifically, Bruce contended that instances of church growth did not threaten the integrity of the secularisation thesis. To illustrate this point, Bruce used a metaphor based on hat wearing – the trend of hat wearing has decreased in popularity since the 1950s and haberdashers have accordingly sold fewer hats since that time; this overall decrease in hat wearing does not, however, preclude the opening of any new hat shops at all or even an increase in hat wearing within some pockets of society (2013a, p. 279). In the same way, Bruce suggests that it is possible for religiosity and the relevance of religion to be declining at the same time that some churches are innovating/growing while other new churches are opening. Bruce argues that what is important in the context of discussing the secularisation thesis is the net numbers of church attendance; that to argue for a reversal in the secularisation thesis (or even to argue that the secularisation thesis is, in some way, flawed) scholars exploring church growth would need to demonstrate that observed increases in church attendance offset decreases in church attendance elsewhere (i.e. that growing churches were not effectively drawing attendees away from other declining churches). Furthermore, Bruce goes on to argue that any observed net increase in church attendance would also need to be considered against the backdrop of local population data; if a town or city’s population drastically increased across a study measurement period a simple up-lift in attendance would not, in Bruce’s view, be sufficient to argue against the secularisation thesis – it would need to be shown that the *rate* of church attendance had also increased (2013a, p. 279).

In his response to Bruce's analysis, Goodhew (2013) argues firstly that something more than simply net church decline is required to support the secularisation thesis and secondly that methodological limitations currently inhibit the ability of academics to be certain about net measures of church attendance.

Goodhew's rationale that proof of net church decline alone is insufficient to defend the secularisation thesis is made on the basis that over recent decades the secularisation thesis has operated on the assumption of "overwhelming, ongoing decline" (p. 301) and the assertion that "church growth is either not happening at all or happening only on a peripheral scale" (p.301). Goodhew concluded, therefore, that any significant observations of church growth, irrespective of whether this was indicative of net church growth, "would seriously call into question both explicit and implicit versions of the secularisation thesis, as they have been expressed with regard to Britain" (p. 301).

In considering the methodological challenges associated with reliably identifying net growth or decline, Goodhew highlights issues with two of the most commonly used metrics in studies of church growth and decline – church membership and usual Sunday attendance. Goodhew explains that the two metrics often show contradictory trends, noting that "one key source calculates that net church membership in England is much more robust than usual Sunday attendance – holding steady at between 3,600,000 and 3,700,000, whereas net attendance is markedly falling"² (Goodhew, 2013, p. 311; Quoting: Brierley, 2011, 1.1 and 13.8). Goodhew also notes that another key metric, number of churches, is also prone to data quality issues, observing that "existing figures for the number of new churches likely exhibit undercounting" (p. 311-312). Notwithstanding a degree of undercounting which Goodhew believes to be present in this dataset, he goes on to note that various data sources suggest that the number of churches opening in recent years have exceeded the number of churches closing over the same time period, thus, indicating net growth in the number of church congregations (p. 312). While this narrative does not address the problem posed by Bruce in identifying whether the church is experiencing net growth or net decline, it does demonstrate some of the key challenges in arriving at a useful answer to this problem. It is clear that there is a valuable role for innovative new methods to play in helping to improve studies, and the resultant data, on church attendance.

All of this argumentation is set against a backdrop of uncertainty, raised by Goodhew, as to the overall merits of assessing the presence of secularisation based only on net church growth decline data *even if* such data could be reliably attained *and even if* such data could be described as sufficiently substantive to credit/discredit

² The original publication (Goodhew, 2013) refers to usual Sunday attendance as uSa, however, it has been modified in the above quotation for clarity and ease of reading.

the secularisation thesis. Goodhew notes that to focus so closely on net data could mean that much of the nuance underpinning that data is lost; Goodhew remarks:

“Bruce argues that the only question worth asking is the broad-brush question of whether there is ‘net’ congregational growth or decline. Church Growth in Britain questions the dominance of that question, since it closes down discussion and obscures key regional, ethnic, and ecclesial differences.” (Goodhew, 2013, p. 302).

In a further follow-up paper Bruce (2013b) refutes the implication that he believes the question of whether there is net congregational growth or decline is the only one worth asking when considering the relevance and applicability of the secularisation thesis (p. 317). He goes on to clarify that in order to properly appraise the secularisation thesis, scholars “need data on all the churches and chapels because otherwise we cannot distinguish internal movement from growth or decline” (p. 318).

It is clear from this intellectual exchange that there are severe limitations to the existing evidence base around church attendance and, therefore, church growth and church decline. The denominational structure of the church in the UK means that different denominations and different congregations within denominations measure church membership, church attendance and church activities in different ways, which are not always directly comparable. It is also clear that the research methods employed in previous studies mean that different metrics are not always comparable (e.g. church membership and usual Sunday attendance) or even reliable (in cases where undercounting is observed). The question of establishing whether the church is seeing net growth or decline is too complex to be addressed in one doctoral dissertation, however, the thesis described in this dissertation sets out to begin to address the issue of poor quality church attendance data by investigating whether it might be possible to use Twitter data to mitigate some of these shortcomings and, therefore, support the research community in driving forward research to better understand patterns of church growth and church decline.

2.2 Church Growth in London

To consider the applicability of Twitter data in the academic quest to better understand church growth and church decline, the doctoral thesis described in this dissertation explores data related to church attendance and church growth in Greater London. London was selected as the focus of this study as it had the most reliable and up-to-date church attendance data at the time the study commenced in April

2014; this data was largely drawn from Peter Brierley's 2012 London Church Census (Brierley, 2013).

The 2012 London Church Census was conducted on 14 October 2012; when published, "it showed an estimated 720,000 people were attending church on a Sunday across London's 4,800 churches, or 9% of the population" (Brierley, 2013, p. 3). Brierley noted that this represented a notable increase, when compared to a previous study which was conducted in 2005 as part of a country-wide census; the number of churches was up 17% from 4,100 churches in 2005 and the number of attendees was up by 16% from 620,000 the same year (Brierley, 2013, p. 3). High volume immigration to London was attributed as being a key driver of this observed growth, with Brierley observing that "one of the growth points has been churches specifically catering for particular nationalities and languages" (Brierley, 2013, p. 4).

In terms of methodological approach, ahead of the census the London Church Census team mailed letters to 4,500 churches across London which they had a record of on their database. Within the letter the team advised leaders of the forthcoming census and asked the church leaders to advise the team of any new churches which had started locally which might not feature on the team's database. The team received a strong positive response to this request and were eventually able to send out the census form to just over 4,800 churches, although Brierley notes that some of those church details proved to be out-of-date. As a result of this snowballing technique combined with desk-based research, Brierley estimated that there was a total of 4,791 churches. (Brierley, 2013, p. 13). This reliance on an existing database of London's churches combined with a snowballing technique could be considered a limiting factor of the London Church Census; it is within the realm of possibility that in a city where churches regularly form, merge, move, split and close several churches could slip through the net of such research, resulting in a degree of undercounting. It is also unlikely, going forward, that as London's churches become increasingly diverse the leaders of those churches will continue to invest much time or effort in making contact with other local church leaders to support research efforts. One aspect of the doctoral research presented within this dissertation (**P4**) was specifically designed to explore this possibility of undercounting arising through these limitations, by seeking to identify potential innovative approaches to mitigating this challenge, should it be demonstrated to exist.

Although there are limitations to the London church attendance data provided by Brierley, these are probably the best and most reliable data available, and they have been widely used by academics across the community. The picture they paint is supplemented by a range of other studies which tell a similar story of recent church growth in London's capital. Rogers (2018), for example, presents evidence that at least 240 black-majority churches have been started in the London Borough of

Southwark alone since the 1950s, numerous of those in recent years. Many of those new churches were concentrated in particular areas within the borough – for example 118 such churches were identified in Peckham. Similarly, Marchant (2018) presents evidence of around 350 new churches in the London Borough of Newham between 1975 and 2015, of which, over 220 were believed to be in operation at the time that study was published. When considered together, this evidence points to notable church growth within the UK capital. Such church growth is not, however, ubiquitous across all denominations and styles of worship – Piggot (2018), for example, shows in his published study how Methodism has been observed to be in decline since 1980.

The story of church growth and decline is not always straightforward. While it is often assumed that church growth is mainly happening within modern Pentecostal and charismatic churches, the more traditional denominations have seen church growth too. The Church of England Diocese of London, for example, was in decline until 1990 but has seen sharp church growth since that point; usual Sunday attendance, for example, was observed to have grown by 18% between 1990 and 2014 while weekday attendance between 2001 (the point at which this measure commenced) and 2014 was observed to have grown even more quickly than that (Jackson, 2018). In contrast, however, the Church of England Diocese of Southwark has seen recent church decline; the electoral roll of Southwark, for example, fell from 49,000 in 1992 to 42,000 in 2014 (Jackson, 2018). The Diocese of London and the Diocese of Southwark share similar demographics, so these differences in outcomes cannot be quickly and readily attributed to differences in rates of migration. It is unclear precisely why such different outcomes have been observed, however, it has been argued that leadership, intent and strategy could play their part in driving this difference, with the Diocese of London being much more deliberately focused on promoting church growth than the Diocese of Southwark (Jackson, 2018).

A recent report published by the think tank Theos (Bickley & Mladin, 2020) explored religiosity and associated social attitudes within London, observing that “London’s religiosity means that the city is less left-wing, welfarist, and more authoritarian than it would otherwise be” (p. 61). This demonstrates that a clear understanding of the trajectory of the Christian faith within London is likely to be of increasing importance to societal leaders and policymakers, as well as to academics and religious practitioners.

As has been demonstrated, there is much nuance and detail hidden within the available statistics on church attendance in London. What is clear, however, is that London is seeing church growth on a notable scale. More work is required, however, to address some of the key data quality issues associated with church attendance data for London. The research summarised in this dissertation seeks to begin to address

this challenge by exploring the potential applications of church-related Twitter data to better understanding church attendance.

2.3 Twitter Research

This doctoral study considers Twitter research to be a sub-set of the wider field of Social Media Research, which **P1** defined as:

“The study of social media websites and applications to consider how such platforms are used, why such platforms are used and the impact arising from the use of such platforms” (**P1**: Cooper, Mann, Sutinen & Phillips, 2021).

Commenting on the importance and value of Twitter data in his book on the impact of Twitter on social communication, Murthy (2013) importantly observes that “Twitter has the potential to increase our awareness of others and to augment our spheres of knowledge, tapping us into a global network of individuals who are passionately giving us instant updates on topics and areas in which they are knowledgeable or participating in real-time” (p. x).

A plethora of previous studies have explored the use of Twitter data to address research questions and it is beyond the scope of this chapter to provide a comprehensive and systematic review of that extant work. However, this section aims to provide a high-level summary of some of the most relevant previous work.

Content analysis of Twitter data has been of interest to the research community for some time. In an analysis conducted more than ten years ago, for example, Stephen Dann (2010) presented a new framework for qualitative tweet content classification which was based on sixteen existing Twitter studies. Studies focused specifically on the analysis of religious discourse on the Twitter platform are, however, less common. A study by Codone (2014) notably explored the ways in which two prominent American church leaders, Rick Warren and Andy Stanley, used Twitter. Both leaders had a high Twitter following (1.4 million and 392,000 followers respectively) and were, therefore, influential figures online. In conducting that analysis, Codone drew on an extant framework for categorising tweets, devised by Naaman, Boase and Lai (2010). Codone concluded that Warren’s most frequent tweet types were:

- ‘random thoughts’
- ‘information sharing’
- ‘self promotion’

While Stanley’s most frequent tweet types were:

- ‘self promotion’
- ‘anecdotes others’
- ‘anecdotes me’
- ‘information sharing’

In a related study, Cheong (2016) explored the ways in which church leaders used Twitter to maintain legitimacy, normality and hierarchy. She reported that the church leaders studied drew on the following five communication practices to achieve these goals:

- Careful selection of account names and bio descriptions
- Strategic promotions
- Encouragement of tweets which promote interactions
- Reporting on behind the scenes and intimate activities related to their work
- The sharing of instructions and/or lessons, often quoting other Christians and/or scripture

These studies are useful in that as well as demonstrating how prominent church leaders use Twitter they also give an indication that the ways in which individuals use Twitter depend very much upon their personal circumstances and the scenario in which they are using their account; a church leader is very likely to use Twitter differently to a member of the public. This insight inspired the design of **P1**, which seeks to explore the ways in which people (not limited to church leaders) use Twitter when discussing church. This will enable **P1** to make an original contribution to knowledge and further academic understanding about Twitter use cases within the religious space.

In a separate article, Holmberg, Bastubacka and Thelwall (2016) presented the results of a study in which they conducted a content analysis of tweets directed at the @God account handle on Twitter. As part of that study the authors designed a coding framework for that particular exercise, providing further evidence that content analyses of tweets on different topics, even within the field of religion, require bespoke coding frameworks. This provided confidence that if **P1** were able to devise a robustly tested qualitative coding framework for church-related tweets, this would be of value to scholars exploring similar tweets in future.

More recently, Karjalainen and Halonen (2019) presented the results of a qualitative analysis of Twitter data in which they considered the content of 937 church-generated tweets (which included re-tweets of posts generated by other Twitter users) posted during 2017. The authors found that, at the time of their

analysis, the key issues discussed were asylum seekers, laws related to equal-marriage and human rights.

In a paper on religious communication and microblogging rituals, Cheong (2010) described various Twitter use cases which apply to the Christian community. These use cases include using Twitter hashtags to share prayer requests, using Twitter-based services to support Christians in praying at set times throughout the day and using carefully-worded twitter posts in order to ‘Twitness’ to non-believers and in doing so, spread the message of faith and create new faith connections.

Previous published studies have reported on work using machine learning to automate qualitative coding of data (e.g. Crowston, Liu, Allen & Heckman, 2010; Scharkow, 2011) and have reported on some of the key challenges related to the use of machine learning to automate qualitative coding (e.g. Chen et al., 2016). In the context of applying machine learning to the qualitative coding of social media data, a plethora of published studies have reported on findings which show potential for real-world impact. Le and Nguyen (2015), for example, proposed an approach for conducting sentiment analysis of Twitter data using Naïve Bayes and Support-Vector Machine. Separate research published the same year by O’Dea, Wan, Batterham, Calexar, Paris and Christensen (2015) explored whether it was possible to identify the most concerning tweets which indicated suicidal intent using human coders and machine learning. They found that both human coders and machine learning were effective in achieving this task, with the machine learning approach correctly identifying 80% of the most concerning tweets. They noted, however, that improvements were required to make this approach suitable in the context of increased data volumes. In a more recent study, Ahmad, Asghar, Alotaibi and Awan (2019) described their work to apply deep learning-based sentiment analysis to code Twitter data as extremist or non-extremist in nature. Studies such as these offer real promise for the potential application of machine learning analysis using Twitter data to address challenges and solve problems. While much attention has been given specifically to the application of machine learning to sentiment analysis challenges using Twitter data, there has been comparatively little previous academic attention specifically focussed on the use of machine learning to automate the process of qualitative thematic content analysis of Twitter data. This gap in the evidence base informed the design of **P2**, which explores whether machine learning algorithms can be used to automate the process of qualitatively coding church-related tweets using the qualitative thematic content analysis framework presented in **P1**.

Published studies have considered the relationship between tweet content and real-life activity. Thelwall, Buckley and Paltoglou (2011), for example, published the results of a study exploring the relationship between tweet content and events, finding that the results provided strong evidence that popular events tended to be associated with an increase in the strength of negative sentiment. In a separate paper,

Suero Montero, Haddad, Mozgovoy and Bechikh Ali (2016) presented the results of an analysis of Twitter data which was able to identify emotion spikes and their likely causes, using a dataset comprising tweets posted by Barack Obama, Bill Gates and the Dalai Lama. These findings inspired the design of **P3** which sought to explore whether there was a relationship between tweet content and the level of engagement in offline activity – in this instance, church attendance. In exploring this potential relationship, **P3** draws on SentiStrength – a freely-available sentiment strength detection tool, which has been reported on in the academic literature (Thelwall, Buckley, Paltoglou, Cai and Kappas, 2010).

A number of previous studies have explored the use of the geolocation data contained within tweet metadata. Cheng, Caverlee and Lee (2010) reported the results of an evaluation of a probabilistic approach to identify the city location of Twitter users, based on the content of their tweets. The evaluation found that the approach was able to put “51% of Twitter users within 100 miles of their actual location” (p759). Similar research by Steiger, Westerholt, Resch and Zipf (2015) explored semantic and spatiotemporal clustered tweets, reporting a strong correlation when compared with workplace census data, “being a good indicator and representative proxy for analysing workplace-based activity” (p255). Both of these studies demonstrate the value of using Twitter geolocation data in research; the study by Cheng et al. (2010), in particular, demonstrates that there are known instances where there is a relationship between tweet content and tweet location. This is highly relevant to this doctoral study which utilises the content and geolocation of church-related tweets in a variety of ways to better understand church attendance.

In a study very relevant to this doctoral research, Walther and Kaisser (2013) monitored all tweets from a specific location and analysed those tweets, using machine learning, to identify whether or not clusters of high Twitter activity represented real-world events; they demonstrated that their approach was able to do so with high precision and recall. The findings of that paper provided the theoretical basis on which this doctoral study assumed that it might also be possible to identify community locations (in this case, churches) using geotagged Twitter data – a research question addressed in **P4**.

This theoretical basis was supported by a study published by Frias-Martinez and Frias-Martinez (2014), in which they demonstrated that machine learning could be utilised to identify land use based on tweet content. This provided further reassurance that the research questions, focused on identifying church locations based on the content of geotagged Twitter data, and the proposed approach set-out in **P4** to answer those research questions were valid and had the potential to yield useful results.

2.4 Digital Theology

Although computing technology has been used in various religious contexts since the birth of modern computing in the 1960s, digital theology as a formal field of academic enquiry is relatively new and has recently captured the interest of scholars from a range of disciplines, most notably computer science, theology³, sociology of religion and the social sciences.

Owing to the fact that the field of digital theology is still finding its way, there are a small number of definitions of the field simultaneously in use, each of which presents digital theology in a slightly different light. Steinhart (2012) provided the first notable definition of digital theology:

“Many recent writers have developed a rich system of theological concepts based on the technologies of computation (especially artificial intelligence, robotics, digital networks, and virtual reality). This theological system is referred to by some as nerd theology (Kelly, 1999) or apocalyptic AI (Geraci, 2010). I shall refer to it here as digital theology (or digitalism).” (Steinhart, 2012, p. 133)

This definition was crafted by Steinhart mainly as a convenience, to replace two previously used subject names. Digital theology, as described by Steinhart, is very theoretical in nature and focuses very closely on the philosophical concepts of identity, mortality and self. Artificial intelligence plays a key role in this definition and it could well be argued that other labels such as computational theology or computational philosophy might better suit this definition.

Other, more recent, definitions of digital theology have taken a slightly different approach and focus. Kolog, Sutinen and Nygren (2016), for example, proposed the following:

“An integration of technology into the understanding of the concept of God and the nature of religious ideas” (Kolog et al., 2016)

This definition is much simpler than that proposed by Steinhart and is, therefore, more accessible. The use of the word ‘integration’ is of particular interest and importance within this definition, as it suggests that digital theology is about more than using technology to study God or religion; it places technology at the centre of the study of God or religion. A key limitation of this definition, however, is that it focuses only on exploring the role of technology within the study of God or religion;

³ Within this dissertation, theology is considered to refer to the academic discipline which considers “Christian reflection on the ideas of faith” (McGrath, 2018, p. xiii).

it does not, reciprocally, propose that exploring the role of God or religion within the study of technology is a central component of digital theology.

Phillips, Schiefelbein-Guerrero and Kulberg (2019) published a recent paper dedicated to the topic of defining digital theology. Within that paper, they proposed the following definition:

- “DT1: The use of digital technology to communicate or teach theology as a traditional academic subject” (Phillips et al., 2019, p. 37)
- “DT2: Theological research enabled by digitality or digital culture” (Phillips et al., 2019, p. 38)
- “DT3: Intentional, sustained and reflexive theologically-resourced engagement with digitality/digital culture” (Phillips et al., 2019, p. 39)
- “DT4: A prophetic re-appraisal of digitality in the light of theological ethics” (Phillips et al., 2019, p. 39)

While this definition is certainly very comprehensive, there are studies taking place within the field which would not readily fit this proposed definition of digital theology. Examples include research developing virtual and augmented reality worship spaces and research to develop emojis specifically intended for religious discourse. Such examples do not fit this definition as the definition is designed principally with theology in mind and is focused closely on theory and high-level academic concepts. It therefore lends itself less well to studies designed predominantly from a computer science perspective.

It is for that reason that the author crafted and proposed an alternate definition of digital theology with Sutinen (Sutinen & Cooper, 2021, p. 17) in supporting publication **P10**. That definition stated that:

“Digital Theology is the field of study and design at the intersection of computer science and theology/religion, which:

- Applies theological thinking and ethics to the field of digital technology.
- Applies computational thinking, processes and approaches to the field of theology.
- Applies digital technology to the practice and study of theology.
- Facilitates meaning making of faith through digital expression.
- Implements research approaches at the intersection of computer science and theology.”

The intent behind this suggestion was to offer a definition of the field of digital theology which provided coverage for the range of studies, particularly those with a strong basis in computer science, which were not previously in-scope of the extant definitions. While offering this definition from a computer science perspective, it was of vital importance to ensure that it captured the range of studies in the field being led not just by computer scientists but also by theologians, sociologists of religion, social scientists and scholars from other backgrounds. In defining digital theology this way, the idea was to acknowledge that digital theology is a truly multi-disciplinary field which is deeply rooted in practice. It is clear that the work presented in this computer science doctoral dissertation fulfils this proposed definition of digital theology, particularly in that it ‘applies digital technology to the practice and study of theology’ and ‘facilitates meaning making of faith through digital expression’. This doctoral study is unique, therefore, in that while it draws primarily on computer science analyses and approaches (e.g. machine learning, sentiment analysis and geospatial analysis), it also contributes original knowledge to a range of different disciplines which come together in a new, innovative and still emerging field of research and study.

It is worth noting that there are overlaps between the fields of digital theology and digital religion. This thesis considers that digital religion “explores the integration of technology within the phenomenon of religion” (P10: Sutinen & Cooper, 2021, p1) while “the perspective of Digital Theology is that of a given faith and its intellectual conceptualisation as digital representation” (P10: Sutinen & Cooper, 2021, p1). Therefore, this thesis proceeds to consider the doctoral research presented within the context of digital theology, as defined in P10.

Given the breadth of the field of digital theology, the topics covered by the published literature are inherently varied and wide-ranging. Studies include analysis of the extent to which the online activities of the Pope can be considered to form the basis of a Digital Papacy (Campbell & Vitullo, 2019), considerations around engagement with the Bible in a digital age in the context of increased digitisation of texts and libraries (Phillips, 2018; Phillips, 2020) and exploration of *missio Dei* in a digital age (Kurlberg & Phillips, 2020). Amidst this emerging literature, various studies have considered social media in the context of digital theology. Lewis (2018), for example, explored practice of the Christian faith in online spaces, including social media platforms, which can be regarded as online surveilled public spaces. Taking a case study approach, Karjalainen and Halonen (2019) presented the results of a qualitative analysis of the content of church-generated tweets posted during 2017, presenting an indication of key topics of conversation at that time. The research findings presented in this doctoral dissertation build on the existing evidence base on church-related tweets, furthering contextual understanding of social media usage within the field of digital theology.

3 Research Design

3.1 Research Questions

To begin to address the identified shortcomings in existing church attendance datasets, this thesis set out to answer a number of research questions to further knowledge on the applicability of Twitter data to church attendance studies and to contribute to the academic evidence base on the response of the church globally to the COVID-19 pandemic and to consider what this might mean for future studies of church attendance.

Prior to commencing this thesis, the author conducted a small-scale study of a sample of tweets containing the word ‘church’ which were posted from within the London Borough of Camden over a 10-week period (Cooper, 2014). The sample size for that preliminary investigation was very small ($n = 108$), however, this was not deemed to be problematic as the aim of that study was merely to assess the feasibility of content analysis using church-related tweets, rather than to robustly assess the content of discussion contained within church-related tweets. As an output of that preliminary study, the author proposed an initial coding framework which might be of use in future more robust studies on the content of church-related tweets.

As a starting point, therefore, this thesis seeks to repeat that work using a bigger sample of tweets containing the word ‘church’ which were posted from across all London Boroughs over the same 10-week period. The first research question addressed within this thesis can, therefore, be defined as follows:

RQ1: Is the existing qualitative coding framework for church-related tweets fit-for-purpose? [This research question is addressed in **P1**].

Having considered the existing qualitative coding framework and either validated it or proposed a new alternative, this thesis then seeks to apply any resulting coding framework to explore the content of church-related tweets across London over the 10-week time period, in order to answer the following research question:

RQ2: What topics are discussed by social media users in church-related tweets? [This research question is addressed in **P1**].

On completion of the planned qualitative content analysis of church-related tweets, this thesis seeks to place these findings in the context of wider research in the field of sociology of religion, by considering how the lessons learned from this exercise might be more widely applied to future studies in the field, thus, answering the following research question:

RQ3: What can sociologists of religion learn from the content of church-related tweets? [This research question is addressed in **P1**].

Recognising that one of the key limitations of qualitative content analysis is the time and resource-intensive burden which such analysis places on researchers and analysts, this thesis then goes on to explore whether the process of qualitative content analysis of church-related tweets can be automated, addressing the following research question:

RQ4: Can content analysis of church-related tweets be automated using machine learning? [This research question is addressed in **P2**].

To ensure that learning from this analytical exercise is captured to inform future research studies using machine learning to investigate the content of Twitter data, this thesis proceeds to address the following research question:

RQ5: What are the implications of using machine learning to classify church-related tweets? [This research question is addressed in **P2**].

Having explored the content of church-related tweets and investigated potential approaches to automate and, thus, speed-up such analyses, this thesis is then concerned with driving forward the evidence base on the potential uses of such data in studies of church attendance. It seeks, therefore, to empirically investigate whether there is a measurable relationship between the content of church-related tweets and offline church attendance activity by exploring the following research question:

RQ6: Is there a relationship between the net sentiment score of church-related tweets and the presence of church growth in the borough from which the tweets were posted? [This research question is addressed in **P3**].

On the basis that a provisional study (Cooper, 2014), prior to this thesis, had identified that tweets were sometimes posted from *within* church buildings, this thesis goes on to explore whether an innovative approach to analysing church-related tweets might enable researchers to identify church buildings, as expressed in the following research question:

RQ7: Can geolocated Twitter data be used to identify church locations? [This research question is addressed in **P4**].

As the overarching aim of this thesis is to address an identified issue with the quality of existing church attendance data, it is essential that any new approaches/methods explored within this thesis research are presented in the context of improving the quality of church attendance study data. This thesis will, therefore, go on to explore whether any churches identified as part of the work to address **RQ7** were excluded from previously published church attendance studies, thus answering the following research question:

RQ8: Can geolocated Twitter data can be used to uncover previously hidden churches? [This research question is addressed in **P4**].

Recognising that all of the work conducted to address research questions **RQ1** to **RQ8** fits within the emerging field of digital theology, this thesis concludes by presenting the results of highly responsive research conducted in light of the COVID-19 pandemic which had a major impact on churches globally throughout 2020 (and which continues to affect the church at the time of writing this dissertation). To begin to explore the impact of COVID-19 on the church, this thesis will address the following research question:

RQ9: How have churches around the world responded to the COVID-19 pandemic? [This research question is addressed in **P5**].

To frame the learning from research conducted to address **RQ9** against the bulk of the work in this thesis which is focused on supporting studies of church attendance, this thesis will conclude by addressing the following research question, which focuses specifically on church attendance:

RQ10: How has the COVID-19 pandemic affected church attendance at the surveyed churches and what considerations might be required to better monitor this going forward? [This research question is addressed in **P5**].

When considered holistically, it is intended that the answers to these ten research questions will further current knowledge on potential innovative approaches to improving the quality of available church attendance data. It is hoped that this will, in turn, support academics in the design of such future studies.

3.2 Data Collection

This chapter outlines the data collection approach taken during the research summarised within this thesis. Several of the publications made use of a common dataset, the design and structure of which is described here. The research and analysis methods used within the component studies are described in the next chapter, which summarises each publication in turn.

Four of the five publications (**P1** to **P4**) comprising the doctoral thesis described in this dissertation summarised research conducted using a Twitter dataset (referred to hereafter as ‘the thesis dataset’) collected in 2014. Those research studies sought to investigate whether Twitter data could be used to complement traditional research datasets and approaches to better understand church attendance.

The final publication (**P5**) comprising the thesis described in this dissertation summarised research conducted using qualitative survey data to better understand the response of the church to the COVID-19 outbreak which commenced in late 2019 and, throughout 2020, fundamentally increased the use of online technology by churches around the world.

Data collection for the thesis dataset commenced on Sunday 20 April 2014, which was Easter Sunday that year. The thesis dataset comprises sample data collected using the Twitter Application Programming Interface⁴, with the assistance of Mozdeh⁵, a free software tool designed to support the analysis of social media text data.

Data was collected each Sunday over a 23-week period. Each tweet in the sample was posted on a Sunday within the data collection date range, was written in the English language, contained the word ‘church’ and was posted from within a 60-kilometre radius of an arbitrary point in central London.

During the data capture phase, a total of 42,661 tweets were collected. Of the 42,661 tweets, 2,605 tweets, posted by 1,765 unique users⁶, contained geodata (i.e. WGS84⁷ values for the latitude and longitude of the point from which the tweet was posted) and were retained for analysis. The remaining tweets which did not contain geodata were discarded from the dataset. The low proportion (6.1%) of tweets which contained geodata was consistent with previous studies – Cheng, Caverlee and Lee (2010), for example, reported that their analysis of a random sample of over one million Twitter users identified that fewer than 0.42% of all tweets contained geodata. It is outside the scope of this thesis to explore why, despite being a low percentage, the 6.1% value was actually much higher than comparator values in other studies such as Cheng, Caverlee and Lee (2010). However, one reason might be to do with the motives which drive users to activate/deactivate their geodata. Further research in this regard would be of real value to the research community.

Owing to the radius search approach used, some of the tweets within the dataset were posted from outside London, which for the purposes of this dissertation is considered to be the area within the Greater London Authority. The issue of non-

⁴ See: <https://dev.twitter.com/> (last accessed: 29 December 2020).

⁵ See: <http://mozdeh.wlv.ac.uk/> (last accessed: 29 December 2020).

⁶ Determined using the AuthorURL field within the Twitter data.

⁷ The World Geodetic System 1984.

London tweets was addressed during the data cleaning and analysis phase of the various research studies and associated publications.

The first three publications (**P1** to **P3**) used a subset of the thesis dataset, comprising tweets collected across the first 10 weeks of data collection. In total, there were 1,234 tweets containing geodata within this 10-week subset. This represented 6.2% of the 19,894 tweets collected during that time period. During the data cleaning phase of the studies described in these publications, the MapIt Application Programming Interface⁸ was used to identify which London Borough each tweet was posted from. In total, 230 tweets were at this stage identified as having been posted from outside London and were accordingly discarded from the dataset. This was expected, due to the radius search approach used during data collection. In total, 1,004 tweets, posted by 723 unique users⁹, were retained for analysis in the first three publications.

The fourth publication (**P4**) required a larger dataset than the initial thesis publications, owing to the research approach being applied, which required multiple tweets to be posted from the same location to identify potential church locations. That study commenced, therefore, with the 2,605 tweets containing geodata in the 23-week dataset. During the analysis phase, the locations of the potential churches identified were investigated and those deemed to be outside of London were discarded at that point in the analysis.

⁸ See: <http://mapit.mysociety.org/> (last accessed: 29 December 2020).

⁹ Determined using the AuthorURL field within the Twitter data.

4 Component Studies: Summary and Results

Having detailed the data collection approach employed by the studies comprising this thesis, this chapter will consider each thesis publication in turn, detailing the aim and design of each study alongside the research results.

4.1 P1: Understanding London's Church Tweeters: A Content Analysis of Church-Related Tweets Posted from a Global City

The main aim of this study was to replicate a preliminary study conducted previously by the author (Cooper, 2014), this time using a bigger dataset comprising a sample of tweets collected from all London boroughs across the same 10-week period. This study, therefore, made use of data from the first 10 weeks of the thesis dataset described in the previous chapter. As described, 1,004 geolocated tweets which were posted from within London were taken forward for analysis.

The author led the qualitative coding of tweets within this study, using the proposed coding framework from the previous preliminary study (Cooper, 2014), as shown in Table 1, as a guide.

Table 1: Thematic coding labels used in a previous study to categorise 108 church-related tweets from the London Borough of Camden (Cooper, 2014).

| Thematic coding label |
|--|
| Discussion about church attendance |
| Discussion about church service content |
| Discussion about theology, the Bible of belief |
| General church discussion |
| Miscellaneous |
| Non-church discussion |
| Unknown theme |

The author worked through each tweet, one at a time, allocating as many labels from the coding framework to each tweet as was deemed necessary, such that each tweet possessed one to many coding labels. Where the author felt that a new coding label might be valuable in qualitatively describing a particular tweet, a note was made of this, to permit discussion with the second coder during a later stage of the analysis when considering the appropriateness of the proposed coding framework.

On completion of the initial coding phase, a second coder (who was also the second author of the publication) was provided with the full dataset, including the allocated coding labels. Their role was to work through each tweet, one at a time, and assess whether the allocated coding labels were appropriate. Where the second coder felt that a new coding label might be of value, they also made a note of this, to permit discussion with the primary coder.

Once the secondary coder had completed their analysis, both coders worked together to discuss, in turn, each tweet which they had coded differently. For each such tweet, the coders worked to reach consensus on the most appropriate coding labels. This was achieved, resulting in a coded dataset which both coders were content with. Following this coding exercise, the coders also considered the appropriateness of the proposed coding framework, in light of the potential additional coding labels they had noted down. The coders reached agreement that the proposed coding framework was appropriate, subject to the addition of only one new coding label – Discussion about Hillsong. This addition was not very surprising; the preliminary study of church-related tweets posted from the London Borough of Camden had remarked that “at least 38 of the 108 tweets (approximately 35%) within the sample appeared to be related to Hillsong church” (Cooper, 2014). This adjustment to the proposed coding framework was, therefore, considered to be minor.

The study was, therefore, able to provide an affirmative answer to **RQ1**, by demonstrating that the existing qualitative coding framework for church-related tweets is fit-for-purpose.

The coding framework and associated tweet volumes identified within the study are summarised in Table 2.

Table 2: Thematic coding labels with the volume of tweets captured by each thematic coding label and percentage of total tweet volume for 1,004 tweets from all London boroughs (P1: Cooper, Mann, Sutinen & Phillips, 2021).

| Thematic coding label | Volume of tweets captured by each thematic coding label and percentage of total tweet volume | |
|--|---|-----|
| Discussion about church attendance | 605 | 60% |
| Discussion about church service content | 72 | 7% |
| Discussion about theology, the Bible or belief | 82 | 8% |
| General church discussion | 339 | 34% |
| Miscellaneous | 201 | 20% |
| Non-church discussion | 85 | 8% |
| Discussion about Hillsong | 51 | 5% |
| Unknown theme | 47 | 5% |

Having presented the results holistically, the study went on to break-down the label allocations by Borough, presenting the results for each coding label. Analysis of the results at that level of detail revealed that there appeared to be a broad range of coverage for each coding label across each London Borough. The study found that with the exception of the Unknown coding label, there were very few instances of labels with zero tweets against them in the different London boroughs, thus demonstrating that church-related tweets are diverse in content across all of London’s boroughs and that the coding framework used is of relevance and benefit to studies exploring church-related tweets in all of London’s boroughs.

To assist readers in interpreting these results and grasping these topics, the paper also provided definitions for each of the coding labels used within the framework:

- Discussion about church attendance (DCA): Tweet content which explored presence in church, e.g. indication that the tweeter, or tweet subject, had previously attended a church, was present in a church, might attend a church in the future or had made a decision not to attend a church.
- Discussion about church service content (DCSC): Tweet content which explored the content of a church service, e.g. a sermon, acts of worship or social interactions taking place within a church service.
- Discussion about theology, the Bible or belief (DTBB): Tweet content which explored religious belief or theory, the nature of God or interpretation of religious texts.
- General church discussion (GCD): Tweet content which contained discussion of church other than church attendance or church service content, e.g. church buildings, church bells or the role of church in society.

- **Miscellaneous (MISC):** Tweet content related to church but fairly unique and varied in nature and volume, such that the content did not naturally align to the other thematic coding labels used within the framework.
- **Non-church discussion (NCD):** Tweet content which was un-related to the Christian church, e.g. tweets related to the surname Church or bars/venues called church.
- **Discussion about Hillsong (HLSG):** Tweet content which was specifically related to Hillsong church.
- **Unknown theme (UNK):** Tweet content which could not be thematically coded, due to coder inability to understand/interpret the nature of the discourse.

The study was, through providing original new insight into the content of church-related tweets, able to provide an answer to **RQ2**, enabling academics in the fields of digital theology and sociology of religion to grasp the types of church-related topics which are discussed on Twitter.

The study concluded by exploring the learning points identified through analysis of the content of 1,004 church-related tweets. While the study was careful not to suggest, based on the sample size, that the findings were generalisable across all future church-related tweets posted within the Greater London Authority, it did note that there were useful lessons which could be of benefit to future work in the field. The study noted, firstly, that digital theologians and sociologists of religion now have a baseline for the content of church-related tweets posted from within London. While this is interesting in and of itself, it adds further value to the academic community by permitting future repeat studies to investigate changes against this baseline; across the whole of the city or within particular boroughs. Being able to monitor changes in the content of church-related tweets across time or location could be interesting, for example, as the church within the UK grapples with key theological issues. It might be valuable, for instance, for a repeat study to investigate whether the types of topics discussed in church-related tweets change as church thinking of issues of sexuality and gender progress and receive increased public and media attention. Secondly, the study noted that the qualitative coding of over a thousand church-related tweets paved the way for researchers in future to make use of other computer science and statistical methods of analysis when examining the dataset, benefitting from an additional datapoint for each tweet. Finally, the study noted that the fact that a prominent church-related discourse had been identified at all on Twitter was a useful finding for digital theologians and sociologists of religion; particularly at a time where the secularisation thesis posits that religion is of decreasing importance to a rapidly modernising society.

In making these observations and raising these points, this study has been able to address **RQ3**, providing insight on some of the key lessons which digital theologians and sociologists of religion can learn from the content of church-related tweets.

4.2 P2: Exploring the Use of Machine Learning to Automate the Qualitative Coding of Church-Related Tweets

The main aim of this study was to explore whether the qualitative coding described in **P1** could be automated using machine learning. This was an important research objective as the discussion within **P1** had clearly highlighted the benefits of being able to qualitatively code church-related tweets, however, the manual process employed, which required two human coders, was both time and resource-intensive and was likely to become insurmountable for future studies using even bigger datasets.

To begin to test the efficacy of machine learning algorithms at performing this task, **P2** made use of the same 10-week subset of the thesis dataset utilised in **P1**. That dataset comprised 1,004 geolocated tweets which contained the word ‘church’ and were posted from within London.

Having defined machine learning and applied the definition of machine learning to the use case of automating the qualitative coding of tweets, the paper went on to discuss how machine learning methods are typically divided into two types – classification (also known as supervised) approaches which utilise a partially coded dataset to apply rules in an attempt to try and correctly code the rest of the dataset and clustering (also known as unsupervised) approaches which group datapoints based on an inferred similarity or proximity between datapoints. The former approach is of particular use in studies seeking to quickly and automatically label large datasets. The latter approach is of particular use in studies seeking to divide large datasets into subgroups or subcategories based on similarity between datapoints. As the thesis dataset had already been coded in **P1** using classes identified in a previous preliminary study, this study focussed on the former approach, classification, to explore whether the qualitative coding of church-related tweets could be automated using machine learning.

To do so, the study conducted machine learning analyses using WEKA¹⁰ (Waikato Environment for Knowledge Analysis), a software package designed to facilitate machine learning analysis. The machine learning analyses conducted were Support-Vector Machine, J48 Decision Tree and Naïve Bayes. These classifiers were

¹⁰ See: <https://www.cs.waikato.ac.nz/ml/weka/> (last accessed: 30 December 2020).

selected based on their application in previous studies and their good performance in text classification tasks (see, for example: Dharmadhikari, Ingle & Kulkarni, 2011). **P2** defined each of these classifiers as follows:

Support-Vector Machine

“Support-Vector Machine (SVM) is one of the widest-used supervised machine learning classification algorithms. The algorithm works by using transformations to analyse data points (known as vectors) in a multi-dimensional vector space and then optimally dividing those vectors, based on the position of extreme vectors, using hyperplanes. The hyperplanes are the decision boundaries of the data points/vectors. Data points on either side of the hyperplane are considered a class. However, to separate two classes of data points, there are many possible hyperplanes that could be chosen. The objective is to find a plane that has the maximum margin between data points of both classes.” (**P2**: Cooper, Kolog & Sutinen, 2019, p. 152).

J48 Decision Tree

“J48 is a decision-tree classifier for supervised classification problems. The algorithm is used to generate a decision tree developed by Ross Quinlan. As the J48 decision tree extends the Iterative Dichotomiser 3 algorithm (ID3), it additionally works by accounting for missing values, decision tree pruning, continuous attribute value ranges and derivation of rules, among other approaches. The algorithm generates rules from which the particular identity of that data is generated. The objective is the progressive generalization of a decision tree until it gains equilibrium of flexibility and accuracy (Kaur & Chhabra, 2014).” (**P2**: Cooper, Kolog & Sutinen, 2019, p. 152).

Naïve Bayes

“Naïve-Bayes utilizes a set of supervised learning algorithms based on applying Bayes’ Theorem (a formula for calculating conditional probabilities) with the “naïve” assumption of conditional independence between every pair of features given the value of the class variable (Taheri, Mammodov and Bagirov 2011)” (**P2**: Cooper, Kolog & Sutinen, 2019, p. 152-153).

The results of the three classifiers are presented in Tables 3 to 5.

Table 3: Support-Vector Machine classifier evaluation [overall shows weighted average] (P2: Cooper, Kolog & Sutinen, 2019, p. 153).

| Themes | Recall | Precision | F-measure |
|----------------|--------------|--------------|--------------|
| DCA | 0,773 | 0,886 | 0,839 |
| DCSC | 0,655 | 0,428 | 0,588 |
| DTBB | 0,474 | 0,437 | 0,453 |
| GCD | 0,662 | 0,659 | 0,697 |
| HLSG | 0,438 | 0,423 | 0,617 |
| MISC | 0,509 | 0,525 | 0,494 |
| NCD | 0,943 | 0,415 | 0,698 |
| UNK | 0,543 | 0,409 | 0,440 |
| Overall | 0.686 | 0.677 | 0.700 |

As can be seen from Table 3, the Support-Vector Machine classifier met the acceptable threshold of 70% for the overall F-measure¹¹, however, the overall recall¹² and precision¹³ measures both fell below 70%. As can be seen from the results, the precision of this classifier was varied – from 88.6% for discussion about church attendance to 40.9% for the unknown theme.

¹¹ The F-measure is the harmonic mean of Precision and Recall. Therefore, $F\text{-measure} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$.

¹² Recall is a measure of the proportion of positives which were actually classified. Therefore, $\text{Recall} = \text{True Positives} / (\text{True Positives} + \text{False Negatives})$.

¹³ Precision is a measure of the proportion of positive classifications which were correctly identified. Therefore, $\text{Precision} = \text{True Positives} / (\text{True Positives} + \text{False Positives})$.

Table 4: J48 Decision Tree classifier evaluation [overall shows weighted average] (P2: Cooper, Kolog & Sutinen, 2019, p. 154).

| Themes | Recall | Precision | F-measure |
|----------------|--------------|--------------|--------------|
| DCA | 0.902 | 0.785 | 0.836 |
| DCSC | 0.429 | 0.511 | 0.446 |
| DTBB | 0.413 | 0.428 | 0.418 |
| GCD | 0.726 | 0.669 | 0.695 |
| HLSG | 0.639 | 0.714 | 0.653 |
| MISC | 0.489 | 0.536 | 0.507 |
| NCD | 0.733 | 0.938 | 0.812 |
| UNK | 0.420 | 0.511 | 0.434 |
| Overall | 0.719 | 0.688 | 0.700 |

Table 4 shows that the J48 Decision Tree classifier met the acceptable threshold of 70% for the overall F-measure and overall recall, however, the overall precision measure fell below 70%. The highest precision score for this classifier was 93.8% for non-church discussion, while the lowest precision score for this classifier was 42.8% for discussion about theology, the Bible and belief.

Table 5: Naïve Bayes classifier evaluation [overall shows weighted average] (P2: Cooper, Kolog & Sutinen, 2019, p. 155).

| Themes | Recall | Precision | F-measure |
|----------------|--------------|--------------|--------------|
| DCA | 0.839 | 0.851 | 0.845 |
| DCSC | 0.646 | 0.567 | 0.599 |
| DTBB | 0.663 | 0.567 | 0.604 |
| GCD | 0.632 | 0.693 | 0.659 |
| HLSG | 0.960 | 0.776 | 0.889 |
| MISC | 0.627 | 0.601 | 0.613 |
| NCD | 0.602 | 0.754 | 0.658 |
| UNK | 0.480 | 0.533 | 0.500 |
| Overall | 0.716 | 0.731 | 0.722 |

Table 5 shows that the Naïve Bayes classifier exceeded the acceptable threshold of 70% for recall, precision and the F-measure. The Naïve Bayes classifier also had the

smallest precision range; discussion about church attendance had the highest precision score at 85.1% while the unknown theme had the lowest precision score at 53.3%. As Naïve Bayes is documented for requiring only a small amount of training data to produce reliable results (Dharmadhikari, Ingle & Kulkarni, 2011, p. 163), it might be that the sample size of the thesis dataset lent itself best to this classifier.

The positive results obtained for the Naïve Bayes classifier demonstrate that it is possible to automate the qualitative coding of church-related tweets. This, therefore, provides an affirmative response to **RQ4**.

Having evaluated the three classifiers tested, **P2** proceeds to address **RQ5** by exploring the implications of the use of machine learning to automate the qualitative coding of church-related tweets.

The paper notes that the automation of the qualitative coding of church-related tweets paves the way for future studies to repeat qualitative content analyses of this type, at scale and at speed. This lowers the barrier to entry for studies similar in design to **P1** and, therefore, offers the potential for researchers in future to be able to track how the themes discussed in church-related tweets change over time, or even across locations. Researchers conducting future studies should be mindful, however, that the themes discussed, and therefore the coding labels used in analysis, might also change over time, requiring a degree of flexibility when planning future studies. The paper also notes that if research approaches were to be able to be designed to facilitate statistical analysis of church-related Twitter data to predict church growth/decline, the rapid qualitative coding of those tweets would make such studies quicker and easier to conduct. This has the potential to add tangible benefits to future studies of church growth, which might need to draw upon innovative approaches to mitigate against the limitations posed by poor quality church attendance data.

The paper concluded by discussing some of the limitations associated with the study – for example the requirement for fully coded test and training data and the reliance upon a tested qualitative coding framework. The paper noted that it might, in future, be possible to mitigate such challenges through the use of clustering (unsupervised learning) approaches and encouraged future research in this space to test this idea. Separately, the paper observed that the process of conducting the study described had demonstrated the importance of being able to form multi-disciplinary teams to conduct digital theology research, to ensure that such studies are able to draw upon a range of quantitative and qualitative skills as well as an understanding of the sociological and theological context in which the studies are set. Finally, the paper concluded by noting that although the study had focused on church-related tweets, the findings were not restricted to a digital theology use case; the results indicated that it was possible to automate the qualitative coding of tweets more generally, therefore offering insight to scholars conducting research in a range of

wider fields. While machine learning is already widely applied across a variety of fields to address problems such as sentiment analysis of Twitter data, it is currently less frequently used to investigate the topics of conversation expressed in tweets. This dissertation therefore demonstrates a further use case which might add value outside of the field of digital theology.

In considering these points and drawing on new insight, **P2** addressed **RQ5**, by making clear the implications of the use of machine learning to automate the qualitative coding of church-related tweets.

4.3 P3: Assessing the Possible Relationship Between the Sentiment of Church-Related Tweets and Church Growth

The main aim of this study was to assess, using inferential statistics, whether a relationship existed between the net sentiment of church-related tweets and the presence of church growth in the London boroughs from which the tweets were posted. This was an important question to ask, as a positive result in this regard would indicate a potential future use of datasets comprising church-related tweets, to predict the presence, or absence, of church growth in areas or regions where church attendance data is not readily available or is not of sufficient quality to draw reliable conclusions.

To approach this task, this paper made use of the same 10-week subset of the thesis dataset utilised in **P1** and **P2**, which comprised 1,004 geolocated tweets which contained the word ‘church’ and were posted from within London.

Within this study, sentiment was considered to be the polarity (positivity and negativity) expressed within the content of the church-related tweets analysed. To identify the sentiment of each tweet, the dataset was run through SentiStrength (Thelwall, Buckley, Paltoglou, Cai and Kappas, 2010), a freely-available sentiment analysis tool, designed to identify sentiment within short informal texts, along with an indication of the strength of that sentiment¹⁴. At the core of the SentiStrength tool is a sentiment word strength list, which is supported by a training algorithm which optimises the strength of those sentiment words. A number of additional features, such as spelling correction and handling of repeated letters ensure that SentiStrength is well-suited to the analysis of the very informal content styles

¹⁴ To provide an example use case of SentiStrength outside the context of the analysis of church-related tweets, the tool was used to dictate the colours in which the London Eye was lit up during the 2012 London Olympic Games. See: <http://www.https://www.telegraph.co.uk/technology/news/9408783/Happy-Olympic-tweeters-to-light-up-London-Eye.html> (last accessed: 10 April 2021).

which are typically encountered in social media posts (Thelwall, Buckley, Paltoglou, Cai and Kappas, 2010). SentiStrength provides two scores for each tweet: a score for the magnitude of positive sentiment (in the range 1 to 5) and a score for the magnitude of negative sentiment (in the range -1 to -5). The author summed the two scores provided for each tweet to arrive at a score expressing the net sentiment of each tweet (in the range -4 to 4), which indicated whether the net sentiment of each tweet was negative, neutral or positive, with an indication of the magnitude of that net sentiment. It is important to remember when considering and working with net sentiment scores, that they can obfuscate a level of detail within the textual content of a tweet. To illustrate this point – a particular tweet might, for example, have a net sentiment score of 3, however that tweet might have contained both positive and negative sentiments, with the positive sentiment outweighing the negative sentiment. Net sentiment scores do not, therefore, provide a perfect description of tweet sentiment but do allow the analysis of large datasets on the basis of considering whether a tweet was *mostly* negative, neutral or positive.

The church growth data used in this study was taken from the 2012 London Church Census (Brierley, 2013) and the measure used was “the rate of growth between 2005 and 2012 in the number of people within each borough who attend church” (P3: Cooper, 2017, p. 43). These growth rates are presented in Table 6.

Table 6: Percentage increase/decrease in church attendance by London borough. (P3: Cooper, 2017, p. 44).

| Borough | Borough % Growth 2005 to 2012 |
|------------------------|-------------------------------|
| Barking and Dagenham | 31% |
| Barnet | 15% |
| Bexley | 11% |
| Brent | 5% |
| Bromley | 20% |
| Camden | 6% |
| City of London | 24% |
| Croydon | 24% |
| Ealing | 17% |
| Enfield | 6% |
| Greenwich | 31% |
| Hackney | -10% |
| Hammersmith and Fulham | 6% |
| Haringey | 28% |
| Harrow | 1% |
| Havering | -5% |
| Hillingdon | 11% |
| Hounslow | 15% |
| Islington | 5% |
| Kensington and Chelsea | 5% |
| Kingston upon Thames | 0% |
| Lambeth | 60% |
| Lewisham | 15% |
| Merton | 30% |
| Newham | 33% |
| Redbridge | 4% |
| Richmond upon Thames | 11% |
| Southwark | 52% |
| Sutton | 13% |
| Tower Hamlets | 27% |
| Waltham Forest | 18% |
| Wandsworth | 17% |
| Westminster | 17% |

Note: Standard deviations are not included in the published data.

Source: Brierley (2013: 66–67).

Having obtained net sentiment scores and church growth data, the mean average tweet net sentiment score was calculated for each London Borough, as presented in Table 7.

Table 7: Mean average net sentiment score of church-related tweets, by London borough. (P3: Cooper, 2017, p. 45).

| Borough | No. of Tweets in Sample | Average Sentiment Score | Standard Deviation |
|------------------------|----------------------------|----------------------------|--------------------|
| Barking and Dagenham | 56 | 0.39 | 1.00 |
| Barnet | 25 | 0.28 | 0.84 |
| Bexley | 17 | 0.41 | 1.23 |
| Brent | 24 | 0.17 | 0.82 |
| Bromley | 26 | 0.73 | 1.19 |
| Camden | 108 | 0.36 | 1.12 |
| City of London | 16 | 0.69 | 1.14 |
| Croydon | 33 | 0.48 | 0.94 |
| Ealing | 23 | 0.13 | 1.29 |
| Enfield | 32 | 0.13 | 0.75 |
| Greenwich | 49 | 0.33 | 1.05 |
| Hackney | 42 | 0.14 | 1.10 |
| Hammersmith and Fulham | 9 | 0.44 | 1.01 |
| Haringey | 23 | 0.17 | 1.23 |
| Harrow | 12 | 0.50 | 0.67 |
| Havering | 22 | -0.05 | 1.25 |
| Hillingdon | 18 | 0.56 | 1.25 |
| Hounslow | 15 | 0.53 | 0.92 |
| Islington | 30 | 0.00 | 1.17 |
| Kensington and Chelsea | 29 | 0.45 | 0.99 |
| Kingston upon Thames | 12 | -0.17 | 0.94 |
| Lambeth | 46 | 0.46 | 1.03 |
| Lewisham | 41 | 0.41 | 0.89 |
| Merton | 26 | 0.00 | 0.57 |
| Newham | 28 | 0.00 | 1.02 |
| Redbridge | 16 | 0.38 | 0.72 |
| Richmond upon Thames | 14 | 0.14 | 1.10 |
| Southwark | 73 | 0.27 | 1.16 |
| Sutton | 6 | 0.83 | 1.33 |
| Tower Hamlets | 21 | 0.29 | 1.06 |
| Waltham Forest | 19 | 0.42 | 0.61 |
| Wandsworth | 25 | 0.28 | 1.21 |
| Westminster | 68 | 0.51 | 0.95 |
| | 1,004 | | |

To further prepare the dataset for analysis, the tweets were divided into two groups – those posted from areas of church attendance decline or stagnation (76 tweets) and those from areas of church growth (928 tweets). Of tweets posted from areas of church attendance decline or stagnation, the average tweet net sentiment score was 0.04 (SD = 1.11). Of tweets posted from areas of church growth, the average tweet net sentiment score was 0.35 (SD = 1.03).

To compare the mean average net sentiment of tweets posted from the two groups, an independent-samples t-test was conducted. This test identified a statistically significant difference (at the $p < 0.05$ level) between the mean average net sentiment of tweets posted from the two groups, such that $t(1,002) = -2.51$, $p = 0.01$. These results provide an affirmative response to **RQ6**, demonstrating that there is a relationship between the net sentiment score of church-related tweets and the presence of church growth in the borough from which the tweets were posted.

To build on these findings, this paper went on to conduct further analysis, to explore whether a relationship existed between the net sentiment of church-related tweets and the rate of church growth in the borough from which the tweets were posted. To conduct this analysis, the tweets were re-grouped into 20% growth bands and mean average net sentiment scores were calculated for each growth band, as shown in Table 8.

Table 8: Mean average net sentiment score of church-related tweets, by 20% growth band. (P3: Cooper, 2017, p. 46).

| Borough | No. of Tweets in Sample | Average Sentiment Score | Standard Deviation |
|-------------------|-------------------------|-------------------------|--------------------|
| -19% to 0% Growth | 76 | 0.04 | 1.11 |
| 1% to 20% Growth | 557 | 0.38 | 1.02 |
| 21% to 40% Growth | 252 | 0.30 | 1.01 |
| 41% to 60% growth | 119 | 0.34 | 1.11 |
| | 1,004 | | |

A one-way ANOVA (analysis of variance) test was conducted, to compare the mean average net sentiments of tweets between the different growth bands. The test did not detect a statistically significant difference (at the $p < 0.05$ level) between the mean average net sentiments of tweets in the different growth bands, such that $f(3, 1,000) = 2.42$, $p = 0.06$. Therefore, the paper concluded that there was insufficient evidence to assert that there is a relationship between the average net sentiment of church-related tweets and the rate of church growth in the borough from which the tweets were posted.

A key limitation of this study was that it was not possible to understand, based on the research methods used, *why* a relationship existed between tweet sentiment and church growth. Therefore, the analysis section of the paper concluded by noting that the identification of a relationship between the net sentiment of church-related tweets and the *presence*, but not the *rate*, of church growth in the borough from which the tweets were posted warrants further future research to better understand the factors underlying the relationship identified.

Although it was possible to obtain a statistically significant result in addressing **RQ6**, the low sample size of tweets posted from areas of church attendance decline or stagnation (76 tweets) was another limiting factor of this study. Therefore, there would be merit in future repeat studies which are designed to ensure higher sample sizes across all church attendance groups, in order to verify the results of this study. Furthermore, given the low sample size of tweets posted from areas of church attendance decline or stagnation, it would also be interesting for future studies to explore whether there is a relationship between the presence of church growth and the volume of church-related tweets.

The results and conclusions presented in this paper furthered academic understanding within the field by demonstrating that there is potential in future work to explore whether social media analysis metrics might be able to supplement missing data and/or low-quality data in future church attendance studies. The identification of a relationship between online and offline activity more generally within the context of churches and church attendance is also of wider interest in the field of digital theology, as scholars increasingly seek to understand how online data might be used to explore offline activities, behaviours and beliefs.

4.4 P4: Using Geotagged Twitter Data to Uncover Hidden Church Populations

The main aim of this study was to investigate whether church-related tweets could be used to identify the location of churches and, if so, whether any such churches discovered had been previously excluded from the 2012 London Church Census (Brierley, 2013). This study was very closely focused on the challenge of improving the quality of church attendance data; the ability of researchers to quickly analyse freely available Twitter data to identify churches for inclusion in future studies has the clear potential to improve study coverage and increase the accuracy of results obtained.

To approach this problem, this study made use of the full thesis dataset, comprising 2,605 geolocated tweets which contained the word ‘church’ and were posted from locations within close proximity to London over a 23-week period. Whereas all of the tweets contained within the 10-week subset of data used in **P1** to **P3** were posted from within a London borough (as verified using the MapIt Application Programming Interface), some of the tweets within the dataset used for this study were posted from outside of London. This is because of the radius search technique used by the Twitter Application Programming Interface. To mitigate this, the locations of potential churches identified in the analysis stage were considered to ensure they fell within a London borough, before being considered for further investigation. This enabled clarity within the results that all churches identified were, indeed, located within London.

The analysis approach taken by this study relied upon the application of several mixed-method techniques. To commence the analysis the WGS84 geolocation data associated with each tweet were rounded to four decimal places. The tweets were then grouped together, to identify groups of tweets which were posted from the same location. On the basis that rounding the WGS84 geolocation data to four decimal places gives approximately 11.1 meters¹⁵ of error and that mobile phone GPS error was believed to be around 8.5 meters (Zandbergen & Barbeau, 2011), this meant that tweets with matching geolocation values during this stage of the analysis should have been posted from within approximately 20 meters of each other. Given typical church size, this distance was considered to be appropriate for the analysis being conducted.

In total, 211 geolocations were identified during this stage of the analysis as containing multiple tweets during the 23-week period. These tweets were then qualitatively analysed, one location at a time, to consider the tweet content and assess whether it was possible to infer a possible church location. During this analysis, groups of tweets were labelled either as a 'possible church location' or a 'location of no further interest'. For example, the location of a tweet which stated "I'm at St George's Catholic Church in Brent, England" (Posted at 12:10 on 14 September 2014) was recorded as being posted from a possible church location, while the location of a separate tweet which stated "Hozier - Take Me To Church @BBCR1 via BBC iPlayer Radio" (Posted at 12:28 on 14 September 2014) was recorded as being a location of no further interest. As a result of this qualitative process of analysis, 78 of the locations were determined to be possible church locations while 133 of the locations were determined to be of no further interest.

As discussed earlier in this chapter, it was to be expected at this point that some of the possible church locations identified would have been outside of London, owing to the arbitrary centre point within London which was selected for radius searching when gathering tweets using the Twitter Application Programming Interface. The 78 possible church locations were, therefore, plotted onto a map at this stage in the analysis to identify any locations outside of London. This review found that 12 of the possible church locations were outside of London; these locations were discarded from the dataset. The remaining 66 possible church locations which were determined to be within London were accordingly taken forward to the next stage of the analysis.

¹⁵ "The distance between the North Pole and the Equator is 10,000,000 meters. This distance can be divided by 90 degrees, to calculate the approximate distance of each degree – 1,111,11.1m. Dividing this distance again by 10,000, gives 11.1m for four decimal places. As such, a tweet with geodata containing four decimal places has an error of 11.1m." (P4: Cooper, 2018, p. 147).

The 66 remaining possible church locations were considered further, to identify whether churches could be identified at those locations. During this second round of qualitative analysis, the content of each tweet was inspected again to pull out any church names which were included. In addition, desk-based research was conducted on each location, using Google Street View¹⁶ and web-based searching¹⁷, to identify evidence of churches at the locations being considered. During this stage of investigation, 24 possible church locations were discarded from the analysis for various reasons, including: “failing to identify a church, insufficient information to suggest the presence of a church and identifying duplicate locations for the same church” (P4: Cooper, 2018, p. 140). In total, 42 locations were confirmed as actual church locations, therefore affirmatively answering RQ7. The 42 churches identified in this analysis are presented in Table 9.

¹⁶ See: <https://www.google.co.uk/streetview/> (last accessed: 9 January 2021).

¹⁷ Mindful that not all churches look and feel like churches (e.g. some churches could meet in houses or public social spaces), web-based searching was vital to identify evidence of churches at the locations being considered.

Table 9: The 42 churches identified in this study using geolocated Twitter data. (P4: Cooper, 2018, p. 141).

| Church Name | Postcode |
|---|-----------------|
| All Souls Church | W1B 3DA |
| Bow church | E3 3AH |
| Christ Church East Greenwich | SE10 9EQ |
| Christ Church Radlett | WD7 7JJ |
| Christ Church Spitalfields | E1 6LY |
| Christ Church with St Phillip | KT4 8LG |
| Christ Church, Mermaid Theatre | EC4V 3DB |
| Dartford Community Church | DA1 2HW |
| Hillingdon Park Baptist Church | UB10 9LS |
| Hillsong Bermondsey | SE16 3LP |
| Hillsong Central London, Theatre Royal, Drury Lane | WC2B 5JF |
| Hillsong Central London, Dominion Theatre | W1T 7AQ |
| Holy Trinity Brompton | SW7 1JA |
| Jesus House for all the nations | NW2 1LT |
| King's church Walton | KT12 3JB |
| Macedonian Orthodox Church | W1D 4NQ |
| Regeneration Church | RM2 6DH |
| Rivergate Church | IG11 0FJ |
| Sacred Heart Church | SW19 4LU |
| Sacred Heart Church | NW6 4PS |
| St Alfege Church | SE10 9BJ |
| St Andrew's Church Watford | WD17 4PY |
| St Bride's church | EC4Y 8AU |
| St Cedds Church | RM8 2HQ |
| St George's Catholic Church | HA0 2QE |
| St James Roman Catholic Church | W1U 3QY |
| St John the Baptist Greek Orthodox Church | N8 0LY |
| St John-at-Hackney | E8 1GY |
| St John's Wood Church | NW8 7PF |
| St Margaret's | SW1P 3JX |
| St Martin's Church | HA4 8DG |
| St Mary Abbots Church | W8 4LA |
| St Mary the Virgin Church | EN4 8XD |
| St Mary with All Souls | NW6 4SN |
| St Matthew's Church Brixton | SW2 1JF |
| St Pancras Old Church | NW1 1UL |
| St Paul's Cathedral | EC4M 8AD |
| St Paul's church | W13 9XW |
| Temple church | EC4Y 7BB |
| The Redeemed Christian Church of God, Victory House | SE17 1TJ |
| United Reformed Church Hounslow | TW3 1UF |
| Westminster Abbey | SW1P 3PA |

To identify which, if any, of the 42 churches identified by this study were excluded from the 2012 London Church Census (Brierley, 2013), the details of the churches set out in Table 9 were provided to Dr Peter Brierley, so that he could compare those churches with the list of churches included in the 2012 London Church Census. During this stage of the analysis, nine of the 42 churches identified by this study were determined not to have been included in the 2012 London Church Census, therefore affirmatively answering **RQ8**. The nine churches identified by this study as having been excluded from the 2012 London Church Census are presented in Table 10.

Table 10: The nine previously excluded churches identified in this study using geolocated Twitter data. (P4: Cooper, 2018, p. 142).

| Church Name | Postcode |
|---|----------|
| Bow church | E3 3AH |
| Dartford Community Church | DA1 2HW |
| Hillsong Bermondsey | SE16 3LP |
| Hillsong Central London, Theatre Royal, Drury Lane | WC2B 5JF |
| King's church Walton | KT12 3JB |
| Macedonian Orthodox Church | W1D 4NQ |
| Rivergate Church | IG11 0FJ |
| St Pancras Old Church | NW1 1UL |
| The Redeemed Christian Church of God, Victory House | SE17 1TJ |

Figure 1 shows the 42 churches identified within this study (including the nine previously excluded churches) plotted onto a map, to show the geographical spread of the churches across London. The churches were broadly spread across the London boroughs, with a concentration within the central zone one area, as might be expected.

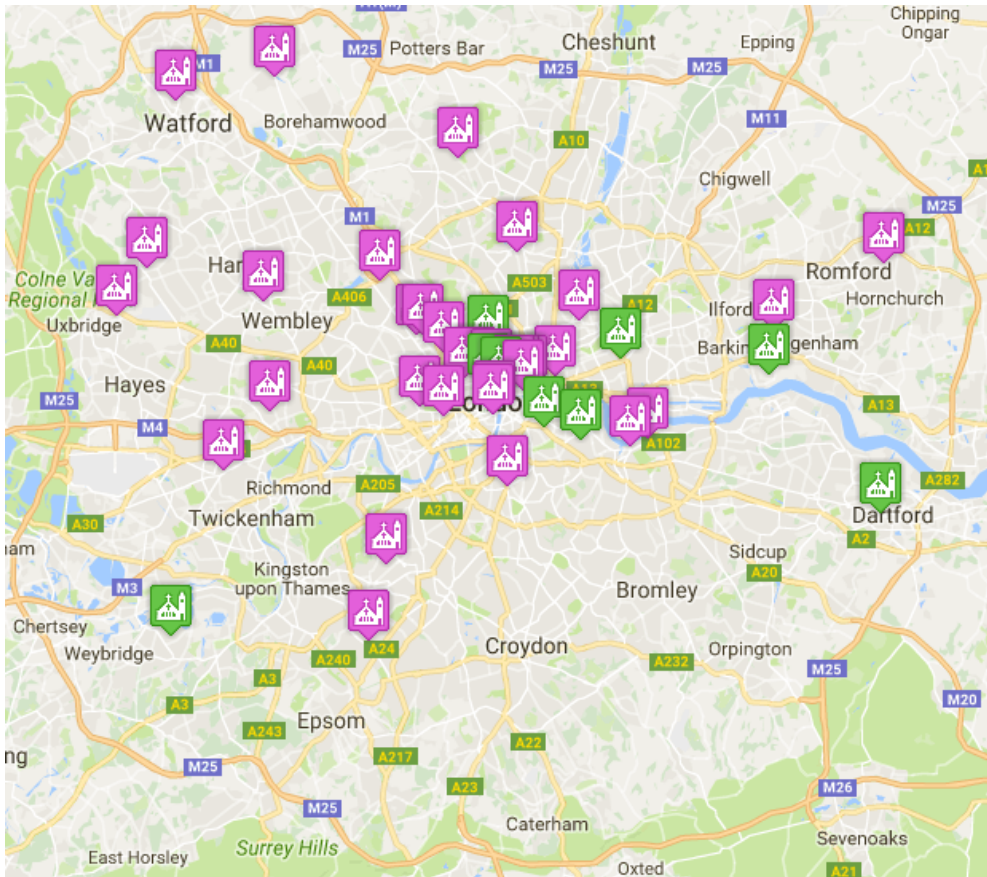


Figure 1: The 42 churches identified within this study; the previously excluded churches are shown in green, while the remainder of the churches are shown in pink.

It is worth noting that, owing to the time difference between the London Church Census in 2012 and the data collection for this study in 2014, it is possible that some of the nine churches identified as excluded from the London Church Census were actually new churches which did not exist at the time of the London Church Census. To mitigate this risk as far as practicable, web-based searching was conducted to understand the context and background of the churches identified and, in some cases, church leaders were contacted to request further information. Notwithstanding these steps taken, it is possible that a subset of the nine churches were in fact started since the London Church Census.

This research has clearly demonstrated that it is possible to use freely-available Twitter data to improve the sampling base of church attendance studies, to improve the accuracy and quality of church attendance data going forward. That this was achievable using only a sample of the available Twitter data indicates that it is likely

that this method will yield even greater results if able to make use of bigger and more complete Twitter datasets. This study concludes, therefore, by noting that the findings likely demonstrate a degree of undercounting in recent church attendance studies and by recommending that Twitter data is used in future church attendance studies with a view to improving data quality and completeness.

4.5 P5: Faith Communities Online: Christian Churches' Reactions to the COVID-19 Outbreak

This final study represented a slight shift in focus towards the end of the doctoral research process, in response to the COVID-19 pandemic which took the world by surprise and, in the first half of 2020, forced countries around the globe to impose restrictions to help reduce the spread of the virus. Recognising that the rapid uptake of online technology by churches, in response to restrictions on in-person services, would change the way in which people engage with and interact with churches and, thus, require new thinking by scholars focused on church attendance, the author designed a study to investigate the response to the pandemic taken by a sample of churches in three different countries.

To explore the church response to the COVID-19 pandemic, the author designed a qualitative survey which was issued to a convenience sample of churches in the UK, Finland and Namibia. Convenience sampling was utilised within this study to enable rapid data collection and analysis, to permit the research team to provide timely analysis and insight which might be of tangible benefit to the research community while the pandemic continued to affect the global church.

P5 provides a broad overview of the survey findings, broken down by country. Broadly, the study found that many of the churches surveyed had been quick to respond to the pandemic, rapidly creating and increasing their online service offerings in response to uncertain and changing events. The churches reported using a wide range of different platforms to facilitate their online service offerings, including: Facebook, Zoom, Google Meet, YouTube, Instagram, WhatsApp, Discord, Twitter, Switcher Studio, WordPress, Skype, Microsoft Teams, Mailchimp and email. The variety of platforms embraced demonstrates a willingness of the churches surveyed to experiment with a range of technologies and to operate in the wide range of online spaces utilised by their attendees and members.

The survey also found that the churches surveyed offered a wide range of services and activities online, in addition to the typical Sunday services, very soon after the COVID-19 restrictions were implemented. These services and activities included: Sunday services and masses, holy communion, daily readings and *Lectio Divina*, Holy Week services, Bible studies, video stations of the cross (during Lent),

small groups, children and youth activities, church and prayer meetings, quizzes, afternoon coffee meetings, choir practices and confirmation teaching. It was clear from the findings of the surveys that many of the church leaders at the surveyed churches had made a focused effort on ensuring that their churches went beyond simply offering religious activities and events; this can be evidenced through the offering of activities such as quizzes and coffee mornings. This demonstrates that, soon after the COVID-19 restrictions were implemented, church leaders were keen to utilise online technology to mirror as many of their offline activities as practicable. It was also noteworthy that several of the activities moved online related to Easter and Lent, which took place soon after the COVID-19 restrictions came into force. It is clear, therefore, that church leaders made a particular effort to ensure that their transition to online activities was quick and seamless and that key services of religious significance would not be missed. This demonstrates a degree of resilience across the churches surveyed, which were able to provide a high-level of support to their attendees at what was, otherwise, a challenging time.

When asked about their future intentions, with regards to online service and activity offerings, none of the responding churches indicated that they would fully cease their online activity; similarly, none of the responding churches indicated that they would fully maintain their online activity. Ten of the online churches indicated they were undecided around next steps with regard to their online activity. One of the survey respondents, Dr. Bonnington of King's Church Durham, remarked when discussing the maintenance of online service and activity offering:

“life won't return to 'normal' ante-CV19 especially for older/vulnerable people so we will continue with 'dual' worship for a long time.” (P5: Cooper, Jormanainen, Shipepe & Sutinen, 2021, p 113).

It was unknown, at the point at which the survey was conducted, quite how long the COVID-19 pandemic would persist for, with many people hoping for a swift return to normality. Fast-forwarding to Christmas 2020, the pandemic had persisted for far longer than many had anticipated. This comment, by Dr. Bonnington, therefore proved to be very insightful at an early stage in the pandemic – it now seems clear, at the time of writing this dissertation, that normality is still some way off and even when COVID-19 restrictions are eased, it is likely that a degree of flexibility will be requiring, meaning that many churches may well need to continue maintaining their online offering well into the future.

Through analysing the responses of surveyed churches in the UK, Finland and Namibia, this paper addressed **RQ9**, by providing timely insight into the ways in which churches around the world responded to the COVID-19 pandemic.

Having explored the approaches taken by churches to increase their online offering, this paper gave consideration to the impacts of the COVID-19 pandemic on church attendance and on studies exploring church attendance going forward. Although this qualitative study was not configured to provide robust church attendance data, survey respondents were asked to comment on how the switch to online impacted on church attendance. Many of the respondents noted a positive increase in attendees at online services, with some of the Finnish churches reporting a ten-fold increase in attendance at Easter Sunday services during 2020, compared to attendance at Easter Sunday services in 2019. Several of Durham's church leaders reported a similar story; Emmanuel Church, for example, reported 500 views for their first Sunday service (the church has a usual Sunday attendance of 300) and estimated that around 480 people had watched their 2020 Easter Sunday services, compared to around 220 people who attended the 2019 Easter Sunday services. However, this observed pattern of growth was not ubiquitous; St Cuthbert's Catholic Church estimated that their live stream videos had achieved 50 viewers, compared to a usual Sunday attendance of around 320 people. This demonstrates that while the COVID-19 pandemic might be driving church growth in some churches, that growth is not stable and consistent across the board. There would clearly be further benefit in future research exploring the drivers of church growth and church decline during the pandemic, to better understand the responses of church attendees. It would be interesting, for example, to consider the extent to which people remained with their usual churches during periods of remote church attendance versus the extent to which people switched churches (perhaps to attended services at familial churches or churches they have previously attended), given the removal of geographical barriers.

Carefully focused on **RQ10** and the impact of COVID-19 on church attendance and the design of future church attendance research, this paper also highlighted some of the challenges which researchers will need to be cognisant of when designing church attendance studies going forward. The platforms used for online church services will present challenges for some studies – one church leader from Durham, for example, reported that it was possible to view the number of rooms being used in Zoom, but not the number of people in each room. In addition to platform issues, there are a number of fundamental methodological questions which will need to be addressed during the study design stage – for example, as devices join and leave online services, how can church attendance studies maintain an accurate picture of the number of unique devices? Similarly, what inferences should church attendance studies make with regard to the number of people likely to be watching a particular service behind a single shared device? The answers to these questions will be essential to ensure that the findings of such future studies are as robust as possible and are as insightful as possible, particularly when it comes to comparisons between

data including a high percentage of online attendees compared to older data which covered mainly in-person attendees. There is clear potential for future studies of church-related tweets to provide support to some of these new challenges. Analysts could consider, for example, the ways in which tweet content related to presence at online church services or even the act of joining/leaving online church services might be used to better understand participant behaviour when using devices to access online church service content. This could provide valuable insight to support the interpretation of existing metrics such as the number of devices connected to online services.

Through exploring the impact of the COVID-19 pandemic on churches surveyed and through addressing some of the key considerations which designers of future church attendance studies will need to explore, this study has addressed **RQ10**, providing historical information on the situation in 2020 and actionable insight into the work required to support future research work in this field. This study has also complemented the findings of other research studies which were quickly conducted following the introduction of measures to restrict the spread of COVID-19. In reporting on the results on a survey of church leaders based on the Island of Ireland, Ganiel (2020) similarly noted a fast increase in the use of technology by churches to provide online worship opportunities. The author also noted that “70 percent of respondents agreed that they would retain aspects of their online ministries when restrictions on public gatherings are lifted” (p. 5); a finding which resonates with the insight presented in P5 that none of the responding churches indicated that they would fully cease their online activity. In a separate study (Village & Francis, 2020), survey research found that the invitation rate for activities which congregants could participate in during acts of online worship was “relatively low for things such as prayer or reciting the liturgy, though most who were invited to do so did join in” (p. 58). This finding could be of relevance when considered alongside the insight presented in P5 regarding the platforms which survey respondents selected when providing online access to services and activities; going forward, church leaders might be advised to take into account to extent to which platforms permit increased congregant participation when selecting which one to use. Future research on this topic might explore the extent to which platform selection and opportunities for congregant participation impact upon online church attendance and engagement.

5 Discussion

Having provided a description of the problem-space in which this doctoral study was conducted, this dissertation will proceed to provide a summary of the findings alongside an analysis of why the results matter, an explanation of the key limitations of the study and, finally a series of recommendations for future work in the field.

5.1 Summary of Answers to Research Questions

The research questions addressed within this doctoral dissertation are listed in Table 11, alongside a summary of the answers obtained during the doctoral research studies conducted.

Table 11: The ten research questions addressed within this doctoral dissertation with accompanying answers obtained during doctoral research.

| Research Question | Associated Paper | Answer |
|---|------------------|--|
| RQ1: Is the existing qualitative coding framework for church-related tweets fit-for-purpose? | P1 | Yes, P1 was able to test the existing qualitative coding framework using a bigger dataset and confirmed that it was fit-for-purpose, subject to the addition of one new coding label – Discussion about Hillsong. |
| RQ2: What topics are discussed by social media users in church-related tweets? | P1 | P1 was able to provide insight on the topics discussed in church-related tweets. This insight is presented in Table 2. |
| RQ3: What can sociologists of religion learn from the content of church-related tweets? | P1 | P1 was able to identify and summarise a number of key insights which are presented in this dissertation. These key insights included the potential for future studies to use the framework presented to explore changes in the content of church-related tweets over time or across different geographical areas. |
| RQ4: Can content analysis of church-related tweets be automated using machine learning? | P2 | Yes, P2 demonstrated that it was possible to automate the qualitative coding of church-related tweets using the Naïve Bayes classifier. |

| Research Question | Associated Paper | Answer |
|--|------------------|--|
| RQ5: What are the implications of using machine learning to classify church-related tweets? | P2 | P2 was able to identify and summarise a number of key insights which are presented in this dissertation. These key insights included the potential for automated qualitative coding of church-related tweets to lower the barrier to entry for studies similar in design to P1 , which creates opportunities to more easily track changes in the content of church-related tweets over time. Another key insight was the potential benefit of clustering to reduce reliance on human-coded datasets in future similar studies. |
| RQ6: Is there a relationship between the net sentiment score of church-related tweets and the presence of church growth in the borough from which the tweets were posted? | P3 | Yes, P3 demonstrated, using an independent-samples t-test, that there is a statistically significant relationship between the sentiment of church-related tweets and the presence of church growth in the borough from which the tweets were posted. |
| RQ7: Can geolocated Twitter data be used to identify church locations? | P4 | Yes, P4 demonstrated that it is possible to identify church locations; in total 42 London churches were identified within the analysis conducted. |
| RQ8: Can geolocated Twitter data can be used to uncover previously hidden churches? | P4 | Yes, P4 demonstrated that it is possible to identify previously hidden church locations; in total nine of the 42 London churches identified within the analysis were excluded from the 2012 London Church Census. |
| RQ9: How have churches around the world responded to the COVID-19 pandemic? | P5 | P5 provided a detailed description of the ways in which churches surveyed in the UK, Finland and Namibia responded to the COVID-19 pandemic. Key insights from these findings are summarised in this dissertation. |
| RQ10: How has the COVID-19 pandemic affected church attendance at the surveyed churches and what considerations might be required to better monitor this going forward? | P5 | P5 provided an indication of the impact of COVID-19 on church attendance based on self-report data from surveyed churches in the UK, Finland and Namibia. P5 also highlighted some of the key considerations which the recent shift to online church will create for future studies of church growth, including the requirement for new and clearly defined metrics to capture and understand online church attendance. |

5.2 Key Insights and Relevance of Findings

This doctoral study has driven forward knowledge in the fields of computer science, digital theology and the sociology of religion. Set in the context of identified data quality issues affecting studies of church attendance, this study has demonstrated that freely available Twitter data can be used to explore the context of church attendance. It has also identified a relationship between the net sentiment of church-related tweets and the presence of church growth in the locations from where the tweets were posted, indicating that it might, in future, be possible to use church-related Twitter data to supplement missing or low-quality church attendance data. This thesis has also demonstrated that Twitter data can be used to identify churches to be included in church attendance studies going forward. It has also provided timely analysis and insight into the impact of the ongoing COVID-19 pandemic on church attendance, providing recommendations for future considerations which will need to be taken by scholars seeking to research church attendance in the context of the dramatic recent shift toward church online.

The coding framework presented in **RQ1** means that, going forward, digital theologians and sociologists of religion have a meaningful starting point from which to commence analysis of the content of church-related tweets. The fact that the study, using a sample of 1,004 tweets from across all London Boroughs, was able to validate a coding framework quickly pulled together during a previous preliminary study using a very small sample of only 108 tweets from one London Borough, demonstrates that exploratory studies limited to small sample sizes can still be useful in setting direction for subsequent more robust studies using bigger sample sizes.

The process of answering **RQ2** provided useful context to this doctoral study, by presenting an analysis of the types of topics which Twitter users discuss in church-related tweets. This set the scene for the rest of this doctoral study, enabling the author to better understand how Twitter is used in church-related discourse, how tweets are structured and how they can be applied to wider analysis. The findings of that study are also useful in that they help digital theologians and sociologists of religion to better understand how Twitter is used as a platform and how churches are discussed online. In answering **RQ3** the study noted a number of interesting possibilities that the emergence of a robustly validated framework opens up – including the potential to conduct repeat studies in future to assess changes in the content of church-related tweets compared to this baseline set using the thesis dataset from 2014. The study did note, however, that the approach taken to analysing this larger sample of tweets was very laborious in nature, meaning that future studies could be limited to size and scope, unless a more time-efficient approach was identified.

RQ4 addressed this observed challenge, by testing whether the qualitative-coding of church-related tweets, as conducted in **P1**, could be automated using

supervised machine learning. The study tested three machine learning approaches – Support-Vector Machine, J48 Decision Tree and Naïve Bayes. The study concluded that the results for the Naïve Bayes classifier had demonstrated that it was possible to automate the qualitative coding of church-related tweets. This finding paves the way for future studies into the content of church-related tweets to take place at greater scale and with greater speed and efficiency. **RQ5** considered the implications of this finding, noting that the automation of the qualitative coding of church-related tweets could be particularly valuable to digital theologians and sociologists of religion, by enabling them to quickly conduct studies exploring how the content of church-related tweets changes over time or across locations. This would permit straightforward identification of changes which might impact on other studies making use of church-related tweets. In addition to enabling studies to better understand the content of church-related tweets, the findings of this study create opportunities for scholars conducting future analysis of church attendance with a view to quantifying church growth or decline. Assuming some model could be devised to conduct predictive analysis using datasets comprising enriched (e.g. with content analysis labels and sentiment analysis scores) church-related tweets, the ability to conduct that enrichment for content analysis labels very quickly could pave the way for the rapid development of metrics to complement missing or low quality church attendance data.

RQ6 complemented these findings by demonstrating that there was a relationship between the sentiment of church-related tweets and the presence of church growth in the borough from which the tweet was posted. This finding is noteworthy, as it demonstrates the potential for Twitter data to be used to develop metrics to complement low-quality or missing church attendance data in future studies. This could be of particular value to studies going forward, in the context of the recent rapid shift to church online, which has the potential to make future church attendance studies very challenging and, therefore, the potential to undermine the robustness and validity of resultant church attendance data. It was beyond the scope of this study to explore whether it is possible to identify drivers of church growth based on the content or sentiment of church-related tweets. There would be clear benefit in future studies addressing this question, to support the work of clergy in promoting, identifying and measuring church growth across a range of different church styles and activities. That this study was able to empirically demonstrate a relationship between online activity and offline activity is likely to be of interest to academics and researchers in a variety of fields, not just limited to studies set in a religious context. This finding raises interesting questions around the extent to which Twitter data might, in future, be able to be applied more widely to generate proxies for measures of offline activity. Future research in this regard clearly has the

potential to open up exciting new opportunities for computer science researchers operating in a wide range of problem spaces.

Having explored a range of uses related to the content of church-related tweets, **RQ7** built on the existing findings of this doctoral study by exploring the use of tweet metadata, in this case the WGS84 geolocation data which is sometimes attached to tweet data. Analysis identified that it was possible to use church-related tweets to identify the physical locations of church buildings. More than that, **RQ8** demonstrated that it was possible to identify the physical location of church buildings which had been excluded from previous more traditional studies of church attendance. This is an important research finding as it tangibly demonstrates the potential for Twitter data to be used to complement the approaches utilised in traditional church attendance studies to improve the quality of those studies and the robustness of their findings. In addition to identifying the potential for Twitter data to improve church attendance studies, this finding importantly demonstrated that it is likely that previous studies of church attendance in London (and, therefore, likely elsewhere) have experienced a degree of undercounting. This is crucial to understand in a context in which church growth is already being observed and documented in London's capital (see, for example: Goodhew & Cooper, 2018b). This finding flags the importance of future studies re-visiting this observed phenomenon of church growth with a view to improving the accuracy of research to quantify the true extent of church growth.

RQ9 built on the findings of the earlier research questions by exploring the impact of COVID-19 on a sample of churches in three different countries. It was found that those churches had responded quickly to the pandemic and the restrictions which had been implemented in an effort to curb the spread of the virus. In responding, those churches indicated that they had vastly increased their offering of online services and activities across a range of platforms, with a view to meeting church attendees on the platforms which they ordinarily use and are familiar with. In so doing, the churches had been rapid in their response; this meant that they were able to very quickly offer services during Holy Week and Easter Sunday, ensuring that the church was able to mark this very important season in the church calendar in a variety of ways, despite having to take a novel approach in doing so. The churches sampled also demonstrated a variety of online offerings outside of the range of traditional church services and activities, including through online quizzes. This demonstrated a desire by the church to continue developing community and relationship, despite the physical barriers to doing this in the local community. While the study was not configured to collect robust church attendance data, to address **RQ10** it briefly explored the early experiences of the churches sampled with regard to church attendance. It found that several of the churches had seen attendance increase since shifting to online services, some up to ten-fold, although this increase

was not ubiquitous or consistent. Some churches had seen a drop-off in attendance since shifting to online modes of operation. Importantly, this study noted that, going forward, sociologists of religion would have to give careful consideration to study design and approaches used to measure church attendance. This is because of the unique challenges posed by measuring online engagement and attendance – for example, the lack of clarity around how many people a single screen view might represent or around how many unique devices a screen view count might represent in a context where some users might drop-off services and re-connect later, for a multitude of reasons. Clearly, the pandemic has shifted the way in which churches operate and behave and will, accordingly, need to shift the way in which researchers go about understanding, measuring and documenting church attendance. Some of the findings of this doctoral research should, however, support that effort by demonstrating the ways in which Twitter data might be able to complement traditional study data to better understand church attendance and, therefore, church growth and decline.

Considered holistically, these studies have progressed the evidence base around the use of Twitter data in studies of church attendance. They have demonstrated that notable online church-related discourse exists on the platform. They have demonstrated that it is possible to classify that content to better understand the types of discourse which are taking place and that it is possible to automate this process so that it can be repeated quickly and at scale. They have demonstrated that there is a relationship between the net sentiment of church-related tweets and the presence of church growth in the location from which the tweets were posted; therefore, paving the way for studies to develop predictive models of the presence (or absence) of church growth. They have demonstrated that it is possible to improve traditional church attendance studies using Twitter datasets to improve the coverage of those studies, thus improving data quality and robustness. They have demonstrated that, going forward, church attendance studies will need to adapt in response to the recent shift to church online, but that consideration of Twitter data might be able help this adaptation. Overall, while there remains much work to be done to enable digital theologians and sociologists of religion to be able to respond to the key challenges raised by proponents of the secularisation thesis, this doctoral research has demonstrated that Twitter data can be a powerful tool in that response.

5.3 Key Limitations

There are a number of key limitations associated with this doctoral study, which it is important to highlight in this dissertation. The first such limitation is the use of a sample of data provided by the Twitter Application Programming Interface. The application programming interface used does not provide all tweets matching search

requests, but rather a quasi-random time-slice of those tweets¹⁸. The lack of transparency in the approach taken by Twitter in generating that time-slice means that it is not evident how truly random the resultant sample of tweets actually is, nor what proportion of the population of tweets matching the search request any resultant sample represents. This limitation means that the findings of **P1** and **P3** should be interpreted with a degree of caution; future research using population data would be of value to demonstrate that the findings still hold. The findings of **P2** and **P4** are less affected by this limitation. **P2** was focused on exploring whether machine learning could be used to automate qualitative human coding; there is no reason to believe this finding might not hold if the dataset had been comprised differently. **P4** was focused on exploring the use of WGS84 data attached to Twitter data; there is no reason to believe the findings would not hold if the dataset comprised different tweets. **P5** is entirely unaffected by this limitation as it did not make use of the thesis dataset.

The thesis dataset was also limited in that it comprised only tweets posted in the English language, used only the word ‘church’ as a search term (the additional use of terms such as ‘cathedral’ and ‘chapel’ might have added to the composition of the thesis dataset) and gathered only tweets posted on a Sunday (some churches meet on other days of the week). While this limitation will have impacted on **P1** to **P4**, that impact is considered to be relatively minor. The aim of the research presented within this dissertation was to explore the art of the possible using church-related Twitter data – to explore whether such data could be meaningfully qualitatively analysed, whether that analysis could be automated, whether a relationship could be detected between that data and offline activity and whether that data could be used to detect physical church locations. That all ten research questions set out in this dissertation to address those research challenges were able to be addressed using a dataset limited by language, scope and day of the week demonstrates that this limitation did not impact on the intent of the study. More than that, it indicates that there is value in future studies being conducted using a broader dataset, to begin to explore what more we might be able to learn from church-related tweets.

The aim of this study was always to explore the use of church-related Twitter data, which the publications presented in this dissertation have demonstrated was achieved. This exclusive focus on one particular social media platform could,

¹⁸ It is worth noting that since the collection of the thesis dataset described in this doctoral dissertation, Twitter have launched a new application programming interface product to make the full tweet catalogue freely available to researchers. Therefore, future studies in this field will benefit from greater access to more complete datasets. See: <https://developer.twitter.com/en/solutions/academic-research/products-for-researchers> (last accessed: 10 April 2021).

however, be considered a limitation. Expanding this study to explore a wider range of social media platforms (e.g. Facebook, Instagram, YouTube) might have broadened the results and led to greater insight into the online discourse related to church taking place on social media platforms and the relevance and implications of that discourse to studies of church attendance. This limitation does not, however, undermine the findings of this doctoral study; it merely affirms the benefit of future studies exploring a wider range of social media platforms to test the findings of this work and to identify additional insights, potentially exploring the extent to which data obtained from particular platforms may have benefits over data from competing platforms.

The research described in **P5** made use of convenience sampling, to enable the research team to generate actionable insights as quickly as possible while the global COVID-19 pandemic was still ongoing. There would, therefore, be clear benefit in future studies exploring this topic again, using larger and more representative sampling to test the findings of this research and to identify any changes in the response of churches across the globe to the COVID-19 pandemic as the pandemic has unfolded and as churches have had a longer period of time to adjust to online ways of working.

5.4 Opportunities and New Horizons for Future Research

In light of the key limitations of this study, as outlined in the previous section of this dissertation, there would be clear benefit in future studies repeating the application of the methods outlined in this dissertation, making use of extended and enhanced datasets. This could include datasets comprising population data, rather than sample data, datasets making use of a greater variety of search terms, datasets collating posts in multiple languages and posts made to a wider selection of social media platforms. Such studies should seek to validate the findings of this dissertation research across more diverse datasets while also seeking to identify what, if any, additional insights might be achievable using these alternatively comprised collections of posts.

P1 explored the content of church-related tweets, while seeking to assess whether a coding framework generated during a previous study with a smaller sample size could be considered fit-for-purpose when applied to a bigger dataset. While there is clearly value in future studies exploring the content of church-related tweets again, to assess whether there has been any change in topics discussed, there would also be benefit in applying this same research approach to different research questions within the field of digital theology; for example, there would be merit in exploring the content of tweets discussing theological issues before/after major changes in the doctrinal positions of different denominations/churches or in exploring the content of tweets

expressing prayer content. Such studies will help improve the evidence base on the role of Twitter in the daily lives of faith communities.

P2 demonstrated that it was possible to automate the qualitative coding of church-related tweets using supervised machine learning algorithms. This research was dependent, however, on the availability of data coded using an extant framework for the online discourse contained within church-related tweets. Future studies might want to qualitatively code tweets where such a framework does not already exist. There would, therefore, be benefit in building on the findings of **P2**, by exploring the potential benefits which unsupervised machine learning might bring to future studies of church-related tweets. The application of unsupervised learning approaches might, for instance, enable the automation of the process of identifying the clusters which church-related tweets naturally fall into and, therefore, the different topics which church-related tweets can be divided into. Combining unsupervised learning with the supervised learning approach described in **P2** might, therefore, enable future studies to automate the qualitative coding of church-related tweets, even where robust coding frameworks do not exist, thus reducing the reliance on the outputs of laborious human coding work and pre-determined coding frameworks. This would improve both the ease and speed of the qualitative analysis of Twitter data going forward.

P3 demonstrated that there was a relationship between the sentiment of church-related tweets and the presence of church growth in the location from which the tweets were posted. In light of this finding, there would be benefit in future studies exploring whether a predictive model could be derived to predict the presence (or absence) of church growth based on the content and sentiment of church-related tweets. This could be of particular value in future studies where church attendance data is either missing or of low quality. Being able to complement church attendance data with predictive analysis might improve the evidence base available to researchers on likely patterns of church attendance which could, in turn, direct the targeting of future research efforts to continue to progress the evidence base on church attendance in the UK.

P5 began to explore some of the emerging impacts of the COVID-19 pandemic on church attendance around the world, including the rapid shift towards much greater up-take of doing church online. The paper highlighted a number of considerations which scholars of digital theology and the sociology of religion would need to take when designing future church attendance studies in light of increased online church attendance. There would, therefore, be clear tangible benefit in future research to explore how future church attendance studies can best account for online church attendance and how such studies can be designed to make resultant datasets as robust as possible and as comparable to historic analyses as possible. This will strengthen the ability of analysts to draw conclusions on trends and changes over time, despite the major changes in approaches to church attendance.

5.5 Practical Recommendations for Church Leaders

In defining digital theology from a computer science perspective, this doctoral dissertation echoed the call of **P10** that digital theology is a multi-disciplinary field of academic enquiry, which is deeply rooted in practice. It seems fitting, therefore, to conclude this discussion chapter with a series of recommendations for church leaders, formed in light of the findings presented within this doctoral dissertation:

1. **Church leaders should study and explore social media discourse and content related to their churches:** The finding that there is a relationship between the sentiment of church-related tweets and the presence of church growth indicates that there might be beneficial learning which church leaders can draw from tweets related to their church. Where churches are seeing attendance decline or stagnation, there might be potential for those church leaders to understand more about their congregations (or potential congregants) and get a sense of some of the attitudes potentially driving that observation, thus permitting the opportunity to act in response.
2. **Church leaders should have open dialogue with their attendees around their viewing/participation habits when attending church online:** Understanding how attendees do church online, the hardware they use, the challenges they face and other nuances related to the lived experiences of attendees offers the potential for church leaders to drive experience improvements. It also offers the potential for church leaders to give early thought to ways in which they might most accurately measure online church attendance; this might well continue to be of increased importance even as the recent limitations imposed in response to the recent COVID-19 pandemic are withdrawn.
3. **Church leaders should consider the problems and challenges in the digital space which they desire solutions to and seek opportunities to collaborate with digital theologians on research projects:** A recurring observation throughout the studies presented in this doctoral dissertation was the importance of multi-disciplinary teams working together to solve and address real problems and challenges faced by practitioners. Digital theology offers exciting opportunities for co-design, which permit the active involvement of non-technical experts; church leaders should relish such opportunities to drive forward the research agenda within this emerging field of academic enquiry and ensure that outputs are as relevant and practical as possible.

6 Conclusion

This doctoral dissertation has summarised the findings of five research publications produced over a period of almost seven years. The main aim of those publications was to explore the use of social media data, in this case specifically Twitter data, to better understand church attendance and to drive improvements in the quality and robustness of datasets resulting from church attendance studies. Robust studies exploring church attendance are of real importance in a society where continuing secularisation is assumed as inevitable and where the decrease in the relevance of the church is considered to be an unavoidable given.

The findings of the publications which form this thesis have demonstrated that there is tangible benefit in using Twitter data when studying church attendance. Such data can provide valuable insight into the types of church-related topics which are posted online. Automation of analysis using machine learning can make future repeat studies of the content of church-related tweets much quicker, cheaper and easier to conduct, opening up the option of regular repeat studies to monitor changes in online church-related discourse over time and across locations. The publications comprising this thesis have also demonstrated that there is a relationship between the content of those church-related tweets and the presence of church growth in the locations where the tweets were posted; this poses an interesting opportunity to explore the development of predictive models to predict church growth in locations where church attendance data is sparse or of low quality. The research described in this dissertation has also proven that church-related tweets can also be used to identify physical church locations which should be included in traditional church attendance studies to improve the reach and completeness of those studies. Finally, the research presented has explored some of the impacts of the COVID-19 pandemic on churches in three different countries and this dissertation has considered the implications of those impacts on future studies designed to explore church attendance.

Considered holistically, the research findings presented in this doctoral dissertation have driven forward knowledge and understanding of the potential uses of Twitter data to better understand church attendance and to address some of the existing data quality challenges affecting recent church attendance studies. They

have also contributed to the evidence base on the content of church-related online discourse. In making these original contributions to knowledge, these findings have demonstrated to scholars of computer science, sociology of religion and digital theology the benefits of truly multi-disciplinary digital theology research to solve real-world problems and generate actionable insights.

Acknowledgements

I gratefully extend my sincere thanks to my supervisor, Erkki Sutinen, who has provided me with diligent support, direction and guidance throughout my time working on this doctoral thesis, as well as being a good friend and a source of much adventure and fun. I also extend my thanks to my second supervisor, Peter Phillips, who has shaped my learning journey within the field of digital theology. I also thank David Goodhew, who although not a formal supervisor has been a mentor and great support to me throughout my time working on this thesis.

Throughout my time working on this doctoral thesis, I have been privileged to be able to research and write with a diverse group of talented and enthusiastic co-authors – Emmanuel Awuni Kolog, Joshua Mann, Ilkka Jormanainen, Anastasia Shipepe, Samuli Laato, Suvi Nenonen, Nicolas Pope, David Tjiharuka and Jonas Kurlberg. My thanks to all of you for the ideas we have shared and the inspiration you have provided over these past years.

I am extremely grateful to Siiri Toiviainen Rø for her diligent and careful work translating the abstract of this doctoral thesis into Finnish.

I acknowledge and thank the staff at St John's College and Cranmer Hall, Durham University, who oversee the governance and work of my research centre – *The Centre for Church Growth Research*. In particular I thank David Wilkinson, the Principal of St John's College and Philip Plyming, the Warden of Cranmer Hall, for their support.

Particular thanks are owed to my girlfriend, Beth Mortimer, for her enthusiasm and support during these final months of the PhD process and for going with me on the journey of the ups and downs of getting to this final point of completion. Thanks for the encouragement, advice and exciting adventures along the way!

Finally, I thank my family and friends who have been a source of support and reassurance throughout my doctoral research. In particular, I gratefully thank my parents, Diane and Wayne Cooper and my grandparents, Norma and Jack Holland, for their continual enthusiasm for my work and research and for the advice and encouragement they have offered along the way.

September 2021¹⁹
Anthony-Paul Cooper



ANTHONY-PAUL COOPER

Anthony-Paul is Co-Director for the Centre for Church Growth Research. He has a background in computer science and social research, with recently published research topics including new church use of ‘secular’ and ‘sacred’ space and the use of social media data to better understand church attendance and church growth.

¹⁹ The acknowledgements were originally written in May 2021, prior to pre-examination. Additional names have been added in September 2021 prior to publication.

List of References

- Ahmad, S., Asghar, M. Z., Alotaibi, F. M. & Awan, I. (2019). "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques". *Human-centric Computing and Information Sciences*, 9(24).
- Bickley, P. & Mladin, N. (2020). *Religious London: Faith in a global city*. London, UK: Theos.
- Brierley, P. [Ed.] (2011). *UK Church Statistics, 2005–15*. Tonbridge, UK: ADBC Publishers.
- Brierley, P. (2013). *Capital Growth*. Tonbridge, UK: ADBC Publishers.
- Brierley, P. (2020). *UK Church Statistics No 4: 2021 Edition*. Tonbridge, UK: ADBC Publishers.
- Brown, C. (2001). *The Death of Christian Britain*. Abingdon, UK: Routledge.
- Bruce, S. (2002). *God is Dead: Secularization in the West*. Oxford, UK: Wiley-Blackwell.
- Bruce, S. (2013a). "Secularization and church growth in the United Kingdom". *Journal of Religion in Europe*, 6(3), p. 273–296.
- Bruce, S. (2013b). "Further thoughts on church growth and secularization". *Journal of Religion in Europe*, 6(3), p. 316–320.
- Campbell, H. & Vitullo, A. (2019). "Popes in Digital Era: Reflecting on the Rise of the Digital Papacy". *Problemi Dell'Informazione*, 44(3), p. 419–442.
- Chen, N-C., Kocielnik, R., Drouhard, M., Peña-Araya, V., Suh, J., Cen, K., Zheng, X. & Aragon, C. R. (2016). "Challenges of Applying Machine Learning to Qualitative Coding". Paper presented at CHI 2016 Workshop on Human Centred Machine Learning, San Jose, CA, USA, 7–12 May 2016.
- Cheng, Z., Caverlee, J. & Lee, K. (2010). "You are where you tweet: A content-based approach to geolocating Twitter users". In *proceedings of the 19th ACM International Conference on Information and Knowledge Management*, Toronto, Canada, October, 2010: 759–768.
- Cheong, P. H. (2010). "Faith Tweets: Ambient Religious Communication and Microblogging Rituals". *M/C Journal*, 13(2).
- Cheong, P. H. (2016). "Religious authority and social media branding in a culture of religious celebrification". In *The Media and Religious Authority*, edited by S. Hoover, p.81–104. Pennsylvania, USA: Penn State University Press.
- Codone, S. (2014). "Megachurch pastor Twitter activity: An analysis of Rick Warren and Andy Stanley, two of America's social pastors". *Journal of Religion, Media & Digital Culture*, 3(2), p. 1–32.
- Cooper, A. P. (2014). "Unwrapping Camden's church tweeters: A small-scale thematic study of Twitter data". In *Social Media in Social Research: Blogs on Blurring the Boundaries*, edited by K. Woodfield. London, UK: NatCen Social Research.
- Cooper, A. P. (2017). "Assessing the possible relationship between the sentiment of church-related tweets and church growth". *Studies in Religion/Sciences Religieuses*, 46(1), p. 37–49.
- Cooper, A. P. (2018). "Using geotagged Twitter data to uncover hidden church populations". In *The Desecularisation of the City: London's Churches 1980 to the Present*, edited by D. Goodhew and A. P. Cooper. Abingdon, UK: Routledge, p. 134–147.
- Cooper, A. P. & Goodhew, D. (2017). "'Resacralising' secular space: new churches in a northern city, 1980–2012". *Journal of Contemporary Religion*, 32(3), p. 495–511.

- Cooper, A.P., Jormanainen, I., Shipepe, A. & Sutinen, E. (2021). "Faith communities online: Christian churches' reactions to the COVID-19 outbreak". *International Journal of Web Based Communities*, 17(2), p. 99–119.
- Cooper, A. P., Kolog, E. A. & Sutinen, E. (2019). "Exploring the use of machine learning to automate the qualitative coding of church-related tweets". *Fieldwork in Religion*, 14(2), p. 140–159.
- Cooper, A. P., Laato, S., Nenonen, S., Pope, N., Tjiharuka, D. & Sutinen, E. (2021). "The reconfiguration of social, digital and physical presence: From online church to church online". *HTS Theologiese Studies/Theological Studies*, 77(3).
- Cooper, A. P., Mann, J., Sutinen, E. & Phillips, P. (2021). "Understanding London's church tweeters: A content analysis of church-related tweets posted from a global city". *First Monday*, 26(9).
- Crowston, K., Liu, X., Allen, E. E. & Heckman, R. (2010) "Machine Learning and Rule-based Automated Coding of Qualitative Data". Paper presented at ASIST 2010, Pittsburgh, USA, 22–27 October 2010.
- Dann, S. (2010). "Twitter Content Classification". *First Monday*, 15(2).
- Davie, G. (1994). *Religion in Britain since 1945: Believing without Belonging*. Oxford, UK: Blackwell.
- Dharmadhikari, S. C., Ingle, M. & Kulkarni, P. (2011). "Empirical studies on machine learning based text classification algorithms". *Advanced Computing: An International Journal*, 2(6), p. 161–169.
- Frias-Martinez, V. & Frias-Martinez, E. (2014). "Spectral clustering for sensing urban land use using Twitter activity". *Engineering Applications of Artificial Intelligence*, 35, p. 237–245.
- Ganiel, G. (2020). "People Still Need Us: A Report on a Survey of Faith Leaders on the Island of Ireland during the Covid-19 Pandemic". Irish Council of Churches/Irish Inter-Church Meeting. Available at: <https://www.irishchurches.org/cmsfiles/resources/People-Still-Need-Us-May-2020.pdf> (last accessed: 11 April 2021).
- Geraci, R. (2010). *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*. Oxford, UK: Oxford University Press.
- Goodhew, D. [Ed.] (2012a). *Church Growth in Britain, 1980 to the Present*. Farnham, UK: Ashgate.
- Goodhew, D. (2012b). "Church Growth in Britain, 1980 to the Present Day." In *Church Growth in Britain, 1980 to the Present*, edited by D. Goodhew. Farnham, UK: Ashgate. p. 3–22.
- Goodhew, D. (2013). "Church growth in Britain: A response to Steve Bruce". *Journal of Religion in Europe*, 6(3), p. 297–315.
- Goodhew, D. & Cooper, A. P. [Eds.] (2018a). *The desecularisation of the city: London's churches 1980 to the present*. Abingdon, UK: Routledge.
- Goodhew, D. & Cooper, A. P. (2018b). "London's Churches, 1980 to the Present" In *The Desecularisation of the City: London's Churches, 1980 to the Present*, edited by David Goodhew and Anthony-Paul Cooper. Abingdon, UK: Routledge, p. 3–38.
- Holmberg, K., Bastubacka, J. & Thelwall, M. (2016). "@God please open your fridge! A content analysis of Twitter messages to @God: Hopes, humour, spirituality, and profanities". *Journal of Religion, Media & Digital Culture*, 5(2), 339–355.
- Jackson (2018). "The Diocese of London and the Anglican church in London, 1980 to the present". In *The Desecularisation of the City: London's Churches 1980 to the Present*, edited by D. Goodhew and A. P. Cooper, p. 134–47. Abingdon, UK: Routledge.
- Karjalainen, S. & Halonen, R. (2019). "Twitter at the Hands of the Church". 32nd BLED eConference Humanizing Technology for a Sustainable Society, June 16th to 19th 2019, Bled, Slovenia.
- Kaur, G. & Chhabra, A. (2014). "Improved J48 Classification Algorithm for the Prediction of Diabetes". *International Journal of Computer Applications*, 98(22), p. 13–17.
- Kelly, K. (1999). "Nerd theology". *Technology in Society*, 21, p. 387–392.
- Kolog, E. A., Sutinen, E. & Nygren, E. (2016). "Hackathon for Learning Digital Theology in Computer Science". *Modern Education and Computer Science*, 6(8), p. 1–12.
- Kurlberg, J. & Phillips, P. M. [Eds.] (2020). *Missio Dei in a Digital Age*. London, UK: SCM Press.
- Le, B. & Nguyen, H. (2015). "Twitter Sentiment Analysis Using Machine Learning Techniques". In *Advanced Computational Methods for Knowledge Engineering. Advances in Intelligent Systems*

- and Computing, vol 358, edited by Hoai An Le Thi, Hoai Minh Le, Tao Pham Dinh and Ngoc Thanh Nguyen. Springer, Switzerland, p. 279–289.
- Lewis, B. (2018). “Social Media, Peer Surveillance, Spiritual Formation, and Mission: Practising Christian Faith in a Surveilled Public Space”. *Surveillance & Society*, 16(4), p. 517–532.
- Marchant, C. (2018). “New churches in Newham”. In *The Desecularisation of the City: London’s Churches 1980 to the Present*, edited by D. Goodhew and A. P. Cooper. Abingdon, UK: Routledge, p. 134–47.
- McGrath, A. E. (2018). *Theology The Basic Readings*. Chichester, UK: Wiley Blackwell.
- Murthy, D. (2013). *Twitter: Social communication in the Twitter Age*. Cambridge, UK: Polity.
- Naaman, M., Boase, J. & Lai, C. H. (2010). “Is it really about me? Message content in social awareness streams”. CSCW 2010, Georgia, USA.
- O’Dea, B., Wan, S., Batterham, P. J., Cascar, A. L., Paris, C. & Christensen, H. (2015). “Detecting suicidality on Twitter”, *Internet Interventions*, 2(2), p. 183–188.
- Phillips, P. M. (2018). “The pixelated text: Reading the Bible within digital culture”. *Theology*, 121(6), p. 403–412.
- Phillips, P. M. (2020). *The Bible, social media and digital culture*. Abingdon, UK: Routledge.
- Phillips, P. M., Schiefelbein-Guerrero, K. & Kurlberg, J. (2019). “Defining digital theology: Digital humanities, digital religion and the particular work of the CODEC Research Centre and Network,” *Open theology*, 5(1), p. 29–43.
- Piggot (2018). “Growth and decline in London Methodism, 1980 to the present”. In *The Desecularisation of the City: London’s Churches 1980 to the Present*, edited by D. Goodhew and A. P. Cooper. Abingdon, UK: Routledge, p. 134–47.
- Rogers (2018). “Walking down the Old Kent Road: new black-majority churches in the London Borough of Southwark”. In *The Desecularisation of the City: London’s Churches 1980 to the Present*, edited by D. Goodhew and A. P. Cooper. Abingdon, UK: Routledge, p. 134–47.
- Scharkow, M. (2011). “Online Content Analysis Using Supervised Machine Learning—an Empirical Evaluation”. Paper presented at International Communication Association (ICA) Conference 2011, Boston, USA, 20–30 May 2011.
- Steinhart, E. (2012). “Digital theology: Is the resurrection virtual?” In M. Luck (Ed.) (2012). *A Philosophical Exploration of New and Alternative Religious Movements*. Farnham, UK: Ashgate, p. 133–152.
- Steiger, E., Westerholt, R., Resch, B. & Zipf, A. (2015). “Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data”, *Computers, Environment and Urban Systems*, 54, p. 255–265.
- Suero Montero, C., Haddad, H., Mozgovoy, M. & Bechikh Ali, C. (2016). “Detecting the Likely Causes Behind the Emotion Spikes of Influential Twitter Users”. Presented at the *17th International Conference on Intelligent Text Processing and Computational Linguistics*, Konya, Turkey.
- Sutinen, E. & Cooper, A. P. (2020). “Interactive technologies, Missio Dei and grassroots activism”. In *Missio Dei in a Digital Age*, edited by Jonas Kurlberg and Peter M. Phillips. London, UK: SCM Press, p. 122–149.
- Sutinen, E. & Cooper, A. P. (2021). *Digital theology: A computer science perspective*. Bingley, UK: Emerald.
- Taheri, S., Mammodov, M. & Bagirov, A. M. (2011). “Improving Naive Bayes Classifier Using Conditional Probabilities”. Paper 121 presented at 9th Australian Data Mining Conference, Ballarat, Victoria, Australia, 1–2 December 2011, p. 63–68.
- Thelwall M., Buckley, K. & Paltoglou, G. (2011). “Sentiment in Twitter events”. *Journal of the American Society for Information Science and Technology*, 62(2), p. 406–418.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. & Kappas, A. (2010). “Sentiment strength detection in short informal text”. *Journal of the American Society for Information Science and Technology*, 61(12), p. 2544–2558.

- Village, A. & Francis, L. J. (2020). "Coronavirus, church & you survey: Preliminary results: June 2020". Preliminary results paper available at: <https://www.yorks.ac.uk/coronavirus-church-and-you/> (last accessed: 11 April 2021).
- Voas, D. & Crockett, A. (2005). "Religion in Britain: Neither Believing nor Belonging". *Sociology*, 39(1), p. 11–28.
- Wallis, R. & Bruce, S. (1991). "Secularization: Trends, Data and Theory". *Review of Social Science Research on Religion*, 3, p. 1–31.
- Walther, M. & Kaisser, M. (2013). "Geo-spatial event detection in the Twitter stream", in Proceedings of the 34th European Conference on Information Retrieval, ECIR, p. 356–367.
- Wilson, B. (1982). *Religion in Sociological Perspective*. Oxford, UK: Oxford University Press.
- Zandbergen, P. A. & Barbeau, S. (2011). "Positional accuracy of assisted GPS data from high sensitivity GPS-enabled mobile phones". *Journal of Navigation*, 64(3), p. 381–399.



**TURUN
YLIOPISTO**
UNIVERSITY
OF TURKU

ISBN 978-951-29-8656-9 (Print)
ISBN 978-951-29-8657-6 (PDF)
ISSN 2736-9390 (Print)
ISSN 2736-9684 (Online)