



# AI and (De)Radicalisation Interaction Study

Israel/ 6.2 Research Report

November 2022

Sophia Solomon and Kobi Gal – Ben Gurion  
University of the Negev



@ Sophia Solomon Kobi Gal

**Reference:** D.RAD [D6.2]

**ISBN:** [if published]

This research was conducted under the Horizon 2020 project 'De-Radicalisation in Europe and Beyond: Detect, Resolve, Re-integrate' (959198).

The sole responsibility of this publication lies with the authors. The European Union is not responsible for any use that may be made of the information contained therein.

Any enquiries regarding this publication should be sent to us at [Sophia@post.bgu.ac.il](mailto:Sophia@post.bgu.ac.il).

This document is available for download at [www.dradproject.com](http://www.dradproject.com).

## Table of contents

<i>List of Abbreviations</i> .....	4
<i>Glossary</i> .....	5
<i>About the Project</i> .....	6
<i>Executive Summary/Abstract</i> .....	7
<b>1. Introduction</b> .....	<b>9</b>
<b>2. Radicalisation and technology: socio-political context</b> .....	<b>11</b>
2.1 Jihadist agenda: online space for recruitment.....	11
2.2 Jewish retaliation: social media as an arena for incitement.....	13
2.3 Religious-based LBGQTQ+ xenophobia: publicity as inspiration for a Copycat syndrome.....	14
<b>.3 Theoretical concept and methodology</b> .....	<b>16</b>
<b>4. AI and (de)radicalisation: Study findings</b> .....	<b>19</b>
<b>4.1 The preventive approach</b> .....	<b>19</b>
4.1.1 The Anti-Racism Governmental Unit (ARCGU) .....	20
4.1.2 FakeReporter .....	21
<b>4.2 The integrative approach</b> .....	<b>22</b>
4.2.1 Ajeec-Nisped.....	22
4.2.2 Tag Meir.....	23
<b>5. Conclusions</b> .....	<b>25</b>
<b>6. Appendices</b> .....	<b>26</b>
Appendix 1. Interview Guide .....	26
Appendix 2. Pre-interview Sheet .....	28
<b>7. References and Sources</b> .....	<b>29</b>

## List of Abbreviations

**AI** – Artificial Intelligence

**ARCGU** – The Anti-Racism Governmental Unit

**GO** – Government Organisation

**IDF** – The Israeli Defence Forces

**IDI** – The Israel Democracy Institute

**INSS** – The Institute for National Security Studies

**ISA** – Israel Security Agency

**ITIC** – The Intelligence and Terrorism Information Centre, Israel

**KM** – Knesset member

**NGO** – Non-Government Organisation

**RoR** – Risk of Radicalisation

## Glossary

**Hamas:** political party and terror organisation in the Gaza strip.

**Mixed/Diverse cities:** cities that contain both Jewish and Arab residents.

**Hazonot Hadatit:** from Hebrew, meaning “religious Zionism”, a joint far-right religious party including the National Union party, Otzma Yehudit, and Noam.

**Knesset:** the Israeli parliament.

**Lehava:** from Hebrew “Lemeniat Hitbolelut Beeretz Hakodesh”, meaning “The prevention of Jewish assimilation with non-Jews in the Holy Land.”

**Noar HaGvaot:** from Hebrew “Youth of the Hills” – radical settlers of outposts in the West Bank.

**Otzma Yehudit:** from Hebrew, meaning “Jewish Power”, an extreme right-wing political party.

**Noam:** from Hebrew, an extreme ethno-religious political party.

**Tag Mechir:** from Hebrew, meaning “price tag” and referring to acts of violence, including vandalism and/or physical harm against Palestinians.

**Tag Meir:** from Hebrew, meaning “Tag of light” (NGO).

**West Bank:** disputed territory, including Palestinian cities and Israeli settlements, also named Judea and Samaria by the settlers.

**Zionism:** Jewish national ideology materialised by the establishment of Israel as a Jewish state based on the word “Zion” (Jerusalem).

## About the Project

D.Rad is a comparative study of radicalisation and polarisation in Europe and beyond. It aims to identify the actors, networks, and wider social contexts driving radicalisation, particularly among young people in urban and peri-urban areas. D.Rad conceptualises this through the I-GAP spectrum (injustice-grievance-alienation-polarisation), with the goal of moving towards measurable evaluations of deradicalisation programmes. Our intention is to identify the building blocks of radicalisation, which include a sense of being victimised; a sense of being thwarted or lacking agency in established legal and political structures; and coming under the influence of “us vs them” formulations of identity.

D.Rad benefits from an exceptional breadth of backgrounds. The project spans national contexts, including the UK, France, Italy, Germany, Poland, Hungary, Finland, Slovenia, Bosnia, Serbia, Kosovo, Israel, Iraq, Turkey, Georgia, Austria, and several minority nations. It bridges academic disciplines ranging from political science and cultural studies to social psychology and artificial intelligence.

Dissemination methods include D.Rad labs, D.Rad hubs, policy papers, academic workshops, visual outputs and digital galleries. As such, D.Rad establishes a rigorous foundation to test practical interventions geared to prevention, inclusion and deradicalisation.

With the possibility of capturing the trajectories of seventeen nations and several minority nations, the project will provide a unique evidence base for the comparative analysis of law and policy as nation-states adapt to new security challenges. The process of mapping this variety and its links to national contexts will be crucial in uncovering strengths and weaknesses in existing interventions. Furthermore, D.Rad accounts for the problem that processes of radicalisation often occur in circumstances that escape the control and scrutiny of traditional national frameworks of justice. The participation of AI professionals in modelling, analysing and devising solutions to online radicalisation will be central to the project's aims.

## Executive Summary/Abstract

This report presents the results of an experimental process focused on some uses of AI technologies within deradicalisation projects in Israeli society. This study was conducted with institutions (government and non-government) centred on assisting people (“backlash populations”) who are suffering from, witnessing and reporting about radicalisation. We examine how state and civic institutions use digital tools (if any) in their work with certain communities and the general public, focusing on the centrality of the connective link between online and offline radicalisation. Attempts to break this link bring different deradicalisation approaches to the surface.

After a short introduction, section 2 offers some socio-political background to put the study into context for three radicalisation set-ups: jihadist terrorism committed by Palestinians and/or Israeli Arabs; Jewish terrorism performed by supporters of far-right-wing racism and religious nationalism; and xenophobia against LGBTQ+ communities, outcast by religious fanatics. The data was based on previous reports, adding the most important contemporary developments. All three have shown increased activity and statements that emphasise the evolving bond between online radical behaviour and offline violent actions. Here we focus on three factors that relate to online/offline radicalisation: recruitment, incitement, and copycat syndrome.

The following section (3) describes the methodology used for this study. It first includes the application of insights from D.Rad reports (D3, D4), followed by semi-structured interviews (“mapping interviews”) with groups working in deradicalisation activities, supplemented with more research from social media, news reports and official data from state institutions. Here we created a six-step process: 1. Data research; 2. Detection of prominent actors; 3. Mapping interviews; 4. Analysing information; 5. AI assessment; and 6. Setting channels for AI solutions.

Section 4 elaborates on the findings revealed in the research, divided into two sub-sections (4.1, 4.2) describing the relevancy of each institution that agreed to share its daily need for, and uses of, technology. Each sub-section presents a summary of data from four institutions that cooperated with and performed the mapping interview process: The Anti-Racism Governmental Unit (GO), FakeReporter (NGO), Ajeec-Nipsed (NGO), and Tag Meir (NGO). Each participant held a key position within the organisation. We found two main channels by which the work on online/offline radicalisation can be established under the I-GAP spectrum framework: preventive and interactive.

Finally, the last section (5) presents the options for adjusting AI solutions in the fight against radicalisation. The preventive route is based on reports that the public submits to the relevant institution, which is slowly becoming a central practice in the field of online/offline radicalisation. Additional digital mechanisms are required since, even if

institutions can detect radicalisation, it does not seem yet that there are suitable tools to handle large-scale information. This involves a set of practices and particular means to extract ad-hoc threats and recognise patterns. Developing a local I-GAP lexicon by the relevant institution in Hebrew can be the first step to implementing the detection of users at risk of radicalisation (RoR) (see D6.1). RoR refers to users who spend significant time online and are recognised as highly exposed to incitive content.

The integrative channel shows that civic actors emphasise the connection with RoR and/or backlash populations as part of their daily routine. NGOs' knowledge can be valuable to the state while developing deradicalisation initiatives. Working with NGOs has shown that many civic initiatives involve educational activities, which can be used as an anchor to future work with young adults from different sectors. Here we find that I-GAP surveys can be used for data collection. AI technologies can produce analysed data reports for working with, and keeping track of, occurrences within backlash populations.



# 1. Introduction

Increasingly, the internet and social media have contributed to and/or mediated the radicalisation of individuals towards political or social extremism. Identifying online radicalisation has become one of the EU's central political and security challenges.

The goal of this report (D6.2) is to investigate the use of AI technology to detect and alleviate radicalisation in at-risk populations within Israeli society. To this end, we conducted interviews involving relevant stakeholders and examined the compatibility of AI approaches (i.e., machine learning) for implementation among relevant civic and state actors. We focus on two settings, as defined in D3.1 – separatist xenophobia and racism, and ethno-religious based “Jewish terror” – as identified in prior reports. This report directly relates to D6.1, which focused on the use of AI to detect the risk of exposure to radicalisation (RoR).

The research efforts of WP6 extend the analysis of online behaviour for detecting radicalised behaviour into interactive strategies for mitigating it (D.Rad proposal, 2022, 3-4). There are two types of input to this process. First, digital traces of the contemporary online interactions of individuals and groups, including social media, blogs, and discussion forums. Second, stakeholders' reactions to radicalisation scenarios under the socio-political framework the other WPs provide (WP3; WP4).

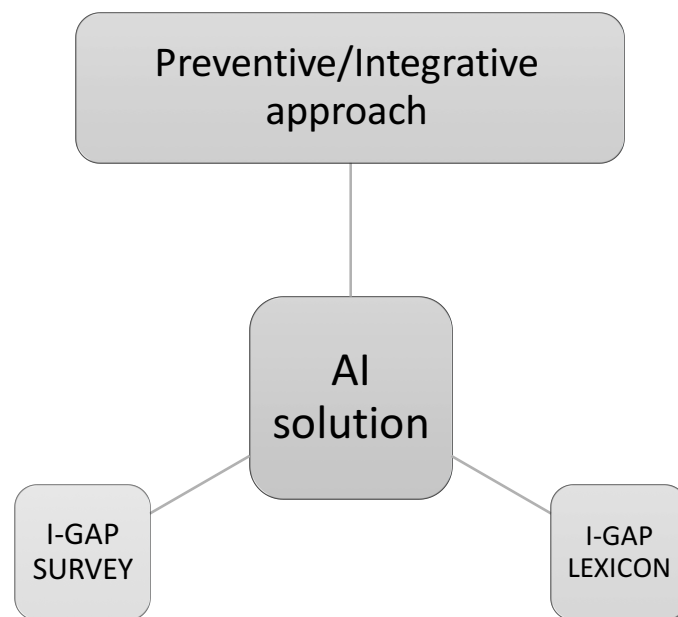
This study was conducted upon institutions centred on assisting people suffering from, witnessing, and reporting about radicalisation (“backlash populations”). We examine how state and civic institutions use digital tools (if any) in their work with specialists and the general public, focusing on the centrality of the connective link between online and offline radicalisation. Attempts to break this link bring different deradicalisation approaches to the surface.

In this report, we attempt to uncover tendencies among relevant stakeholders responding to radical activity occurrences in cyberspace and its exploitation by extremists, and answer the following questions: 1. How are digital tools used (if at all) within the institution's activity framework? 2. What are the needs and requirements of NGOs and state actors to apply AI as an assistive tool? 3. In what way can we apply big data categorisation and the I-GAP thematic approach to relevant institutions? 4. When it comes to online/offline activity in times of increasing radicalisation, what can we learn from attitudes towards AI as an assistive tool?

To address the first two questions, we analysed online data from news media and governmental reports, followed by social media publications, locating the relevant organisations for this experiment. We have come up with four institutions that have expressed the desire to share their knowledge and needs. Next, we have built a semi-structured interview guide oriented to examining the current usage and conditions of digital tools of any kind (“mapping interviews”). To address the third question, we have made an individual summary of each mapping interview and the work process with each institution, comparing similarities and differences between tendencies on the significance of the digital sphere in terms of big data, thematic categorisation, and future aspirations. Finally, considering the findings from the interviews and working processes, we answer the fourth question by offering recommendations on two optional future channels of work with affiliated institutions, as well as the future challenges and accountabilities associated with implementation.

We have found that NGOs and GOs address radicalisation using two main approaches (sometimes intertwined), both basing their work on interacting (one way or another) with the public. The first is the preventative approach, which includes measurements detecting radicalisation, reporting it to state authorities, and revealing information to the public. During confrontations, the necessity for handling big data arises since the quantity of information regarding radical activities increases. The central issue from this approach is the struggle to categorise and analyse big data from social media and news networks. One of its primary necessities is creating reports that reflect radicalisation trends. Creating an I-GAP lexicon on which the reports will be based can be helpful.

The second is the integrative approach, which cooperates with backlash populations via collective activities (e.g., protests, parades, round tables, etc.) and addresses radicalisation-related issues. It also includes special training programs on prevention, targeting community leaders, teachers, and managers of public institutions. The critical matter is utilising AI technologies to expose social media users to deradicalisation activities and information. The spread of I-GAP surveys to backlash populations and other social media users, followed by analytical reports, could be the first step in receiving data to adjust integration plans.



**Figure 1.** Experimental findings: from approach to a solution under the D.Rad framework

Applying AI and big data instruments requires NGOs and GOs to combine manual and digital work, fusing existing knowledge with new tools. This might be challenging for many institutions that so far rely on a limited number of information sources.

## 2. Radicalisation and technology: socio-political context

To approach this acute challenge, we first recognise the prominent trends that best express the online/offline connection. In countries like Israel, handling an existential crisis and not finding a long-term solution to the Israeli-Palestinian national conflict, claims of distress and frustration appear so often. In earlier reports, we have focused on three radicalisation set-ups: Jihadist terrorism committed by Palestinians and/or Israeli-Arabs; Jewish terrorism performed by far-right-wing racists and religious nationalists; and xenophobia against LGBTQ+ communities, outcast by religious fanatics (see D3.1; 3.2). In the past year, all three have shown increased numbers of acts and statements that emphasise the evolving bond between online radical behaviour and offline violent actions. Here we focus on three aspects: recruitment, incitement, and copycat syndrome.

### 2.1 Jihadist agenda: online space for recruitment

While Jerusalem maintains its characterisation as a city in conflict due to its religious complexity, it has affected other towns, and pinpointed connections between offline activities and online organisation and the spread of violence. On 10 May 2021, during the annual “flag parade” on “Jerusalem day” in the old city quarters, Hamas launched rockets towards the Mount Temple, located close to Al-Aqsa Mosque, which drew a comprehensive military response from IDF. The “guardian of the walls” operation lasted 11 days against Hamas’s perpetrators in the Gaza strip (INSS editors, 2021). The military operation did not last long, and even though it led Hamas to cease fire from Gaza, it did not prevent the murderous attacks committed by individuals. Part of it is the ability to recruit Palestinians from different areas of the West Bank and Israel. Hamas uses social media networks to spur violence in the area surrounding Al-Aqsa Mosque by recruiting the locals (Ben Menachem, 2022), as found on multiple occasions (Bohbot, 2022; Kobowitz, 2022); its cyber branch unit also contacts minors in an attempt to incite and encourage them to commit violent acts (ISA, 2022). It follows a parallel path of other cyber attempts to extract information by hacking Israeli soldiers’ cell phones and social media accounts (IDF, 2020).

One point concerned the authorities most – the involvement of education personnel in some ongoing jihadist attacks. The murder of Eliyahu Kay on November 2021 in Jerusalem was committed by a former teacher known to the police since he was active in the Hamas organisation before the attack (Rot-Avneri, 2021). Several months later, two more vicious attacks occurred. The murder that cost four lives in Beer Sheva was the second known case where a former teacher performed a jihadist attack upon Israeli civilians. The teacher was arrested in the past and spent four years in prison for joining ISIS (Kan editors, 2022). The attack was followed by another when two ISIS perpetrators in Hadera injured five and killed two border police soldiers by shooting towards a bus station (Ravid et al., 2022).

Even though the three perpetrators of the attacks above declared that ISIS was the leading driver of their actions, Israeli security scholars and intelligence do not see it as a central enemy. Still, as a possible amplifier of radicalised behaviour, ISIS does not hold on to Israeli-

Arab citizens or Palestinians simply because it stifles their national ambitions for an autonomous state that clashes with the idea of the Islamic state. But even so, it still affects mechanisms of action and similar narratives of jihadism, one of which Arab citizens usually condemn (Schweizer et al., 2022). Some of them are convenient targets for recruitment by Hizballah in Lebanon (Iran's proxy), Hamas in Gaza, or other terror groups. There have not been too many examples of Israeli Arabs that committed attacks recently, but even one is enough to tear apart the delicate relations between populations. Backlash reactions emerged from right-wing political extremists and also from the public in general.

Politicians from the far-right immediately connected the perpetrators to Arab society, saying that there has been a pattern of radicalisation among Arabs and Bedouins in recent years (Srugim editors, 2022). They raised far right-wing Knesset Member (KM) Itamar Ben Gvir's offer, suggesting legislating for a particular law prohibiting teachers from supporting terrorism (Rot-Avneri, 2021). It attracted legislators' attention, raised the issue to a higher priority, and created a series of discussions within the Knesset committee on education, culture and sport, on preventing terrorism from emerging from the next generation of kids and teachers. When an Israeli-Arab teacher supported the perpetrator that committed an attack by saying, "God Bless the Heroes" (Sela, 2022), it showed the blur between normal support for the Palestinian struggle and embracing the jihadist mechanism.

The debates among the committee of the Knesset raised a significant assumption: Arab teachers who get their professional training from the Palestinian Authority (PA) are exposed to incitive content and therefore need to be under closer supervision. The Israeli Security Association (ISA) used to perform background checks until recently, a move that was cancelled by a former minister of education, Yuli Tamir, who claimed it contradicts democracy and taints a whole population (Trabelsi-Hadad, 2022). According to the announcement by the head of the committee, the Israeli education system requires information about terror affiliation among education workers since over 3000 Arab teachers are working within the Israeli system. Between 2010 and 2019, more than half were trained in the PA (Knesset, 2022).

The Arab members of the Knesset protested, saying it was a racist act, and mentioned Jewish teachers with similar affiliations. Accordingly, this racist framework includes all 22% of teachers that got their education within Palestinian education facilities. Yet the chairwoman of the committee, Sharen Haskel, said that the education system is responsible for the messages children get from their teachers that are "preaching" in addition to teaching support for acts of violence, putting "children at risk for future radicalisation" (Knesset, 2022). Here it is relevant to mention what we have referred to as 'Risk of Radicalisation' (RoR) (see D6.1) and its joint manifestation inside the classrooms at schools. This means teachers, considered model figures, lead some kids to be exposed to RoR by attending lessons and spreading their personal views on social media.

The Ministry of Education held close discussions and started cooperating with ISA to locate teachers that raise "red flag" alerts on intentions and connections to terrorism (Adalah, 2022). Accordingly, this step violates the court's ruling from 2004 regarding Arab teachers. But here, it is essential to mention that teachers who commit terror and are even associated with violent content is a topic that doesn't only concern the Israeli government and legislators. A UN Watch

report exposed that UNRWA (United Nations Relief and Work Agency for Palestinians) workers committed over 100 cases of incitement against Israelis and the Jewish community in 2020, and estimated that there are more than cases among its 30,000 employees, based on data from their personal Facebook pages (Maariv editors, 2021).

## 2.2 Jewish retaliation: social media as an arena for incitement

In Beer Sheva, the jihadist attacker published his intentions on Facebook. After his attack, the police and Justice Ministry's cyber unit located over 5800 posts identified by security authorities as incitive, not following social platform regulations. Of them, 2950 were removed by network managers (Fareed, 2022). A similar pattern of using social networks as a platform for spreading violence, whether incitement or a clear call for action, was identified by security authorities in Israel during the attacks in the mixed cities. TikTok was at the centre of violent video footage of Palestinian perpetrators (Fareed, 2022). While more than 4360 missiles attacked the Israeli population during the Guardian of the Walls, the most concerning events were the clashes between Jews and Arabs who share their living space ("mixed cities").

During 11–14 May, Jewish and Arab civilians and residents of mixed cities (e.g., Aka, Lod, Beer Sheva, Bat Yam) attacked each other brutally. For four days, citizens took the law into their own hands and acted instead of the local and central government (ITIC, 2021). News channels filmed one specific case in Bat Yam, documenting a raging mob attacking an innocent Israeli-Arab resident, Saeed Moussa, who had just passed by with his car. Dozens attacked him; a few were criminally charged with attempted murder, a brutal terror act, vandalism and more. It was revealed that arrangements and plans were performed before this event on social media, spreading an incitive "call for Jews" to retaliate against Arabs in the streets of mixed cities and vice versa (Peleg, 2022). Hence, the violence within the diverse cities during Palestinian terror attacks also emphasised Jewish terror attacks that were organised using social media.

What happened during the riots reflected a more profound challenge that has occupied the minds of many states: the social media arena. Some of the effects of a two-front battle hold increasing feelings of fear in the public, the option of organising participation in violent acts, as well as influencing worldwide public opinion, thus attracting antisemitic manifestations (Orpaz and Siman Tov, 2021). The complex digital arena shared by Jews and Arabs includes Facebook, Twitter, WhatsApp, Instagram, and Telegram, with the utilisation of short incitive videos via TikTok, which has become central to the transmission of messages. The messages spread rapidly, increasing the phenomenon of fake news (Ibid, 3-5). Hence, disinformation is used to sow public panic, increasing the general population's fear level (Ibid, 7).

The Israeli State Comptroller's recent report criticised the state's lack of interference and failure to offer civilians protection, based on the murder of three civilians, hundreds injured, and thousands arrested. It estimates that 3200 Arabs and 240 Jews participated in violent acts led by incitement on social media, among other factors (State Comptroller, 2022). This critique uniquely addressed the lack of intelligence and preliminary information regarding the intentions and practice of ad-hoc vandalism. Following the riots within the diverse cities, police commissioner Ya'akov Shabtai recently caused a dispute when he suggested shutting down access to social media platforms in times of extensive riots (Haaretz editors, 2022). The

minister of internal defence objected to his statement, firmly claiming it contradicted the democratic regime's values (Walla editors, 2022).

And yet, the public discourse of “us vs them” deepened polarisation between Jews and Arabs in a fragile texture of shared living. Public spaces that all citizens share have become places to avoid, mostly in mixed cities where Jews and Arabs share the same commercial living areas. Lack of trust in the official authorities, such as the police, is also present among Jews and Arabs since they did not work fast enough to prevent the horrors on both sides. And so, the days of May 2021, in which violent riots were present in the streets among both Jews and Arabs, mainly showed the vivid connection between online incitement and offline violence.

### 2.3 Religious-based LGBTQ+ xenophobia: publicity as inspiration for a Copycat syndrome

Social media is part of incitement mechanisms when an imminent threat comes from posts that might inspire other youngsters to commit violent acts. In Israel, ‘Copycat’ is a phenomenon that usually emerges in days of a bloody conflict, built upon social media incitement, encouraging individuals or groups to execute similar actions as “lone wolves” (ITIC, 2022, pp. 2-4). It is not new, but is present and amplified by social media's primary tool: publicity. The Israeli police said its worst fear came vivid when Israeli Arabs that committed attacks inspired Palestinians from the West Bank to imitate murderous actions (Harel, 2022). The concern continued as the IDF prepared itself for such mimicry scenarios within the West Bank (Blumenthal et al., 2021).

One form of inspiration for copycat actions that has recently developed is manifested via “terror fashion”, a new and popular mechanism that connects offline violent dress-code with online publications and purchases. T-Shirts with a print of an M-16 gun produced in Turkey spread across Arab villages and the PA jurisdiction, using the spread of violent images as inspiration for mimicking extremism. As a response, police have warned fashion stores to stop providing this type of merchandise and selling it, claiming it is prohibited and considered an offence (Eli, 2022). The photos and physical appearances of kids, students and adults wearing gun T-Shirts have caused rage within the Israeli public. The M-16 image represents a violent statement against Israel's military weapons and favouring Hama's armed struggle (Shapira, 2022; Cohen et al., 2022). As a response, separatist groups have printed counter-T-Shirts with the slogan “Death to Terrorists”, used by far-right-wing fanatics (Am Israel Chai, 2022). Again, social media is a central tool that makes it possible for radical ideas to reach their intended audience.

The conflict between Palestinians and Israelis is not the only case where copycat syndrome appears. May is also the national month of Pride events. One can find events across the country such as parades, marches, round-tables, theatre shows and films, usually with municipal support. Even though the “Pride Parade” in Jerusalem has been developing over the past 20 years since the first march took place, it is still not normative for religious residents, who do not agree to allow such a march to take place within the city's borders. The agents spreading hate discourse against the LGBTQ+ community are part of a fanatic religious movement (see D3.1) in which electorates are represented by far-right wing religious Zionist parties such as the “Noam” party.

Some rabbis claim that the LGBTQ+ community at its core is against religion and damages the “proper” structure of a family. For example, Rabbi Kostiner, a member of the Noam party, said to his followers in a live Twitter video, “do not be ashamed to say ‘gays go home’”, adding that LGBTQ+ people are re-engineering the perception of family life (Etinger, 2022). Even though Israel has ministers and other public officials from the LGBTQ+ community, the atmosphere is still hostile to the community from some members of the Knesset. For example, KM Avi Maoz from Noam compared gay people to paedophiles, saying there is no difference between being attracted to kids and being attracted to the same sex while giving a speech in parliament (Ademkar, 2021). In 2019, Maoz, who was not a public representative yet, compared the Nazis’ actions to the LGBTQ+ community (Jewish Pluralism Watch, 2019). Maoz represents a movement that does not support LGBTQ+ rights or even political freedoms such as engaging in a parade. This movement is not necessarily combined with politicians but has little public support from more traditional crowds (see D3.1).

In this context, the parade of 2022 also involved threats that uncovered a copycat phenomenon derived from ultra-orthodox separatists. It emerged as a threatening message



**Figure 2.** "we will not allow the Pride parade to take place in Jerusalem. Jerusalem is the Holy Land. Shira Banki's fate awaits for you..." (from Hebrew).

that was sent to the parade organisers the day before via Facebook and other social media channels, explicitly using the name of Yishai Schlissel, an ultra-orthodox separatist who injured and murdered participants in past parades and is serving in prison (see D3.2). Social media activity can be an example of radicalisation “in action”. As such, a fake Facebook profile was opened under the name “Yishai Schlissel’s brothers”, saying the organisers would share the same destiny as Shira Banki (a 16-year-old murdered in the 2015 Jerusalem parade) (Chalabi, 2022). This statement (see Figure 2) was taken seriously by the authorities. As a response, the state has increased the number of security forces in the area. The parade in Jerusalem took place on 2 June, escorted and secured by thousands of police officers, including the attendance of the minister of homeland security. Israeli parliament chairman, KM Miki Levy, attended the parade as the first police county officer who approved it in 2002 and



said he was horrified by the incitement that came before the event (Hakmon and Bagano, 2022).

The latest extreme event that might suggest radicalised activities spread nationally was the cancellation of the first-ever planned Pride parade in the southern city of Netivot, located in the southwest periphery and characterised by an orthodox population (Yefet, 2022). Locals have also started an online petition in protest against it, reaching 97% of the original goal of 5000 signatures (Azuma, 2022). The government did not intervene or demand to continue with the plans. The organisers did not get the support of the local mayor even though Netivot's population includes an estimated 10% LGBTQ+ members (Yefet, 2022). But eventually, after their car window was shattered and a gun bullet was fired through the mailbox of their private homes, the organisers had to back down from their will to march freely across the town (Levy, 2022). In parallel, the issue of Netivot highlights the differences between the periphery and the centre since the Pride community, who requested the local police to alter the parade into a protest, was declined (Steinman, 2022). Here, the local extremists are imitating death threats against the organisers, like in the Jerusalem case and others (Yefet, 2020), emphasising incidents of online inspiration for offline actualisation.

### 3. Theoretical concept and methodology

WP6 aims to capture critical moments affecting radicalisation by developing specific computational data science tools. Considering the main issues above that pinpoint the essential aspects of radicalisation and technology set up, we are exploring the option of utilising digital tools to assist the ongoing activities of dominant actors within the field. We first define the eminent concepts in data science research, followed by the methodological process.

Artificial intelligence (AI) does not have a single definition. Still, the common assumption is that it refers to computer systems that are similar to, or can function like, the human mind, considering that both cannot become identical. AI can include a variety of techniques, such as heuristic search, game-playing, expert systems, data mining, etc. The AI field has various subdomains that have formed within it, like reasoning, planning vision, natural language processing, robotics, and teaching machine self-learning skills (Wang, 2019, 7-8). Hence, "machine learning" is a set of methods designed to perform cognitive tasks, meaning the artificial ability of the system to learn from examples and apply the knowledge to new information (Even and Siman Tov, 2020, p.29).

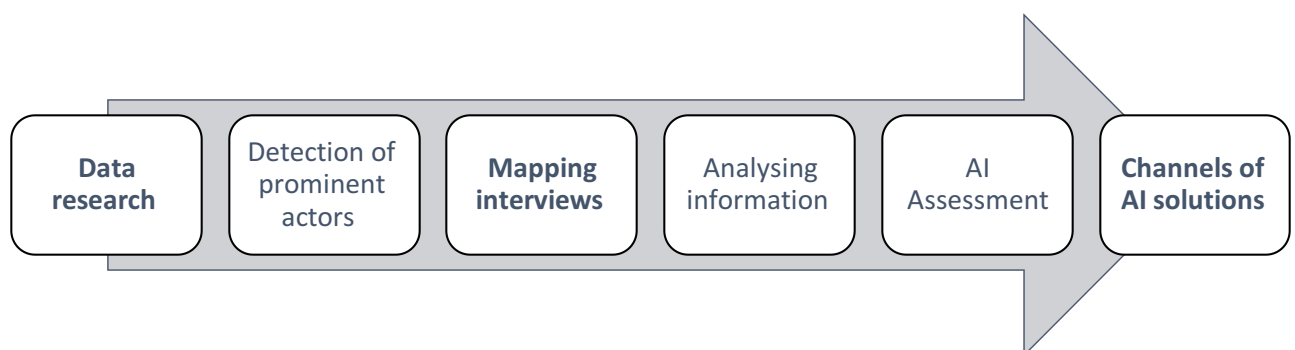
Among others, AI is needed in spaces containing large amounts of data that are required to be analysed rapidly. "Big data" is "a database of data of various types and in diverse formats that come from many sources - from which we try to extract information and draw conclusions" (The Hebrew Linguistic Academy, 2022). It is a field of information that can assist other areas that involve primary civic considerations (i.e., security or public health) (The Israel Academy for Science and Humanities, 2022). Systems that combine machine learning and big data offer the analysis of vast amounts of information by automatic software that can provide knowledge, suggestions or services (e.g., Amazon's "Alexa") (Knesset research and data centre, 2018, p.1-2), each containing relevant features that define it.



A significant amount of data from social media platforms can be used as a sub-type of big data. Social media networks such as Twitter, Facebook and YouTube serve as a platform for the broad exchange of ideas and opinions, sometimes exploited by political groups to publicise their views. In the previous report (D6.1), we used machine learning to predict critical events which lead to radicalisation behaviour by forging a unique lexicon accumulating polarised terms and themes. Our model can potentially assist those providing support for people at risk of radicalisation (RoR) in real time, enabling fast detection even for users with relatively short historical activity on the site.

Israel joined the GPAI (Global partnership for Artificial Intelligence) in 2021 (Eichner, 2021). The lunge of “the national plan for AI” aims to combine it in the public sector and the market. It amplifies the government’s will to fulfil its image as a cybernation (The National digital division, 2022). In our experiment, we approached existing state and civic organisations affiliated with deradicalisation projects centred on public interaction that were willing to participate in our theoretical investigation. Here we focus on the initiatives and actions performed by government organisations (GO) and non-governmental organisations (NGO) affiliated with educational and social interactions within the radicalisation set-ups. We gave special attention to the centrality of social media platforms and the current usage/necessity of AI tools for (de)radicalisation initiatives.

Following this theoretical framework, we develop a methodology based on preliminary data from semi-structured interviews (“mapping interviews”) supported by secondary resources, including samples of big data datasets provided by the participants. The data assists in determining if state and civic institutions can implement the model above (RoR), and otherwise creating additional AI tools needed within the field of radicalisation and online/offline traits. The process involved the steps outlined in Figure 3.



**Figure 3.** The six-stage methodological process

### **Data research**

Accumulating insights from D.Rad reports (D3, D4) involving groups working in deradicalisation activities, we added more research from social media, news reports and official data from state institutions to find suitable participants.

### **Criteria for choosing an institution**

- a) Multi-cultural, co-existence progressing organisations,

- b) who work all year long implementing programs of deradicalisation and/or respond to radicalisation (events or individuals) legally,
- c) and/or work with backlash populations and/or possible RoR groups and individuals,
- d) who have technological needs in observing, organising, and categorising big data,
- e) and/or wishes to encourage digital options to assist deradicalisation initiatives.

### **Detection of prominent actors**

Participation involved sharing information and insights from their daily work and pinpointing tendencies regarding the online/offline sphere. With each chosen institution, we held a few preliminary talks (via digital communication) in preparation for the central interview to identify technological needs and AI affiliations. For that, we created the "pre-interview sheet", mentioning significant issues that emerge from the institution's activity (see appendix 2). The guidelines include preliminary research based on at least 1–3 preparation meetings with chosen institutions. The goal was to a) find the person most relevant within the GO/NGO to apply technological-experimental research, and b) map out existing practices that the institution follows, focusing on the present and future vision of AI and D.Rad activities.

### **Mapping interviews**

The semi-structured interview guide includes five central themes, each offering relevant questions (see appendix 1):

1. Digital usage – to check what digital operations are available to the institution and how it considers the use of technology within its framework.
2. Technology and data – to find out how the institution handles the categorisation and general use of data.
3. Technology and social media – to examine how common social media channels are, and if the institution sees them as valuable and valid information sources.
4. Digital ambitions – to explore what kind of digital ambitions the organisation has, and which ones present an option for joint work.
5. New digital tools – to explore how AI tools can assist the institution (e.g., handling data, developing communication, adjusting digital mechanisms to collect data, etc.).

### **AI Assessment**

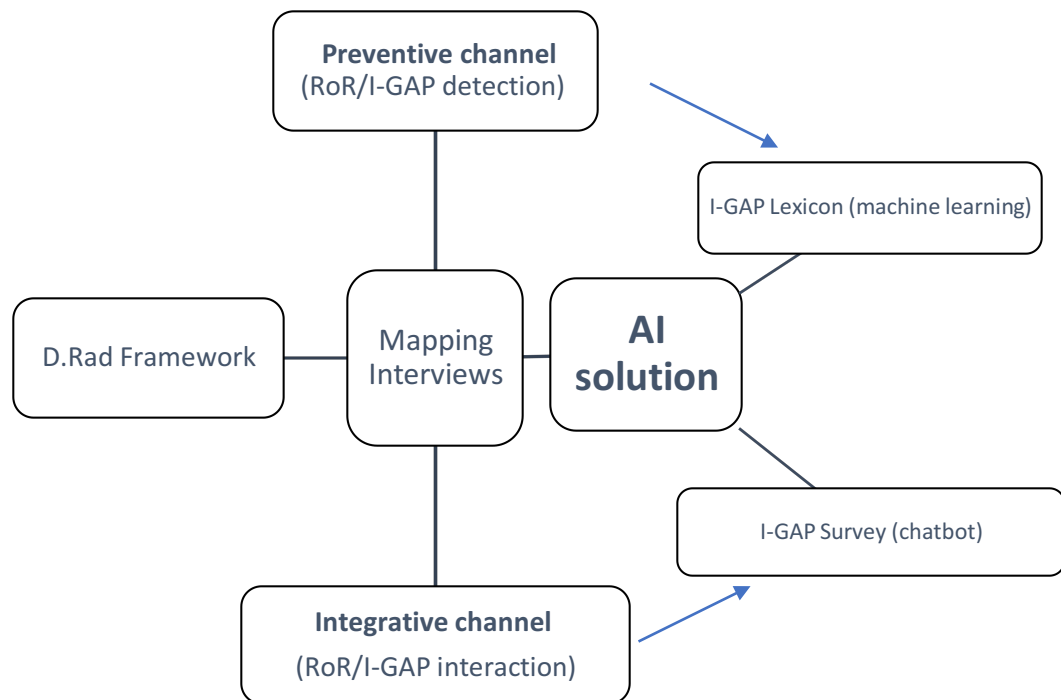
At this stage, we assess options and/or D.Rad activity in practice, determining if AI can assist the particular institution's activities through the information collected earlier. It also relies on the actor's will and commitment to continue with the interaction and wishes to direct resources to perform the cooperation.

### **Channels of interaction**

The final stage is determining the action channel through which we can implement AI technology, meaning what type of working path can benefit and complement current and future work.

## 4. AI and (de)radicalisation: Study findings

Here we present a summary of data from four institutions that cooperated with the study and performed the mapping interviews. The participants gave their consent to participate in the study verbally at the beginning of each recorded interview. The names that appear within the remarks are aliases, to protect the interviewees; each holds a key position within the organisation. We found two main channels through which work on online/offline radicalisation could be established: preventive and interactive (see Figure 4). Each can assist and co-depend on the another, but they require different tools.



**Figure 4.** Process and channels to solutions

### 4.1 The preventive approach

Here we found preventive measures by a government institution and a civic one aiming to stop radicalisation from spreading further, or before extreme actions occur as a result of a dispute. Both rely on public data and reports from individuals and groups regarding the use of physical and verbal violence derived from racism, xenophobia, nationalism, and ethno-religious agendas. This working channel can prevent further harm by breaking the link between online and offline incitive activities, using the 2.0 arena to access valuable information, and relying

on the public's experience. Two institutions are performing their central work with the preventive channel.

#### 4.1.1 The Anti-Racism Governmental Unit (ARCGU)

The Anti-Racism Coordinating Government Unit (ARCGU) under the Ministry of Justice (GO) is a result of protests from the Ethiopian community against racism and police brutality (ARCGU, 2019). The unit works as an institutional entity that could give a proper response to racism in state institutions. Global awareness has raised the subject of police brutality and heavy-handed policing, and revealed that the individual story of the Ethiopian soldier was part of a much larger phenomenon of group association-based racism taking place within official state institutions. The crucial aspect of this incident is that even though the soldier's IDF uniform was supposed to "protect" him from random violence, it did not prevent the policemen from acting out of clearly ethnically based discriminatory motives (Akir, 2021).

As a result, in 2016, a governmental committee was appointed. The "Palmore Committee" was set up and asked to examine the question of police brutality within society. It concluded that the state was actively racist towards the Ethiopian community, and followed by establishing a specific governmental unit to handle institutional discrimination and respond appropriately to racism-based charges (see D4.1). Even though its initial focus was on the Ethiopian community, it has now expanded its investigations to include all types of racism performed based on group association (e.g., ultra-orthodox based on cultural-religious association). Three main branches of practice lead the unit: public complaints, progressing lateral moves, and training and education (see D4.1).

The unit operates in the prevention channel while taking legal actions after receiving information about racism within public institutions, meaning the places where the authorities act unequally and abuse power against minorities. Based on public reporting via a digital form on the website, it analyses the data it receives manually to submit evidence to legal authorities. It works with the police, legislators and the justice department and progresses punitive actions. In addition to the reports from the public, it practises internal monitoring by receiving digital alerts from all over the web, including news sites, social media posts, and more. Finally, it publishes its data annually, is open to the public, and uses legislators and other organisations. The unit also plans to initiate educational programmes by developing teaching methods regarding radicalisation (Akir, 2022).

According to the unit, it receives data from the public that can help detect sources of racist behaviour and follow patterns of extremism within public spaces. For example, the team discovered that parades had become a centre of attention in the past year since extremists are using the public space to enhance racist ideology (e.g., the flag parade in Jerusalem; the Pride parades across Israel):

The online platforms are used to organise and amplify existing events by participating in offline events. **The combination of offline/online initiatives is expressed in public parades, following online 'trends'**. Thus, navigating social media data has become a more significant task for authorities, so social media monitoring has become a central issue concerning the unit (Akir, 2022).

Functioning with a low budget and lacking personnel might become a central challenge of needing a mechanism for handling big data from social media. The unit lacks tools to monitor and control big data. This affects policy recommendations and missing data patterns (e.g., the number of new groups created on social media, violent discourse in comparative analysis, and ways of use among politicians and public officials, etc.). Even though it is active on Facebook, it wishes to be more involved in the sphere through TikTok, Twitter, etc. Theoretically, the unit could use AI to detect, trace, and analyse future data in this case. Still, it needed more human resources to apply such tools, combining manual and digital work.

#### 4.1.2 FakeReporter

This NGO's main goal is to use "mass wisdom" as a tool for reporting violence on the internet, and social media in particular. The NGO operates within digital platforms encouraging the public to pass on information that does not follow democratic and human rights values:

[...] a civil effort against the malicious online activity. Leveraging crowdsourcing via its reporting platform, FakeReporter exposes and disables attacks on democratic systems, inauthentic influence campaigns, fraud, hate speech, incitement, conspiracies, and lies (FakeReporter, 2022).

In this way, the NGO is giving people the power to fight other groups and individuals spreading radical, violent, and baseless information online that state institutions are failing to uncover. Accordingly, some information on social media and other digital platforms is invisible to the state. Sometimes no security institution can locate data in real-time as the general public does. Through reports from the public via a digital reporting form, the NGO is a reporting channel to communities at RoR and/or backlash populations, similar to ARCGU. It allows users to notify the NGO about fake accounts and users, violent discourse and personal offences experienced on social media platforms (Twitter, Facebook, Instagram, and soon TikTok). The types of violations are varied. Not all are committed with the intention of harm but emerge from personal information sharing (Or, 2022).

Public reports allow the NGO to fulfil these main goals: 1. Research to develop more theme categories that allow 2. Operational reporting to the authorities by routing the reports to relevant legal and security institutions and influencing the reduction of radical content. 3. Legal work: the NGO also assists the parliament's legal actions by advising and providing relevant data (Or, 2022).

FakeReporter uses digital tools as its operative mechanism. It is very active on Facebook and Twitter. It receives some of its data regarding fake news from online channels documented by the public and its volunteers. Currently, it carries dozens of thousands of reports, offering a digital plugin that will appear on some platforms and directly link to online reporting, increasing the amount of data. The NGO wishes to empower the public and address the fact that

citizens are exposed to **more digital-based offences** and violent discourse than state institutions consider (Or, 2022).

The NGO alerts state authorities after revealing networks of radicalisation (e.g., Iranian hackers that try to connect with civilians). As the reports will increase over time, there is a need for advanced AI tools to assist in manual and digital categorisation and analysis of information in Hebrew.

## 4.2 The integrative approach

This section presents NGOs that use an integrative approach. Their essence is cooperating with the public, state, and civic actors to replicate deradicalisation initiatives. NGOs offer participative activities like parades and round-tables, emphasising the need for safe space areas in times of tension. They also advance educational programmes for youths in RoR and towards backlash populations negatively affected by radicalisation. In times of radicalisation between Jews and Arabs, these NGOs operate ad hoc to comfort and assist minorities and individuals.

### 4.2.1 Ajeec-Nisped

As an educational organisation, its primary goal is to strengthen the connection between Arabs and Jews within Israeli society. It expressed the difficulty in sustaining an ongoing partnership between weak and robust populations (the Jewish majority and Arab minority). Since 2000, Ajeec-Nisped has operated as an Arab-Jewish institution and offers education initiatives located in the southern region of Israel:

The organisation's goal is to promote leadership and social involvement among teenagers and youth in Arab society, community development in the Arab-Bedouin sector in the Negev, and the development of a shared society for Arabs and Jews throughout Israel. [...] The Arabic word ajeec means "I am coming to you", and this is the principle that motivates us – we approach each other (Ajeec-Nisped, 2022)

At the moment, it is progressing three primary programs for youths and young adults coming from Arab communities (i.e., Bedouins of the south): 1. Cooperative society – including schools and a “transitional year” – trains youths from high school until after acceptance to university or a professional workplace. 2. Youth leadership – developing social activism and enabling tools to return to the community and improve it. 3. Social resilience, which involves assisting women's employment, health programs, etc. The NGO mentioned that the existing educational process might take years until concrete results appear. Thus, it's adopting a new approach named “from meeting zones to shared spaces”:

Its purpose is to make mutual usage of physical and virtual meeting zones into shared spaces that create closeness between populations that share the same places (i.e., shopping centres, sports centres, etc.) (Uri, 2022).

Existing cooperation with the government, including future budgets, relate to education, health and especially the transitional year program. Ajeec aspires to increase the number of youths in the program from hundreds in 25 cities across Israel to thousands. There is also a plan that leans on the cooperation of local municipalities (Jewish and Arab) with geographic proximity,

locating opportunities for joint spaces, and then working together on establishing facilities and economic initiatives:

As a pilot, the NGO has passed questionnaires to participants in the educational program “transition year” to identify tendencies regarding Israeli society integration and the fulfilment of political rights. It was given twice – at the beginning of the yearly program and its end. It uncovered a noticeable decrease in their self-belief in civic fulfilment since they better understood their realistic chances of integration (Uri, 2022).

The NGO wishes to adopt new programmes as a response to the connection between ongoing social media radicalisation and violent results within mixed cities in Israel (after the events of May 2021). On the interview day, a terror attack occurred in a central shopping centre of Beer Sheva, a diverse city that is a meeting place for Jews and Arabs. The perpetrator was an Israeli-Arab citizen from one of the Bedouin communities of the region. This attack emphasised the need to progress more short and long-term responses. Representatives told us that a lot of the communication among the Arab population happens within social media platforms such as WhatsApp, Instagram, Facebook, etc., even though they are considered traditional. It explains the socio-political processes that Arab society is undergoing, moving towards modernisation and dissolving the conventions of the past.

One of the initiatives Ajeec aims to develop in future work with bloggers and YouTubers is to raise their awareness of how to act when a terror event occurs. One of the problems is that inflaming both the Jewish and Arab sides can cause further violence. Ajeec-Nisped wishes to create more accessible digital tools for sustaining connections between Arabs and Jews that lose touch after finishing joint educational programs. It showed interest in educational initiatives involving the distribution of IGAP surveys that will enable mutual information input. The organisation will receive information about the place of radicalisation in the perception of its members. The members will receive the same data after an in-depth analysis of recommendations and adapting the system to emerging trends among its participants.

#### 4.2.2 Tag Meir

This NGO operates in multiple channels of D.Rad activity in the battle against religious-based racism, as part of a campaign to support democratic values. Tag Meir seeks to transcend religious divides, enlisting support from the Israeli spectrum, from secular through reform and conservative to orthodox and ultra-orthodox. The organisation also creates educational programs for schools and youth movements that deal with human dignity, respect for different cultures and religions, and respect for members of minorities, refugees, and foreigners. It also condemns perpetrators of racist actions. Tag Meir is an exemplary NGO that follows the connection between online/offline activities (Eli, 2022).

The NGO operates counter-actions using tolerance discourse, such as the “flower parade” (see D4.1). The parade takes place annually in Jerusalem in response to incitement and vandalism emerging from the “flag parade” (see section 2), in which flowers are given to Arab merchants in the area. Its presence is vital to social radicalisation that involves incitement. Usually, it has a couple of thousand participants; NGOs hope to increase this number in time:

The language of flowers has not been common in Israel in the past two decades, where words like 'peace' and 'tolerance' seem irrelevant to the harsh reality. And despite that, we are evolving to try and change the discourse of violence. If we could make 10,000 people participate in the annual flower parade, it could eliminate the flag parade and its incitive actions against the Arabs (Eli, 2022).

Flowers are a theme for the NGO, since it also sends them to victims of Jewish terrorism (see D4.1). Tag Meir's goal is to comfort casualties of terror attacks and strengthen the connection between Arabs and Jews in Israeli society. It responds to price-tag attacks against the Arab population by doing the opposite of terrorism: spreading positive actions to remind society that not all of the extreme religious Zionists in Israel are acting in a racist manner. Another example involves NGO activity regarding the LGBTQ+ community and the pride parade in Jerusalem, where Tag Meir takes part by commemorating Shira Banki, the 16-year-old teenager (see also section 2.3):

By progressing a mutual discourse of humanity and solidarity between Jews and all other minorities, we are trying to act as a connecting branch between religion and liberalism as an answer to religious fanaticism directed against Palestinians and other minorities (such as LGBTQ+) (Eli, 2022).

The organisation is not dependent on the government; it relies on finance from other NGOs and donations to act freely. For example, when Tag Meir wanted to send flowers to a family in Jenin on the West Bank, they did it through a proxy – a Palestinian that the NGO has been assisting for a few years. It also has an education department, trusted with creating training programs for religious teachers, funded partially by the EU, under the cooperation and authorisation of the anti-racism department in the Ministry of Education. Those training programs involve meetings between Jewish and Arab teachers. They get emotionally complicated, especially in times of security distress (e.g., fear of arriving in Palestinian villages and Jewish ones).

Tag Meir is based on volunteer work, uses digital tools, and is active on Facebook and other social media channels. It receives some of its data regarding racist attacks from online sources such as the public and its volunteers. It also uses digital tools to connect the volunteers spread across Israel, where each one is a member of a WhatsApp group (Tag Meir south, Tag Meir north, etc.). It assists in reporting events and gets an instant recruit of volunteers to protest, give away flowers, and support victims by verifying and sharing their stories. Its Facebook page has more than 37K members, which has increased over the past couple of years. The NGO also has a newsletter sent to registered members. Social media is a channel to receive reports from the media and individuals and spread information. Tag Meir follows online publications manually:

We are using Google searches, they attend radical websites using specific keywords, such as ones seen on graffiti of price-tag actions, and other quotes they find in the offline space (e.g., Revenge; long live king Mashiach, etc.). Another significant issue is the relation to Mount Temple and the potential of destruction around religious fanaticism; **the tension there is high and motivated by online/offline gatherings** by fanatics led by their rabbis and political electors (Eli, 2022).



In this case, AI can assist in locating reports regarding price-tag actions and casualties and automatically gathering data from social media. Another is the aspect of recognition in the activity, which is necessary to expand its agenda across the web and attract more volunteers to join acts of compensation.

## 5. Conclusions

After examining the activities of state and civic actors and their affiliations with technology in general, we came up with a few insights. We have identified two central pathways for deradicalisation: the preventive and integrative channels. These present some differences and similarities in their work mechanisms. The preventive track, which focuses on detecting radicalisation, is based on reports that the public submits to the relevant institution, which is slowly becoming a central practice. It can be captured as a prominent theme within the online/offline radicalisation framework. Additional digital mechanisms are required since, even where they can detect radicalisation, it does not seem yet that there are suitable tools to handle information on a large scale. This involves a set of practices and particular means to extract ad-hoc threats and recognise patterns.

Regarding the GO that participated, the lack of human resources and budget shortage directly influenced its ability to analyse big data and design future responses. Another issue is state bureaucracy, which sometimes holds back the practical use of information from GO and NGO activities. On the other hand, NGOs can capture events sometimes faster than the security authorities since it contains much more data than the GO, using more advanced technology. Yet both mentioned the need for assistance in handling big data, a need that AI could solve. Here we see that developing a local I-GAP lexicon by the relevant institution in Hebrew can be the first step to implementing the detection of users in RoR.

The integrative channel shows that civic actors emphasise the connection with RoR and/or backlash populations as part of their daily routine. NGOs' knowledge can be valuable to the state while developing deradicalisation initiatives. Working with NGOs has shown that many civic initiatives involve educational activities, which can be used as an anchor to future work with young adults from different sectors. Here we find that a possible way of communicating and gathering data from specific online/offline communities is by spreading I-GAP surveys using "chatbots". Depending on the presence of users on a given social platform, a friendly-to-use chatbot (activated with the user's consent) in a Q&A format could enable access to information about tendencies during and after events linked to radicalism. It is a data source that can assist the institution in configuring necessary steps (i.e., professional help). AI technologies can produce analysed data reports for working with, and keeping track of, the occurrences within backlash populations.

What was familiar to all interviewees was the struggle with limited resources as opposed to the will to achieve technological assistance from AI. Applying AI and big-data instruments requires NGOs and GOs to combine manual and digital work, hence the fusion of existing knowledge with new tools. Therefore, they need to invest in particular operators focusing on online activities that will be trained to work with digital tools. It might be challenging for the

work of many institutions that rely on a limited number of information sources so far. Under the growing usage of social media as both a separating and socialising space, states should address this matter to increase deradicalisation initiatives as a response to extreme themes across the web and against the attempts of groups and individuals to establish it as normal behaviour.

Another essential point shared by all four institutions is the importance of publishing their deradicalisation activities. When lacking a vast presence within social media, the knowledge that an organisation possesses stays mainly within it. Assistance in publications can enlarge NGOs' groups of volunteers and improve public awareness of radicalisation processes. It can help reduce violent online discourse by empowering field experts involved in offline deradicalisation organisations. The institutions wished to receive more information from and about their communities, which big data can offer within the social media arena. In the future, we plan to continue working with the most relevant entities from this study, experimenting with AI solutions that will be developed for particular needs.

## 6. Appendices

### Appendix 1. Interview Guide

#### **D.RAD- WP6: AI tools and deradicalisation**

##### **SEMI-STRUCTURED INTERVIEW GUIDE**

Name	
Institution	
Type of institution (NGO or government)	
Position	
Professional responsibility	
D.Rad-related topics of interest	

Place and date of the interview	
---------------------------------	--

**Interview essentials**

- Before the official interview, you will have to perform at least 1-3 preparation meetings to **a)** find the person who is most relevant for applying technological-experimental research and **b)** Map out the main issues and existing practices that the institution follows, focusing on a future vision of technical assistance in the field of D.Rad activities.
- Consent to the interview can be by signing a consent form, and in case the interview is recorded, it can be by verbal consent. In any case, the institution’s approval of the discussion will have to take place before the interview starts, and if it is given verbally, it needs to be recorded.
- Please, introduce yourself and the main issues involving the D.Rad project; explain that the obtained data will be used to prepare reports and recommendations for EU partners, access to the data is limited to the researchers working on the project, and the names of the interviewees will remain anonymous if they choose. It is also necessary to explain the WP goals and the fact that we are examining aspects of technology but cannot offer any recommendations that will break the thin line between security and human rights, such as freedom of speech.
- Recording is not mandatory, but it is recommended to document notes for the future. You may start the recording and stop it anytime the interviewee or you wish to do so. After the introduction, announce to the participants that you are recording and remind them that participation in the study is voluntary, and the participant is free to refuse to answer a question and to terminate it at any point.

**Key questions**

Topic	The goal	question	Related issues
<b>Digital usage</b>	Check what kind of tools are available to the institution and how it considers the use of technology within the institution	<b>Tell us about the institution’s activity in the digital realm: how do you use technology in your everyday routine when it comes to D. Rad issues?</b>	<ul style="list-style-type: none"> <li>✓ Advantages and disadvantages of existing tools</li> <li>✓ Ways to spread out information digitally</li> <li>✓ Data collection and the institution’s use of it</li> <li>✓ Educational and future programs followed by received data</li> <li>✓ Digital surveys</li> </ul>

<b>Technology and data</b>	Find out how the institution handles categorisation and general use of data	<b>How do you receive the data you rely on, and for what purposes do you use it?</b>	<ul style="list-style-type: none"> <li>✓ Multiple channels of information</li> <li>✓ Self-collecting data from digital forms</li> <li>✓ Data from digital platforms</li> <li>✓ Ways of organisation</li> </ul>
<b>Technology and social media</b>	Try to examine how common the usage of social media channels is, and if the institution sees it as a practical and valid information source	<b>Are you involved in social media channels, and if so, what are their practices and benefits?</b>	<ul style="list-style-type: none"> <li>✓ Which ones?</li> <li>✓ How powerful is social media involvement within your field of expertise?</li> <li>✓ Is it considered the main channel of communication with the public?</li> <li>✓ What are the difficulties with the rise of social media involvement?</li> </ul>
<b>Digital ambitions</b>	Find out what kind of digital ambitions the organisation has, and which ones present an option for joint work	<b>How do you see the digital usage of the institution regarding your deradicalisation initiatives?</b>	<ul style="list-style-type: none"> <li>✓ Use of digital data from social media</li> <li>✓ Spread of information across multiple channels</li> <li>✓ Digital initiatives of D.Rad</li> </ul>
<b>New Digital tools</b>	Explore how AI tools can assist the institution (e.g., handling data, developing communication, adjusting digital mechanisms to collect data, etc. )	<b>In what way can AI or any other digital change help progress the institution's existing practices?</b>	<ul style="list-style-type: none"> <li>✓ Data organisation</li> <li>✓ Internal digital survey</li> <li>✓ I-GAP survey</li> <li>✓ Data detection and categorising followed by the I-GAP lexicon</li> <li>✓ Automatic computerised assistance in containing data</li> </ul>

## Appendix 2. Pre-interview Sheet

### WP6: Meeting summary

#### PRE-INTERVIEW

**Date:** \_\_\_\_\_

**Participants:** \_\_\_\_\_

**Meeting duration:** \_\_\_\_\_

**Radicalization category (WP3):** \_\_\_\_\_

**Goal:** \_\_\_\_\_

**Meeting no.:** \_\_\_\_\_

**Name and type of institution:** \_\_\_\_\_ (NGO/GO)

### **Critical issues**

1. NGO as a link to communities at RoR and/or backlash \_\_\_\_\_

\_\_\_\_\_

2. DeRadicalization initiatives \_\_\_\_\_

\_\_\_\_\_

3. Cooperation with government and municipal authorities \_\_\_\_\_

\_\_\_\_\_

4. Digital relevancy \_\_\_\_\_

\_\_\_\_\_

5. Future progress \_\_\_\_\_

\_\_\_\_\_

## **7. References and Sources**

Adalah, 2022. "Adalah and the Monitoring Committee for Arab Education Affairs demanded from the Ministry of Education to receive a report on the content of the meeting between the Director General of the Ministry and the ISA", The Legal Centre for Arab Minority rights in Israel. Available at: [link](#) (Last accessed 19 October 2022).

Ademkar, Y. 2021. "MK Avi Maoz: "If there is nothing wrong with people who are attracted to members of their own sex – what is wrong with paedophiles?". Walla, 2 August [Online], [link](#). (Last accessed 19 October 2022).

Ajeec-Nisped, 2022. "About us". Available at: [link](#). (Last accessed 19 October 2022).

Akir, Y. (Alias), 2022. Interview with the author, Beer Sheva, 7 June. (via zoom).

Am Israel Chai, 2022. "Death to terrorist T-shirts". Available at [link](#). (Last accessed 19 October 2022).

ARCGU, 2019. "Government Decision no. 1958", The Anti-Racism Coordinating Government Unit, Ministry of Justice, Available at [link](#). (Last accessed 19 October 2022).

Azuma, 2022. "Netivot pride parade". Available at: [link](#). (Last accessed 19 October 2022).

Ben Menachem, Y. 2022. "When the ball is in the hands of Hamas". The Jerusalem centre for Public and state Affairs (JCPA) [Online], [link](#). (Last accessed 19 October 2022).

Blumenthal, I., Steinmetz, M., and Berger, G. 2021. "After the attacks: The IDF detects an increase in the motivation of terrorist organisations to carry out attacks in the West Bank". Kan, 6 December [Online], [link](#). (Last accessed 19 October 2022).

Bohbot, A. 2022. "They were recruited through social networks from Gaza - and planned attacks: Hamas infrastructure in the Yash area was exposed". Walla, 19 September [Online], [link](#). (Last accessed 19 October 2022).

Chalabi, E. 2022. "Threats to the organisers of the Jerusalem Pride Parade: "The fate of Shira Banki awaits you". Ynet, 1 June [Online], [link](#). (Last accessed 19 October 2022).

Cohen G., Ben Kimon, E., and Chalabi, E. 2022. "The "weapon shirts" phenomenon is spreading: "This is an incitement to murder Jews". Ynet, 13 May [Online], [link](#). (Last accessed 19 October 2022).

Eichner, I. 2021. " Israel joined the GPAI forum for artificial intelligence". Ynet, 11 November [Online], [link](#) (Last accessed 19 November 2022).

Eli, C. (Alias), 2022. Interview with the author, Beer Sheva, 31 July (via zoom).

Eli, Y. 2022. "The fashion of terrorism: the shirts with the weapon print have become the symbol of the Palestinian struggle". Reshet13, 15 May [Online], [link](#). (Last accessed 19 October 2022).

Etinger, Y. 2022. "Don't be ashamed to say: gays go home": the rabbi and the attack on the LGBT community". Kan, 24 May [Online], [link](#). (Last accessed 19 October 2022).

Even, S. and Siman Tov, D. 2020. "Research in the intelligence community in the age of artificial intelligence". The Institute for National Security Studies (INSS) [Online], [link](#) (Last accessed 19 November 2022).

FakeReporter, 2022. "About us". Available at: [link](#). (Last accessed 19 October 2022).

Fareed, S. 2022. "Since the attack in Beer Sheva, about 5,800 inflammatory posts on social networks have been monitored", Makor Rishon, 13 April [Online], [link](#). (Last accessed 19 October 2022).

Haaretz editors, 2022. "Commissioner: In the next violent conflict, social networks should be blocked", Haaretz, 21 September [Online], [link](#). (Last accessed 19 October 2022).

Hakmon, A., and Bagano, Y. 2022. "The Speaker of the Knesset spoke for the first time at the Jerusalem Pride Parade: "I was shocked by the incitement". Maariv, 2 June [Online], [link](#). (Last accessed 19 October 2022).

Harel, A. 2022. "The security establishment's fear has come true: terrorists from Israel have risen as imitators in the West Bank". Haaretz, 20 March [Online], [link](#). (Last accessed 19 October 2022).

ITIC, 2022. "Palestinian Terrorism, 2021: Summary, Types and Trends", The Intelligence and Terrorism Information Centre, 2 February [Online]. Available at: <https://www.terrorism-info.org.il/en/palestinian-terrorism-2021-summary-types-and-trends/> (Last accessed 11 December 2022).

ITIC, 2021. "The events of the escalation - "Guardian of the Walls" – summary", The Intelligence and Terrorism Information Centre, 23 May [Online]. Available at [link](#). (Last accessed 19 October 2022).

IDF, 2022. " Hamas's attempt to take over your phone has been foiled again. All the details, characters and instructions". Israel Defence Forces report, 16 February [Online], [link](#). (Last accessed 19 October 2022).

INSS editors, 2021. "'The Guardian of the walls' operation". The Institute for National Security Studies [Online], [link](#). (Last accessed 19 October 2022).

ISA, 2022. "ISA exposes the recruitment of minors by Hamas in the Gaza strip to commit terror in the west bank". Israel Security Association [Online], [link](#). (Last accessed 19 October 2022).

Jewish Pluralism Watch, 2019. "The Noam party in a video comparing the actions of the Nazis to the LGBT organisations". Available at [link](#). (Last accessed 19 October 2022).

Kan editors, 2022. "The terrorist in Beer Sheva: a resident of the Bedouin diaspora, with a security background and active in a terrorist organisation". Kan, 22 March [Online], [link](#). (Last accessed 19 October 2022).

Knesset, 2022. "Teachers who support terrorism part 1", The Committee of Culture, Education and Sports news, 6 April [Online], [link](#). (Last accessed 19 October 2022).

Knesset research and data centre. 2018. "Information about Artificial Intelligence". Available at: [link](#) (Last accessed 19 November 2022).

Kobowitz, Y. 2022. "ISA: Seven Hamas operatives were arrested in the West Bank who planned attacks directed by the organisation in Gaza". Haaretz, 19 September [Online], [link](#). (Last accessed 19 October 2022).

Levy, Y. 2022. "Violence against the organisers of the pride parade in Netivot: the windshield of his mother's car was smashed". Mynet, 10 May [Online], [link](#). (Last accessed 19 October 2022).

Maariv Editors, 2021. "A new report reveals: UN teachers publicly support the murder of Israelis", Maariv, 2 August [Online], [link](#) (Last accessed 19 October 2022).

Or, B. (Alias), 2022. Interview with the author, Beer Sheva, 31 August (via zoom).

Orpaz, I., and Siman Tov, D. 2021. "The never-ending campaign: the arena of social networks in "Guardian of the Walls". The Institute for National Security Studies (INSS), Available at: [link](#). (Last accessed 19 October 2022).

Peleg, G. 2022. "The lynching in Bat Yam: The Supreme Court ordered the release of the main accused", N12, 16 March [Online], [link](#). (Last accessed 19 October 2022).

Ravid, O., Veron, G., Nassar, F., and Odem, Y. 2022. "The ISA and the Israel Defense Forces arrested five suspects in Umm El Faham for involvement in the attack in Hadera", N12, 28 March [Online], [link](#). (Last accessed 19 October 2022).

Rot-Avneri, D. 2021. "Ben Gvir's bill: prohibiting teachers from supporting terrorism". Israel Hayom, 29 November [Online], [link](#). (Last accessed 19 October 2022).

Schweizer, Y., Lavie, E., and Elran, M. 2022. "What about ISIS and Israel and its Arab citizens?", The Institute for National Security Studies (INSS) [Online], [link](#). (Last accessed 19 October 2022).

Sela, U. 2022. "A teacher supported a terrorist on Facebook, the Minister of Education ordered to fire her", Walla, 30 March [Online], [link](#). (Last accessed 19 October 2022).

Shapira, N. 2022. "Terrorist support fashion": the new trend of Israeli Arabs – shirts with a print of an M16 rifle". N12, 10 May [Online], [link](#). (Last accessed 19 October 2022).

Srugim editors, 2022. "Radicalisation among Israeli Arabs - a ticking bomb", Srugim, 22 March [Online], [link](#). (Last accessed 19 October 2022).

State Comptroller, 2022. "Publication of a special audit report on mixed cities". Israel state Comptroller and Ombudsman [Online], [link](#). (Last accessed 19 October 2022).

Steinman, T. 2022. "After cancelling the pride parade in Netivot, the police also banned the support demonstration". N12, 19 May [Online], [link](#). (Last accessed 19 October 2022).

Tag Meir, 2022. "About us". Available at: [link](#). (Last accessed 19 October 2022).

The Hebrew Linguistic Academy. 2022. "Big Data- Definition". Available at: [link](#) (Last accessed 19 November 2022).

The Israel Academy for Science and Humanities. 2022. "Big Data". Available at: [link](#) (Last accessed 19 November 2022).

The National digital division. 2022. "Artificial intelligence and the data revolution in government". Available at: [link](#) (Last accessed 19 November 2022).

Trabelsi-Hadad, T. 2022. "A call to renew ISA tests for teachers who studied in the territories: "uproot weeds". Ynet, 24 March [Online], [link](#). (Last accessed 19 October 2022).

Uri, A. (Alias). 2022. Interview with the author, Beer Sheva, 11 July (via zoom).

Walla editors, 2022. "Minister Bar-Lev: "The commissioner's statement regarding social networks is unfortunate". Walla, 22 September [Online], [link](#). (Last accessed 19 October 2022).



Wang, P. 2019. "On defining artificial intelligence". *Journal of Artificial General Intelligence*, 10(2), pp. 1-37.

Yefet, N. 2022. "A pride event in Netivot was cancelled due to threats against the mother of one of the organisers". *Haaretz*, 18 May [Online], [link](#). (Last accessed 19 October 2022).