Government
Communication
Service

RESIST

Counter-disinformation toolkit

Foreword

Disinformation is not a new phenomenon: malicious rumours have always travelled faster than the truth. However a changing media environment means that disinformation can now spread faster than ever, to more people than ever. The rise of disinformation and the multiple threats this poses to our society means that we must respond urgently. And we must do this while continuing to embrace the incredible opportunities open to us to engage with the public in an online world.

We are at the forefront of a growing international consensus on the need to take action against disinformation, regardless of source or intent. Our vision is to strengthen the institutions of democracy and uphold our democratic values by ensuring the public and our media have the means to distinguish true news from disinformation. This starts with us, as government communicators. We hold the responsibility of delivering the truth, well told.

The systematic approach outlined in this toolkit is a crucial starting point. It is designed to help your organisations build resilience to the threat of disinformation step by step, while continuing to deliver effective communications to the public on the issues that matter most.

Alex Aiken

Executive Director of Government Communications

RESIST Disinformation: a toolkit

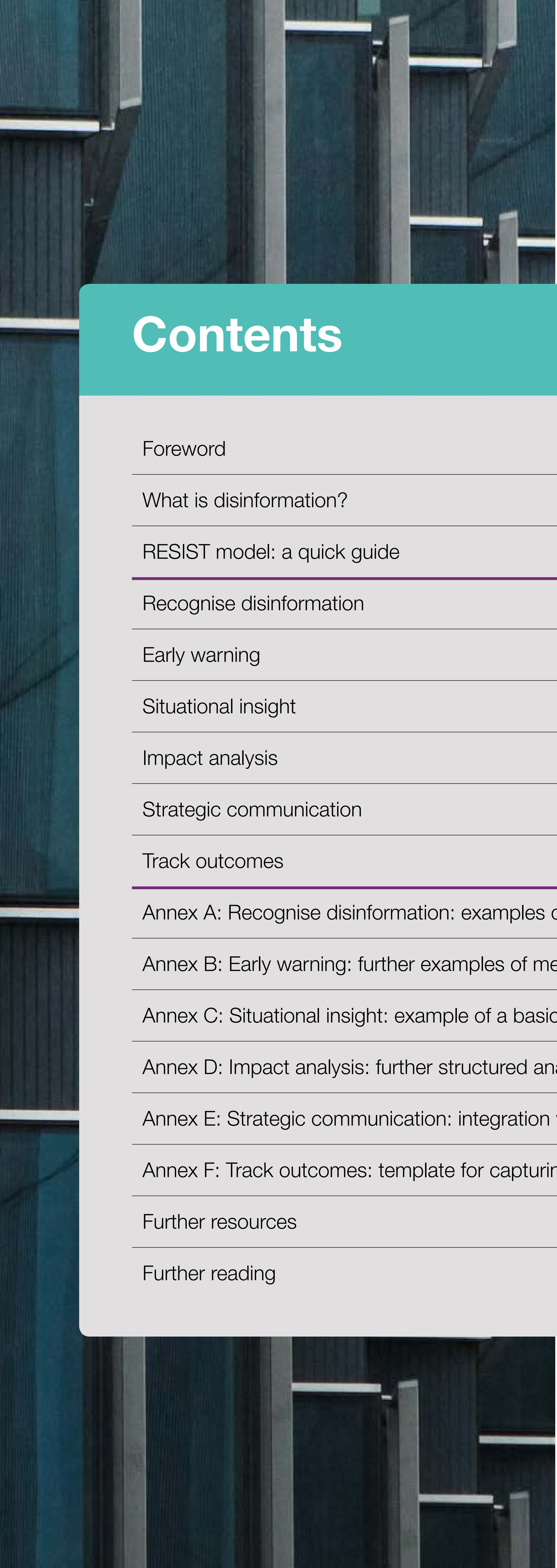
The purpose of this toolkit is to help you prevent the spread of disinformation. It will enable you to develop a response when disinformation affects your organisation’s ability to do its job, the people who depend on your services, or represents a threat to the general public.

What is disinformation?

Disinformation is the deliberate creation and/or sharing of false information with the intention to deceive and mislead audiences. The inadvertent sharing of false information is referred to as misinformation.

Who is this toolkit for?

Government and public sector communications professionals, as well as policy officers, senior managers and special advisers.



Contents

What is disinformation?

Disinformation is the deliberate creation and dissemination of false and/or manipulated information that is intended to deceive and mislead audiences, either for the purposes of causing harm, or for political, personal or financial gain.

When the information environment is deliberately confused this can:

- threaten public safety;
- fracture community cohesion;
- reduce trust in institutions and the media;
- undermine public acceptance of science's role in informing policy development and implementation;
- damage our economic prosperity and our global influence; and
- undermine the integrity of government, the constitution and our democratic processes.

Our aim is to reduce the impact of disinformation campaigns on UK society and our national interests, in line with democratic values. Our primary objective in countering disinformation is to **give the public confidence in information so they are equipped to make their own decisions**.

This toolkit provides a consistent and effective approach to identifying and tackling a range of different types of disinformation that government and public sector communicators may experience. **The RESIST disinformation model** is divided into components that can be used independently or tailored depending on the kind of organisation and the threats it faces.

Communications departments play a central role in recognising and responding to disinformation. You will often be the first to see it. This toolkit helps you develop routines to make informed assessments of risk and to share your insights with other parts of your organisation. It helps you to formulate recommendations and responses, and to evaluate your actions. The approach set out in this toolkit will contribute to a robust **early warning system** for recognising and responding to threats and emerging trends in the information environment.

RESIST can also be used alongside the **FACT** model, which has been developed for quick application in everyday communications activity to tackle misinformation and disinformation online. This model helps media offices identify and act on threats identified on a daily basis.

The toolkit will help you to:

recognise disinformation

use media monitoring for **early warning**

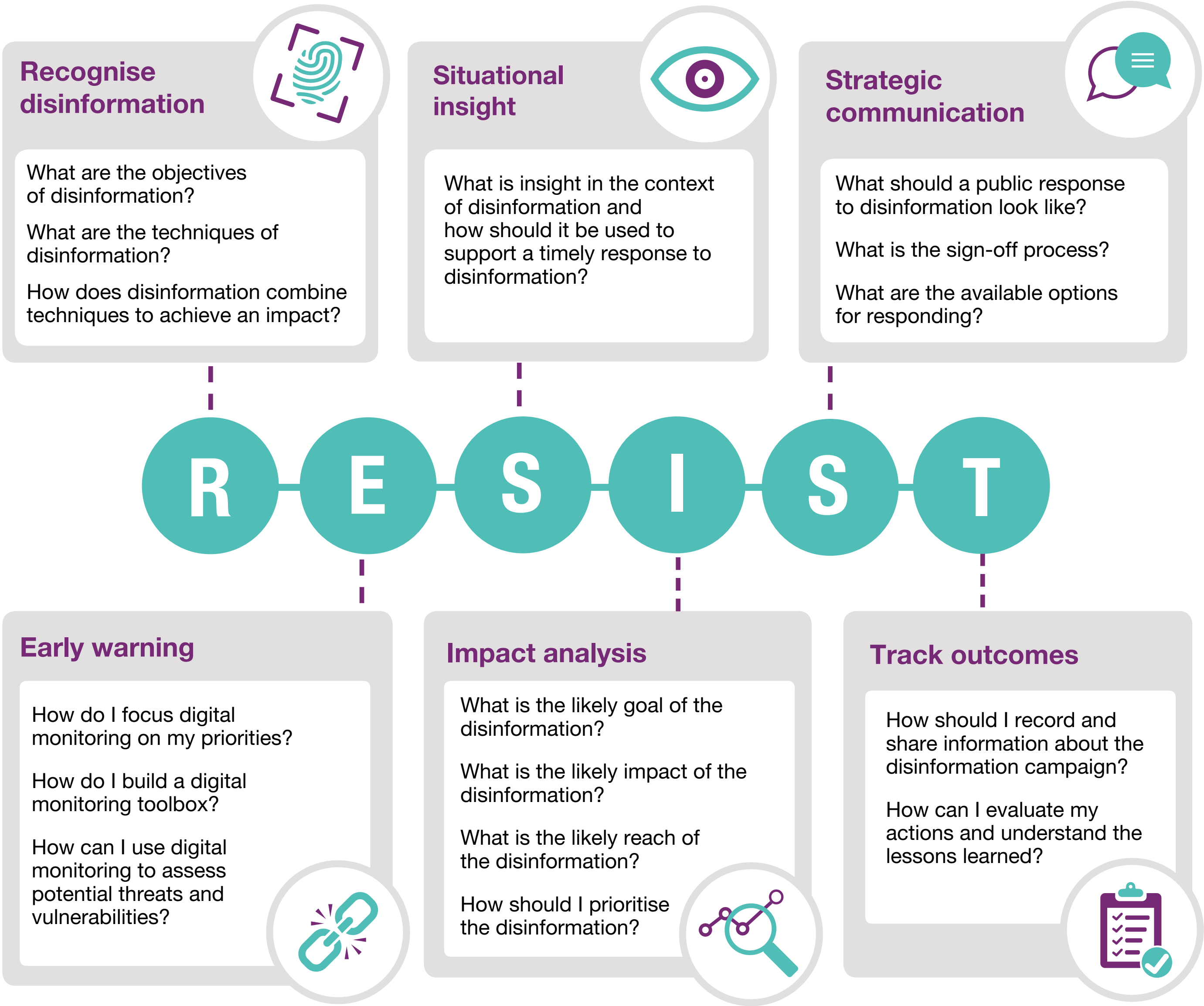
develop **situational insight**

carry out **impact analysis** to better understand the goals, impact and reach of disinformation

deliver **strategic communication** to counter disinformation

track outcomes

RESIST model: a quick guide



R E S I S T

Recognise disinformation

What are the objectives of disinformation?

What are the techniques of disinformation?

How does disinformation combine techniques to achieve an impact?

Early warning

How do I focus digital monitoring on my priorities?

How do I build a digital monitoring toolbox?

How can I use digital monitoring to assess potential threats and vulnerabilities?

Situational insight

What is insight in the context of disinformation and how should it be used to support a timely response to disinformation?

Impact analysis

What is the likely goal of the disinformation?

What is the likely impact of the disinformation?

What is the likely reach of the disinformation?

How should I prioritise the disinformation?

Strategic communication

What should a public response to disinformation look like?

What is the sign-off process?

What are the available options for responding?

Track outcomes

How should I record and share information about the disinformation campaign?

How can I evaluate my actions and understand the lessons learned?

Recognise disinformation

Disinformation is about **influence**. The people who spread it do not want members of the public to make informed, reasonable choices. They try to achieve a goal by deliberately shortcutting normal decision-making processes. The basic techniques are simple – we call them the **FIRST principles of disinformation:**

- **Fabrication** manipulates content: for example, a forged document or Photoshopped image;
- **Identity** disguises or falsely ascribes a source: for example, a fake social media account or an imposter;
- **Rhetoric** makes use of malign or false arguments: for example, trolls agitating commenters on a chat forum;
- **Symbolism** exploits events for their communicative value: for example terrorist attacks; and
- **Technology** exploits a technological advantage: for example bots automatically amplifying messages.

These FIRST principles of disinformation are often **combined to create an impact.**

FIRST principles, combined

1. Look for a social issue that is sensitive or holds symbolic value.
2. Create two or more social media accounts under false identities.
3. Manipulate content to provoke a response within the sensitive issue.
4. Release the content through one account, then criticise it through others.
5. Use bots to amplify the manipulated content to opposing networks.
6. Use memes and trolling to give the impression of a heated public debate.

Potential impact:
Undermine confidence in government or between social groups; contribute to political polarisation; earn money through clicks; go viral and reach mainstream news.

Early warning

You will need to do some preparatory work to better understand exactly **what** you want to monitor. The answers to the below questions will help you to **focus** your digital monitoring on the issues that matter most for disinformation. This step can be used in different stages and kinds of planning.

	Priorities	Attitudes
Policy objectives	What are my **priority policy areas** and objectives?	What are the prevailing attitudes in these areas that could be harnessed for disinformation?
Influencers	Who are the **key influencers** affecting my policy areas?	What are their prevailing attitudes toward my organisation or our objectives that could be harnessed for disinformation?
Audiences	Who are my **key audiences?**	What are their prevailing attitudes toward my organisation or our objectives that could be harnessed for disinformation?

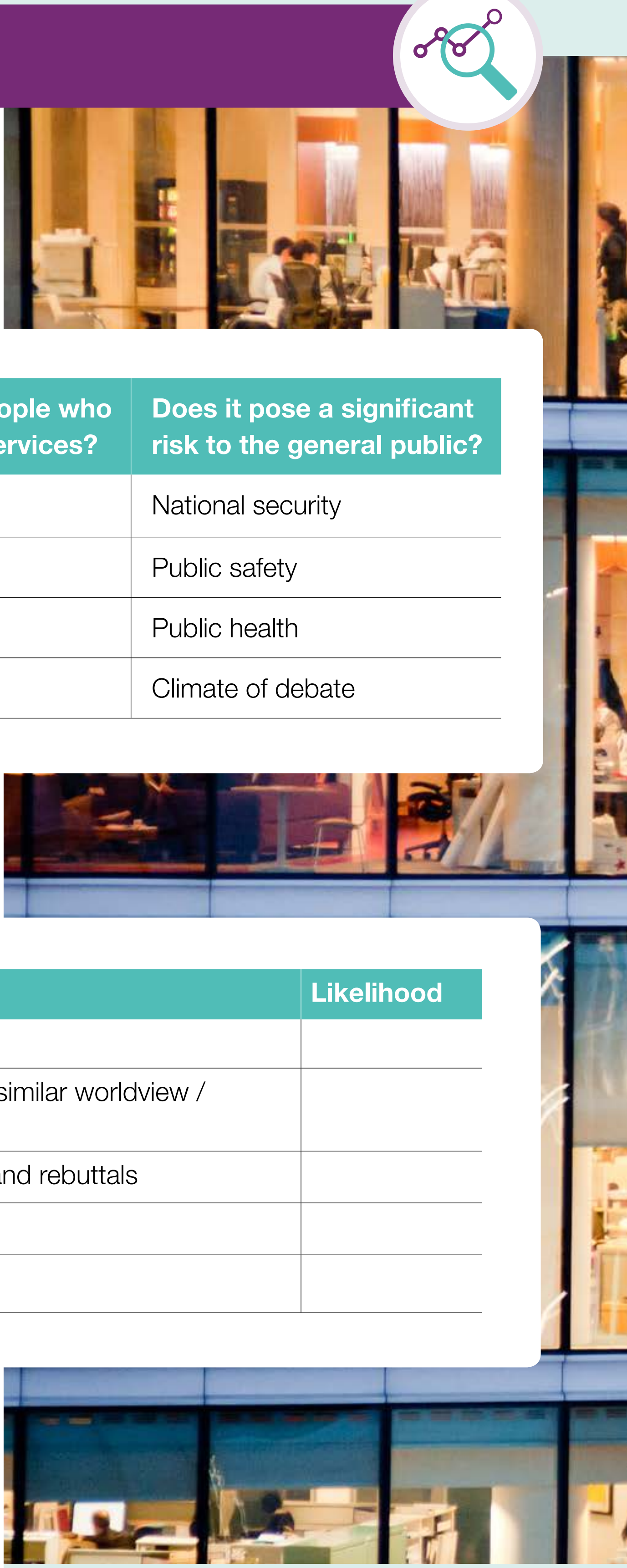
This work can help to guide your digital media monitoring so that you are prepared to identify any indicators of potential threats at the earliest possible stage.

Situational insight

Monitoring becomes valuable when it is turned into **insight**. Insight is a form of analysis that turns **interesting data** into **actionable data**. It answers the question, 'so what?' At its core, insight is about understanding audiences to support communication planning. A disinformation insight product should at a minimum include:

- key insights and takeouts: a top line summary including a short commentary explaining the 'so what' and setting out your recommendations for action; and
- sections on key themes and issues covering:
  - relevant outputs from your department on priority issues, for example a ministerial announcement;
  - examples of disinformation relating to these outputs, including where and how it is circulating;
  - key interactions and engagements;
  - trends and changes in attitudes over time (this can be combined with any polling data you have); and
  - your commentary and recommendations for a response.

Impact analysis



If you have identified some disinformation that relates to your organisation, you should make an assessment of its goals, impact and reach. This is achieved by answering a number of questions which can guide you in deciding whether to respond. For example, you should ask:

Does it affect the ability of your organisation to do its job?	Does it affect the people who depend upon your services?	Does it pose a significant risk to the general public?
Ability to deliver services	Key stakeholders	National security
Reputation	Key audiences	Public safety
Policy areas/goals	Niche audiences	Public health
Individual staff/staff safety	Vulnerable audiences	Climate of debate

You should make an assessment of how extensively you believe the disinformation will be engaged with. Is it likely to disappear within a few hours or does it have the potential to become tomorrow's headlines?

Exposure/reach	Likelihood
Little interest: very limited circulation and engagement	
Filter bubble: some engagement within niche audiences with similar worldview / automated circulation	
Trending: some discussion online, may include open debate and rebuttals	
Minor story: some reporting on mainstream media	
Headline story: affecting day-to-day operations	

Once the previous steps are completed, you should be able to assign a priority level to the disinformation. Is the disinformation likely to become part of a major cross-governmental crisis, like the Skripal poisoning? Or is it enough simply to monitor developments? The principle is that the goal, impact and reach should inform how urgently you prioritise the case.

	Description	Actions	Audiences	Tools
High	The disinformation has the potential to affect national security and has a high likelihood of making headlines. It requires immediate attention and escalation.	Make senior staff, SpAds / policy advisers and other parts of government aware of the issue and its priority. Share insight and analysis. Prepare quickly for a cross-Whitehall response.	- Senior staff - Wider government	- Share insight - Briefings - Prioritise short-term communications
Medium	The disinformation has the potential to negatively affect a policy area, departmental reputation or a large stakeholder group and is trending online. It requires a response.	Make senior staff and SpAds / policy advisers aware of the issue. Share insight and analysis within department. Investigate the issue and prepare press lines based on known facts.	- Senior staff - Policy advisers	- Insight - Briefings - Press lines - Prioritise short and medium-term communications
Low	The disinformation has the potential to affect the climate of debate and has limited circulation. The debate should be routinely followed but intervention is unnecessary/ undesirable.	Share insight and analysis within media department. Investigate the issue and prepare press lines/narratives based on known facts. Conduct a baseline analysis of debate and track any changes.	- Comms officers	- Insight - Press lines - Baseline analysis - Prioritise medium and long-term communications

Strategic communication

You can now consider a range of communicative approaches, such as short-term/reactive options, medium-term/proactive options, and long-term/strategic options.

You can combine them into a tailored communication strategy aligned with the OASIS communications planning model. For example, your response could include:

	Action	Target groups	Tools
Short-term reactive	The disinformation requires an immediate response. Use rapid communications to rebut, correct or counter disinformation in accordance with the established facts.	- Traditional media (journalists/editors) - Stakeholders and influencers - Social media platforms - Key audiences	- Press statement - Minister statement - Brief journalists - Q&A - Paid advertisement - Search engine optimisation (SEO) - Expose actors via friendly influencers

	Action	Target groups	Tools
Medium-term proactive	The disinformation requires a considered response. Use a combination of communications to assert own values/brands. Tie together proactive measures with your normal everyday communications and work with stakeholders/influencers to create consensus around your position.	- Traditional media (journalists/editors) - Stakeholders and influencers - Social media platforms - Wide audiences	- Campaign, narrative and brand development - Community outreach, dialogue and engagement - Facilitate network, stakeholders and influencers - Workshops/training

	Action	Target groups	Tools
Long-term strategic	The disinformation requires a coherent, sustained response to create long-term change. Develop and assert strategic narratives in relation to an issue by shaping the information space to promote your own position and deter others (raising the threshold).	- Traditional media (journalists/editors) - Young up-and-comers - Stakeholders and influencers - Social media platforms - Wide audiences	- Campaign, narrative and brand engagement - Programme funding e.g. for participatory content - Talent spotting and influencer support/ creation - Facilitate network, stakeholders and influencers - Workshops/training - Contingency planning

Track outcomes

You can evaluate your decision-making and actions based on the above steps, using a common format that enables you to share lessons learned.

Recognise disinformation: provide a bottom-line overview of the disinformation techniques used, including visual examples.

- What was the goal of the disinformation?
- What disinformation techniques were used?
- How were the disinformation techniques combined to achieve an impact?

Situational insight: once you have identified disinformation, consider how well your initial analysis and situational briefing supported your team's response.

- Were we able to offer an accurate and timely briefing to colleagues?
- Did we make any incorrect assumptions? On what basis?

Strategic communication: provide an overview of the communicative responses you took broken down into actions, target groups and tools.

Early warning: consider your preparatory work and the extent to which it supported your efforts to handle disinformation.

- Is your digital monitoring sufficiently focused on your priorities?

Impact analysis: consider your assessment of the likely goals, impact and reach of the disinformation.

- Was the disinformation prioritised correctly, based on goals, impact and reach?

Track outcomes: collect this information in a dossier together with your assessments of the actions you took.

- What was the impact of your efforts to handle the disinformation?
- What lessons can be learned from this case?

Recognise disinformation

This section will help you to answer the following questions:

- What are the **objectives** of disinformation?
- What are the **techniques** of disinformation?
- How does disinformation combine techniques to **achieve an impact?**

These steps should be used to help you recognise disinformation if and when it appears, as a first step toward tackling it.

3.1 Objectives of disinformation

Disinformation is about influence. People try to influence one another all the time. For example, the advertising and public relations industries try to influence our behaviour in small ways hundreds of times a day. Disinformation tries to influence us by using falsehoods to achieve an outcome.

The people who spread it do not want us to make informed, reasonable choices. They try to achieve a goal by deliberately shortcutting normal decision-making processes. They lie to make us think or act a certain way. The reasons for this are many, and have differing degrees of severity. Below are five common types.

1. **Economic:** the goal of disinformation is monetary gain. For example, in the case of clickbait, the goal is to obtain a 'click'. This is achieved by providing a headline, multimedia or other signalling that falsely entices you to visit a webpage. A fraction of a penny in advertising revenue can turn into thousands of pounds if a story goes viral. In such cases, the objective is purely economic. However, such websites can be linked to malware or other forms of tracking with ultimately criminal objectives, and can have secondary political consequences related to the content of the articles, where existing fault lines in debates are exploited.

Example: a group of entrepreneurs create controversial stories with clickbait headlines during an election. The headlines do not lead to genuine stories but rather to advertising pages that attempt to automatically install malware.

2. **Because I can:** the goal of disinformation is to achieve something difficult or audacious. This is supported by a 'hacker' or 'gamer' mentality, assuming the view that systems are there to be 'gamed' or technologically exploited. The objective is primarily about the scale of the challenge, personal gain, and earning respect for ability. Secondary consequences can include the hacking of crucial systems, the leaking of sensitive materials, the abuse of algorithms or other digital systems, and unethical use of user data to better target disinformation, for example 'dark' advertisements.

Example: a programmer is challenged by an online contact to manipulate the results of a Twitter poll for Premier League team of the season.

3. **To discredit:** the goal of disinformation is to negatively affect credibility, trust and reputations. This is achieved by targeting an individual or organisation and using falsehoods to undermine them. The explicit target of the attack may not necessarily be the main objective of the disinformation. For example, the objective may be to isolate vulnerable audiences dependent on the services of a certain organisation by discrediting the organisation. Discrediting is one of the most prevalent intentions of disinformation and fits with other intentions such as polarisation and information influence operations.

Example: an actor forges documents which discredit the BBC's leadership. The digital debate is seeded with the narrative that the BBC cannot be trusted, pushing audiences toward alt-left or alt-right news sites.

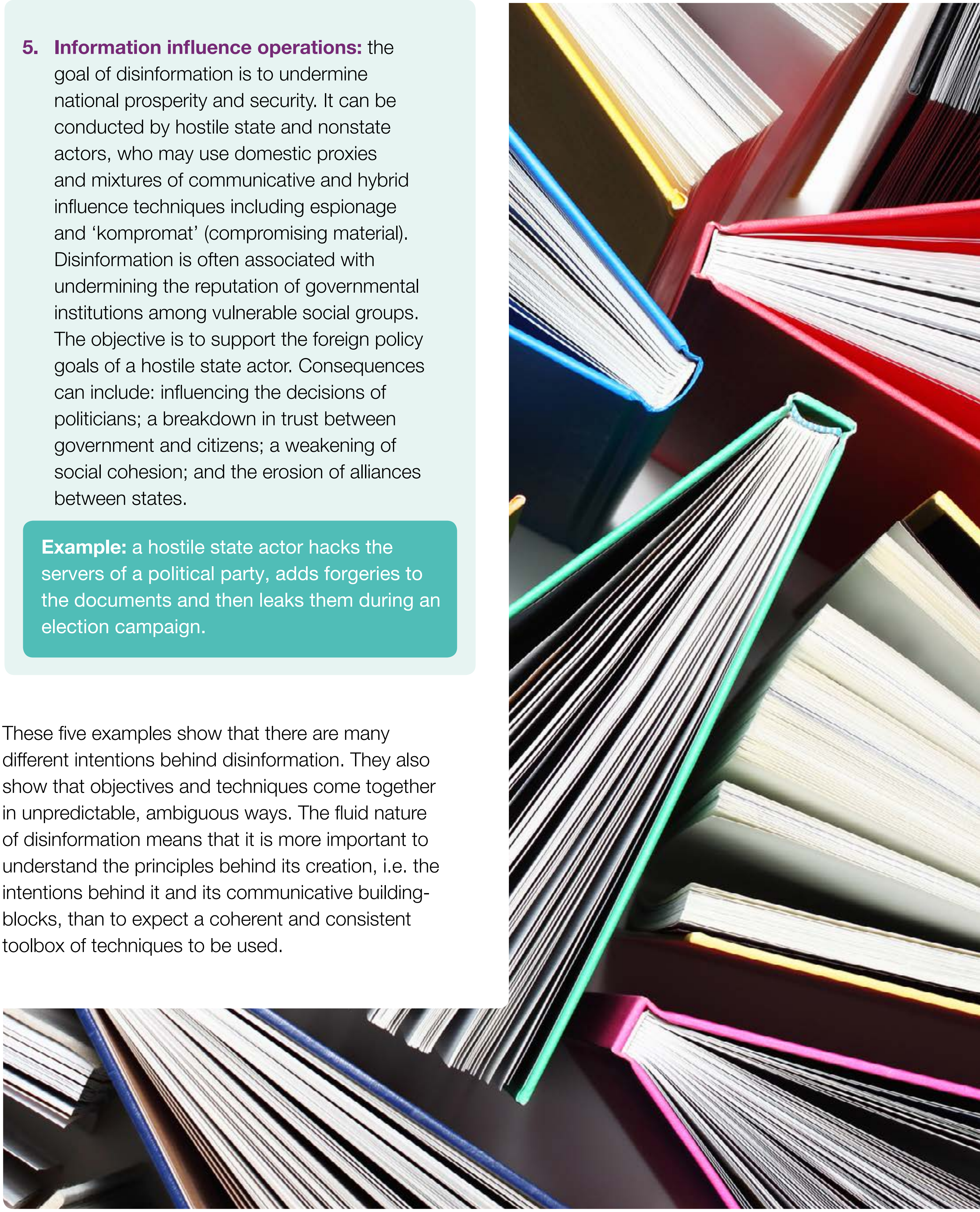
4. **Polarisation:** the goal of disinformation is to contribute to existing tensions by aggravating them. This is achieved by exploiting an existing debate by seeding it with spurious content designed to provoke a response from either side, thereby eroding the middle ground. The objective is usually political or social. Consequences include: damage to reputations or credibility; frequent arguments instead of constructive dialogue, for example from online trolls; an increased polarisation of political debate, for example from stoking of sensitive questions such as migration; provocations with implications for public health, for example in the anti-vaccine movement; through to incitements to violence.

Example: a group flood the comment section of a news story about deaths from flu with falsehoods claiming that the flu vaccine caused the deaths, while simultaneously writing comments that attack the parents of children who don't get vaccines. The common ground for constructive debate is diminished.

5. **Information influence operations:** the goal of disinformation is to undermine national prosperity and security. It can be conducted by hostile state and nonstate actors, who may use domestic proxies and mixtures of communicative and hybrid influence techniques including espionage and 'kompromat' (compromising material). Disinformation is often associated with undermining the reputation of governmental institutions among vulnerable social groups. The objective is to support the foreign policy goals of a hostile state actor. Consequences can include: influencing the decisions of politicians; a breakdown in trust between government and citizens; a weakening of social cohesion; and the erosion of alliances between states.

Example: a hostile state actor hacks the servers of a political party, adds forgeries to the documents and then leaks them during an election campaign.

These five examples show that there are many different intentions behind disinformation. They also show that objectives and techniques come together in unpredictable, ambiguous ways. The fluid nature of disinformation means that it is more important to understand the principles behind its creation, i.e. the intentions behind it and its communicative building-blocks, than to expect a coherent and consistent toolbox of techniques to be used.



3.2 Techniques of disinformation

Disinformation is the combination of a malign intention or goal with a number of unethical communicative principles into a communication technique. The principles are simple. We call them the **'FIRST' principles of disinformation**:

- **Fabrication** manipulates content: for example, a forged document or manipulated image;
- **Identity** disguises or falsely ascribes a source: for example, a fake social media account or an imposter;
- **Rhetoric** makes use of malign or false arguments: for example, trolls agitating commenters on a chat forum;
- **Symbolism** exploits events for their communicative value: for example terror attacks; and
- **Technology** exploits a technological advantage: for example bots automatically amplifying messages.

These FIRST principles of disinformation are combined to create unethical communication techniques. The most common techniques include:

Technique	Example
ASTROTURFING (I) Falsely attributing a message or an organisation to an organic grassroots movement to create false credibility.	A source pays or plants information that appears to originate as a 'people's' movement.
BANDWAGON EFFECT (S) A cognitive effect where beliefs increase in strength because they are shared by others.	A person is more willing to share an article when seeing it is shared by many people – automated and rhetorical techniques can be used to give this impression.
BOT (I,T) Computer code that performs repetitive tasks along a set of algorithms.	Bots can be used to amplify disinformation or to skew online discussion by producing posts and comments on social media forums and other similar tasks.
FILTER BUBBLE (I, T) Algorithms which personalise and customise a user's experience on social media platforms might entrap the user in a bubble of his or her own making.	The social media flow of a user interested in Aston Villa FC gradually adapts to consumed content to eventually only show information in favour of Aston Villa.
FORGERY (F, I) Product or content is wholly or partly fabricated to falsely ascribe the identity of the source.	A false document with an official-looking Government heading is produced to embarrass or discredit the government.

Technique	Example
LEAKING (S, T) Disseminating unlawfully obtained information.	Stolen emails are leaked to compromise individual actors or to undermine public confidence.
MALIGN RHETORIC (R) Linguistic ruses aimed at undermining reasonable and legitimate debate and silencing opinions.	A combination of different rhetorical techniques are applied in online conversation to ridicule and diminish other opinions.
MANIPULATION (F) Alteration of content to change its meaning.	An image is cropped to only show some of the participating parties in an incident.
MISAPPROPRIATION (I) Falsely ascribing an argument or a position to another's name.	A public figure is incorrectly cited or falsely attributed as a source.
SATIRE AND PARODY (R, S) Ridiculing and humouring of individuals, narratives or opinions to undermine their legitimacy.	A public figure is ridiculed using memes where non-factual opinions are ascribed to the public figure.
SOCKPUPPETS (I, R, T) Use of digital technology to disguise identity, to play both sides of a debate.	A user creates two or more social media accounts under opposing identities, i.e. one pro-fox hunting, one against, with the aim of playing the identities against one another.
TROLLING (I, R, S) Deliberate commentating on internet forums to provoke and engage other users in argument.	Social media users deliberately post provocative comments to create emotional outrage in other users.

The techniques develop over time, particularly as new technologies emerge. The FIRST principles will help you see through the clutter and identify the underlying techniques. A more detailed list is in annex A.

3.3. Achieving an impact

Being aware of the goals of disinformation and its main techniques is an important step, but it is only part of the problem. These objectives and communication techniques are usually **combined** to achieve maximum communicative impact.

Disinformation techniques can be used to complement one another to create intricate ‘ruses’ or ‘plays’ that support specific objectives. Actors can utilise a variety of different disinformation techniques to construct complex operations.

Aggravated sockpuppet

1. Identify a social issue that is sensitive or holds symbolic value (symbols).
2. Create two or more social media accounts under false identities (sockpuppet).
3. Fabricate provocative content related to the issue (fabrication).
4. Release the content through one account, then criticise it through others (rhetoric).
5. Use bots to amplify the fabricated content to opposing networks (bots).

Potential Impact: Polarise debate, create confusion, undermine legitimate positions, sow discord.

Alternative narrative

1. Formulate a narrative which supports your objective.
2. Prepare disinformation to support your narrative e.g. false news stories, blogposts, ads (fabrication).
3. Publicise disinformation through own channels or alternative websites (filter bubble).
4. Engage controversial bloggers/opinionists to ‘verify’ and distribute the narrative (misappropriation).
5. Use trolling to attack users who argue against the narrative (rhetoric).

Potential Impact: Divert from real issues, undermine legitimate positions, crowd out legitimate narratives.

Tainted leak

1. Gain access to internal documents and emails from a target organisation through cyber-attacks (such as spear-phishing).
2. Prepare forged documents containing discrediting information resembling the obtained documents in style and form (fabrication).
3. Dilute the leak with forgeries.
4. Disseminate the ‘tainted leak’ over established channels (such as Wikileaks) to get the attention of legacy media.
5. Amplify negative reporting using bots and trolls (bots and rhetoric).

Potential Impact: Discredit and/or falsely incriminate individuals or institutions, undermine trust, create confusion.

Big data targeting

1. Conduct target audience analysis of highly engaged groups on social media to identify psychographic triggers related to key issues (symbolic).
2. Set up closed groups on social media designed to appeal specifically to the designated target audience (filter bubble).
3. Recruit members of the target audiences to the groups by mimicking a legitimate organisation or movement (astroturfing).
4. Distribute disinformation in the form of false news articles and memes to these groups (rhetoric and symbolism).
5. Encourage members of the target audience to action, such as contributing to the spread of disinformation or engaging in public demonstrations (agitation).

Potential Impact: Polarise debate, change of behaviour, undermine trust.

Manipulated quote

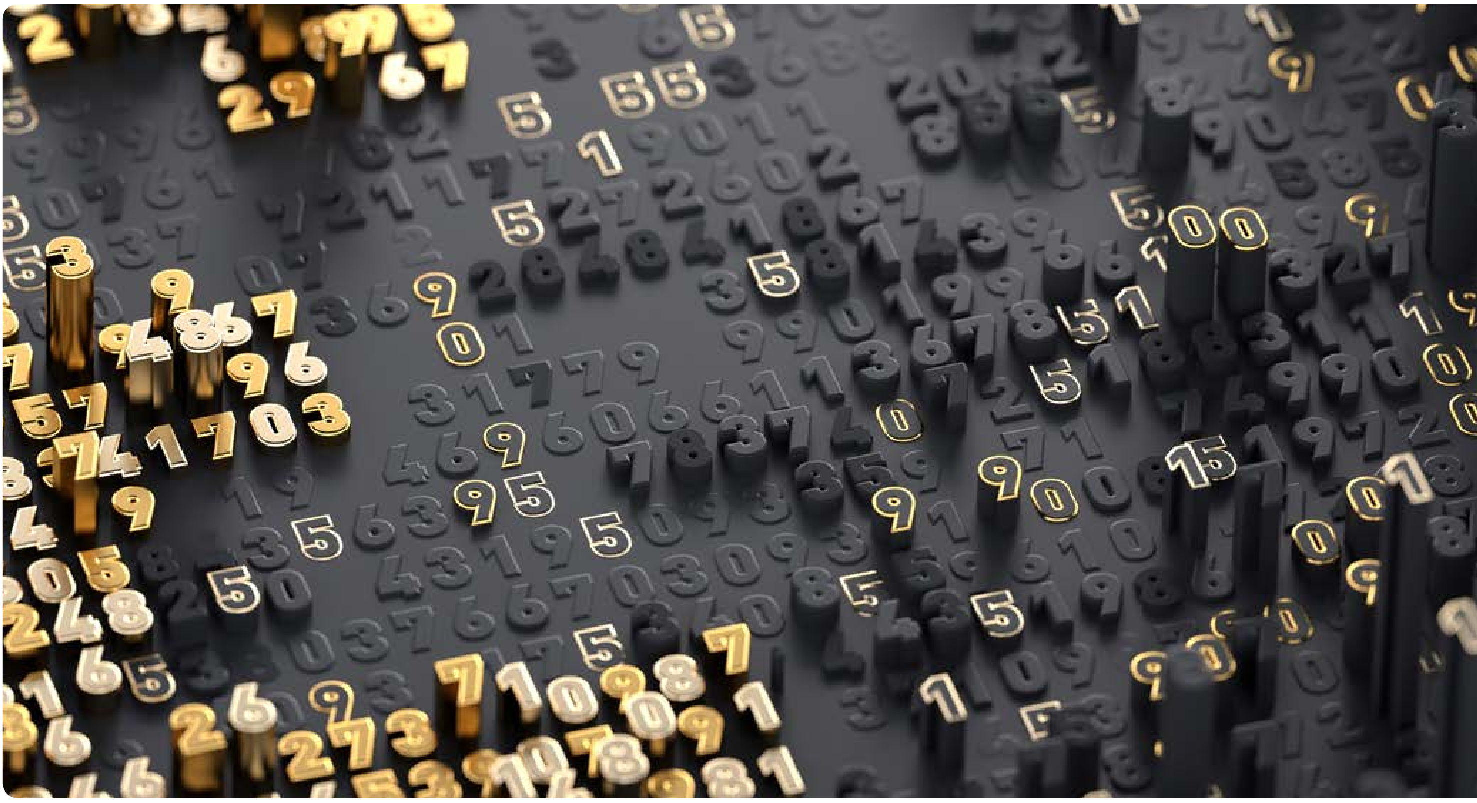
1. Take a quote from a public figure you wish to target.
2. Publish an article where the quote is taken out of its context to frame an issue so that it fits your preferred narrative (misappropriation).
3. Reference sources that mention the citation, across multiple news platforms and languages.
4. Use different actors and platforms to share your misappropriated article with minor changes to the text each time.
5. Refer to these intermediaries as sources for the falsified statement, which in the end has been ‘laundered’ to seem legitimate.

Potential Impact: Obscure truth, legitimise false claims, undermine trust.

Cheerleading

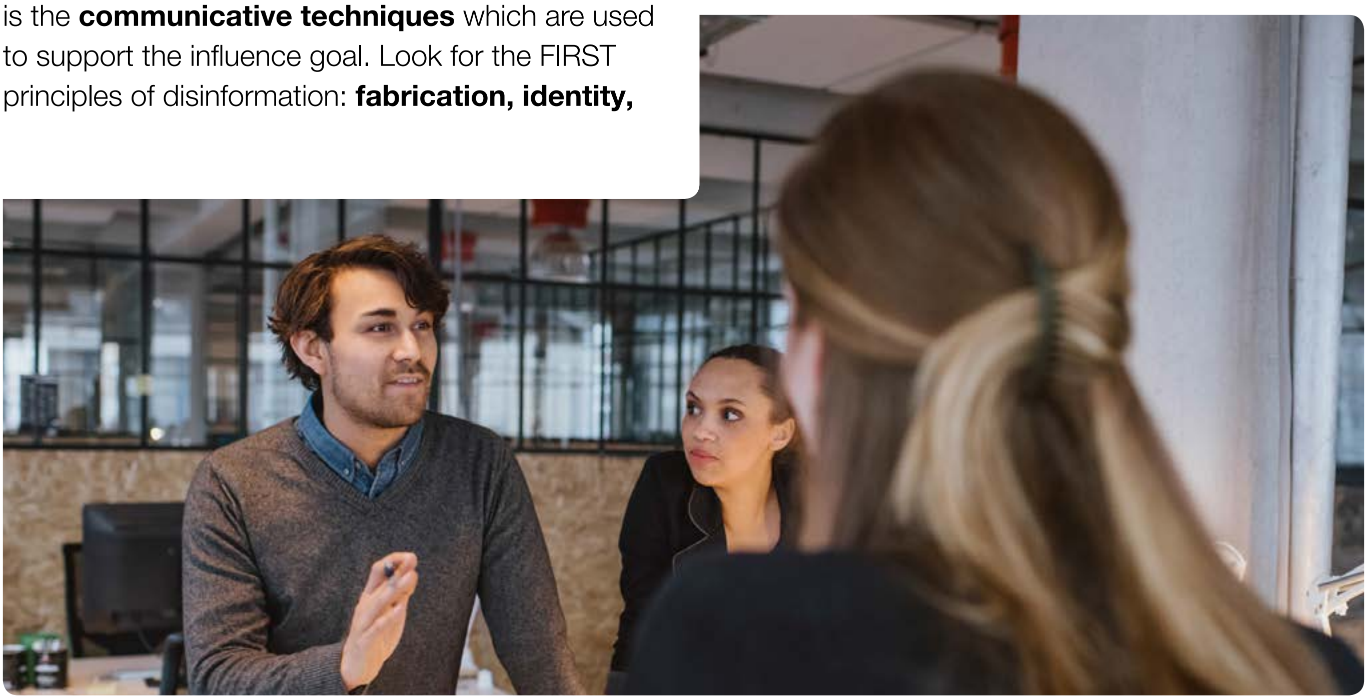
1. Identify dissenting opinions on a subject affecting your interests (rhetoric, symbolism).
2. Flood the information space with positive content (cheerleading) by using bots and trolls.
3. Ensure dissenting opinions are crowded out by positive comments and posts.
4. Create online groups which support your standpoint (filter bubble).
5. Maintain a large army of posters (bots and trolls) ready to get involved in any debate (rhetoric).

Potential Impact: Silence dissenting opinions, overload information space, shift narrative.



Summary

You should look for three factors when attempting to recognise disinformation. The **influence goal** is the first. What is the actor trying to achieve and why? Typical goals include monetary gain, personal respect, discrediting others, polarisation, and the influence operations of a hostile state actor. Second is the **communicative techniques** which are used to support the influence goal. Look for the FIRST principles of disinformation: **fabrication, identity, rhetoric, symbolism** and **technology**. There are dozens of advanced techniques based on these FIRST principles. Third, look for how the intention and the techniques are combined to **achieve an impact**. Together, it should be possible to recognise disinformation when it appears, as a first step toward tackling it.



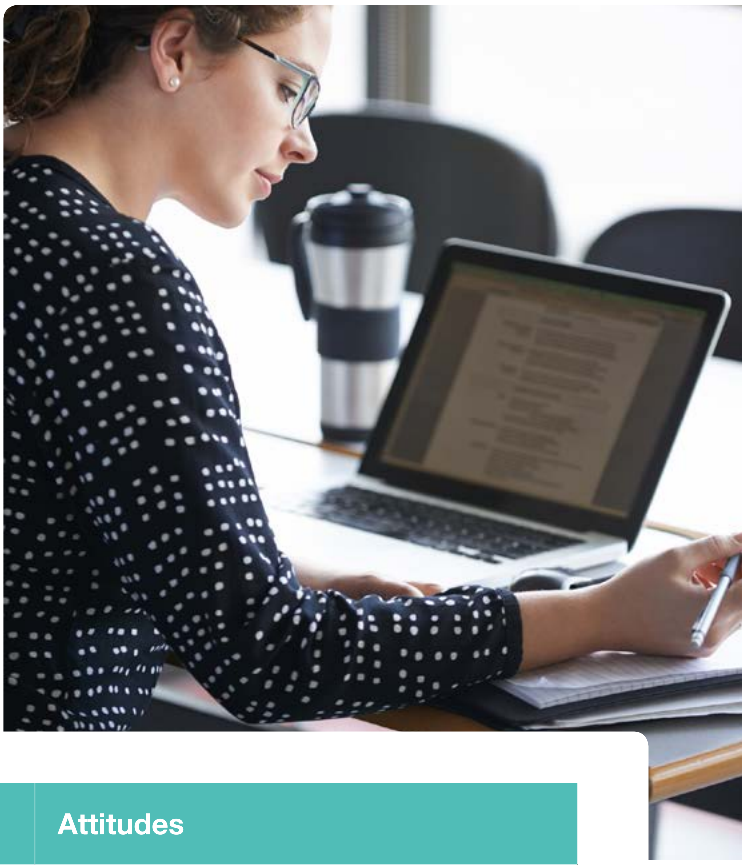
Early warning

This section will help you to answer the following questions:

- How do I focus digital monitoring on my **priorities?**
- How do I build a **digital monitoring toolbox?**
- How can I use digital monitoring to assess potential **threats and vulnerabilities?**

These three questions should be used for long-term monitoring of the information environment, for risk and contingency planning, and for short-term monitoring of trends. You should combine the steps outlined in this section with those in the following section to create **actionable insight** into disinformation trends.

4.1 Focus your monitoring on priorities



You will already conduct some kind of media monitoring, both of traditional and digital media. That means that you have a baseline knowledge of your key audiences, influencers and an understanding of the online debates that relate to your priority policy areas. However, you probably haven't focused this work specifically on disinformation. This section offers advice as to minimum and recommended standards for digital media monitoring and audience analysis for handling disinformation, based on examples of current best practice.

You will need to do some preparatory work to better understand exactly **what** you want to monitor. Ask yourself the following questions and place them into a grid:

	Priorities	Attitudes
Policy objectives	What are my **priority policy areas** and objectives?	What are the prevailing attitudes in these areas that could be harnessed for disinformation?
Influencers	Who are the **key influencers** affecting my policy areas?	What are the prevailing attitudes toward my organisation or our objectives that could be harnessed for disinformation?
Audiences	Who are my **key audiences?**	What are the prevailing attitudes toward my organisation or our objectives that could be harnessed for disinformation?

The answers to these questions will help you to **focus** your digital monitoring on the issues that matter most for disinformation. This step can be used in different stages and kinds of planning. For example, you could use it to assess your biggest issues for the year and/or for weekly or campaign-based planning.

4.2 Build a monitoring toolbox that suits your needs

Resource levels have a major effect on how much effort can be placed in digital monitoring. Thankfully, there is a great deal of support available. You should select from a variety of tools to form a **toolbox** or **dashboard** based on your needs. Tools include products created by specialist Government units, free tools and paid tools.

A number of useful monitoring, analysis and insight products already exist. You should identify which existing HM Government monitoring resources are available to you, and how helpful they can be for monitoring your priorities, influencers and audiences. You can find a list of contacts in the further resources section of this toolkit.

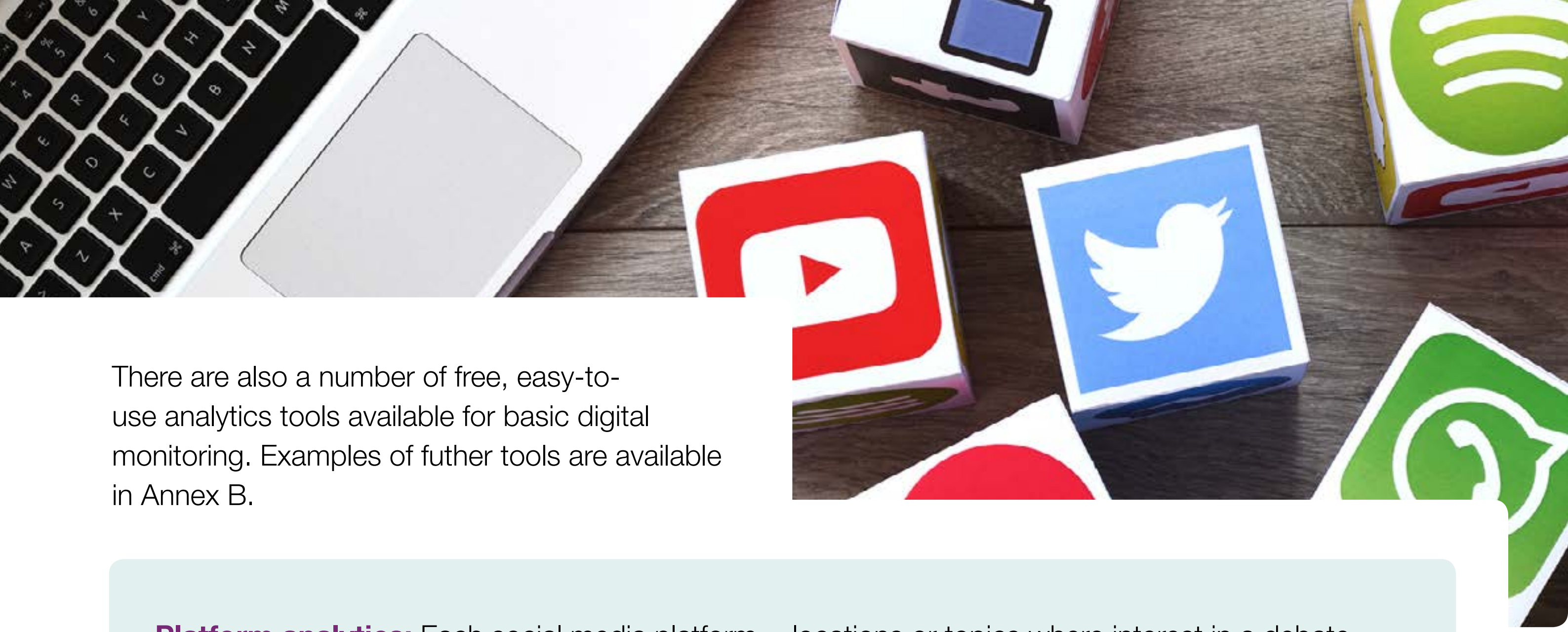
Media Monitoring Unit (MMU): produces daily social media briefings relating to specific topics and monitoring reports on traditional media (radio, TV, print) – based in No.10/Cabinet Office.

Rapid Response Unit (RRU): produces 3x daily email alerts on the top government stories and themes gaining traction online and monitors digital media in real time to respond at speed using the FACT model when mis/disinformation relating to HMG has been identified. Based in No10/Cabinet Office.

Research, Information and Communications Unit (RICU): produces analysis and insight on terrorist, extremist and organised criminal communications. Also home to the Disinformation Analysis Team (DAT), a cross-Whitehall unit responsible for building understanding of the domestic implications of disinformation (who is vulnerable to it, why and how it is impacting on UK society) through provision of specialist advice and insight – based in the Home Office.

Open Source Unit (OSU): provides open source monitoring and assessment of international social media and other open source material – based in the Foreign and Commonwealth Office.

Insight and Evaluation Basecamp: shares insight and evaluation tools and techniques and best practice from government and beyond – your organisation's insight team may also run regular polling.



There are also a number of free, easy-to-use analytics tools available for basic digital monitoring. Examples of futher tools are available in Annex B.

Platform analytics: Each social media platform has an analytics function that provides data on accounts or pages that you own. Platforms that you own pages on are an important source of insight for understanding how people engage with your content.

Google Trends: Shows how frequently terms are searched for on Google. The results can be broken down by time, country, and related queries to focus attention on a specific timeframe, location, and/or topic.

This is useful for revealing spikes in interest and can help guide your attention to specific days, locations or topics where interest in a debate has changed.

TweetDeck: Create a Twitter dashboard to follow multiple timelines, accounts and search terms in real time. Available at tweetdeck.twitter.com.

Browser extensions: There are a number of apps that can be added to your browser to speed up or even automate functions such as translation, image searches and taking screenshots. This is especially useful for speeding up simple tasks that you need to do often.

A number of paid-for social media insight tools are currently used by Government such as Newswhip, Crimson Hexagon and Brandwatch. These tools allow for the monitoring of complex keywords and phrases and automated outputs for factors such as most viewed/engaged posts, influencers, network maps and share of voice.

It should be underscored that none of these services provide a one-size-fits-all solution and teams should focus first on the skills needed to use these tools effectively and invest in the correct training when procuring them. Nor can components such as sentiment analysis be wholly relied upon. Users of any insights tool should also be aware of what data is available to be analysed and the limitations of that data.

Digital media monitoring allows you to form a **baseline** understanding of how your priority policy areas are represented on digital media, how debates are engaged with by key influencers, and how different audience groups are formed. They can help you to better understand where to look, and what to look for. The outcome of this kind of analysis should be a more focused understanding of:

- digital debates that are taking place in relation to your organisation and its work;
- the main attitudes held by key influencers and audiences;
- how influencers and segmented audiences engage on digital platforms with your organisation and its work; and
- changes in trends over time.

The value of this knowledge is that it enables you to improve your preparedness for handling disinformation. It can offer early warnings of potential threats and risks, and give a sense of what is normal and what might involve deliberate manipulation of debates according to **FIRST principles**. The next step is to develop contingency planning around the risks of disinformation.

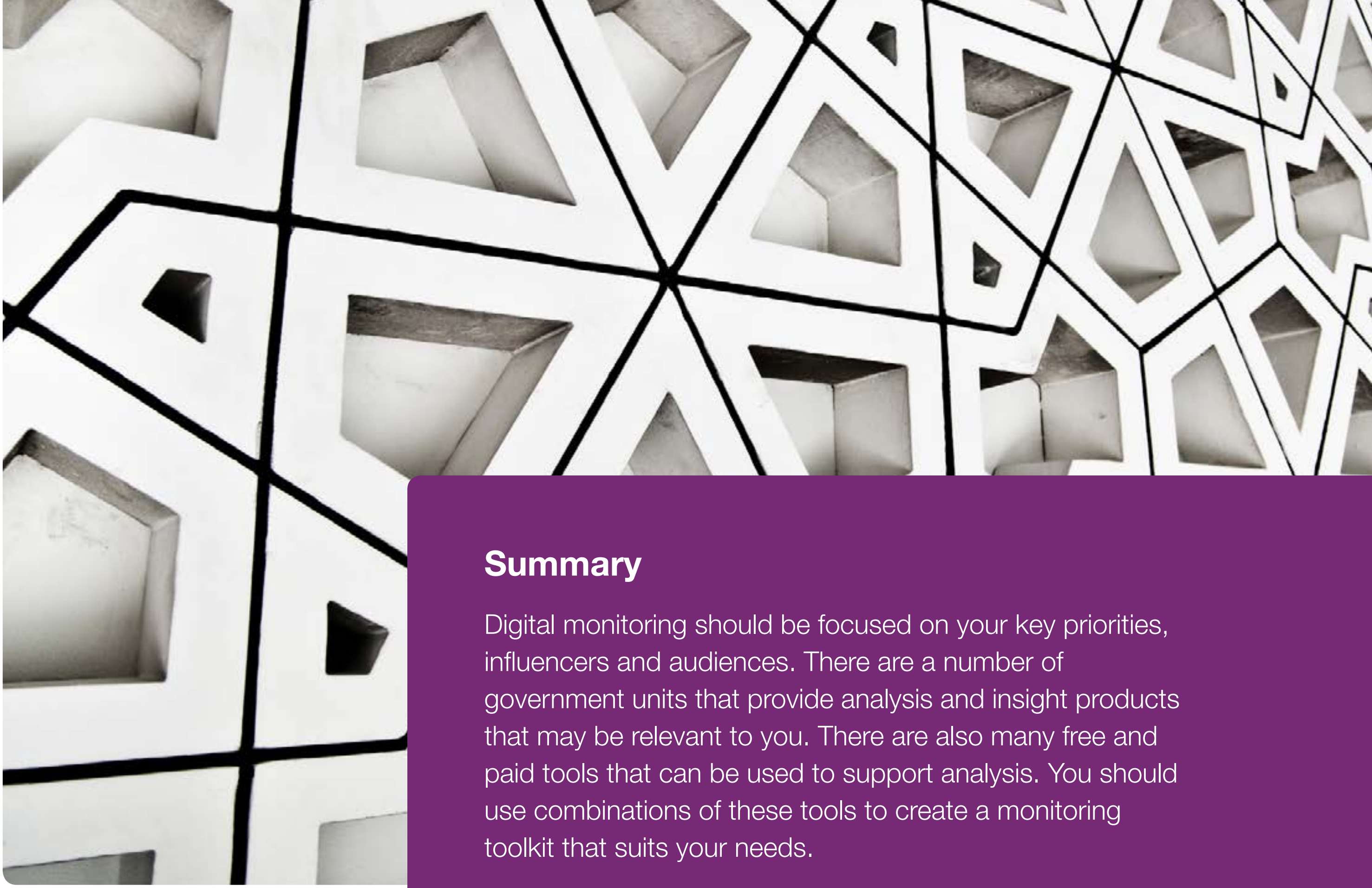
4.3 Work through the threats and vulnerabilities

The first two steps will help you produce a toolbox with focused information about your policy areas, influencers and audiences. While this is a useful exercise in its own right, the main emphasis for this work is on disinformation. You will recall from section 3.1 that disinformation can have a number of goals, such as economic, gaming of the system, discrediting, polarisation and influence operations. The results of the previous steps will help you establish a grid that captures the main public attitudes toward your priority policy areas. The next step is to examine the data more closely to look for indicators of:

- networks that appear to have an interest in sharing disinformation; and
- vulnerabilities in debates/issues/narratives that could potentially be exploited by disinformation.

Look for examples of communication that draw upon the FIRST principles of disinformation: fabrication, identity, rhetoric, symbolism and technology. Is there any indication that there are networks of influencers interested in using such techniques? Is there a market for disinformation among your audiences? Even if you don’t see any warning signs, it is worth considering the potential risks disinformation could seek to exploit, and worst-case scenarios for what disinformation could accomplish. This can be used both for long-term planning, and for weekly planning or campaign planning.

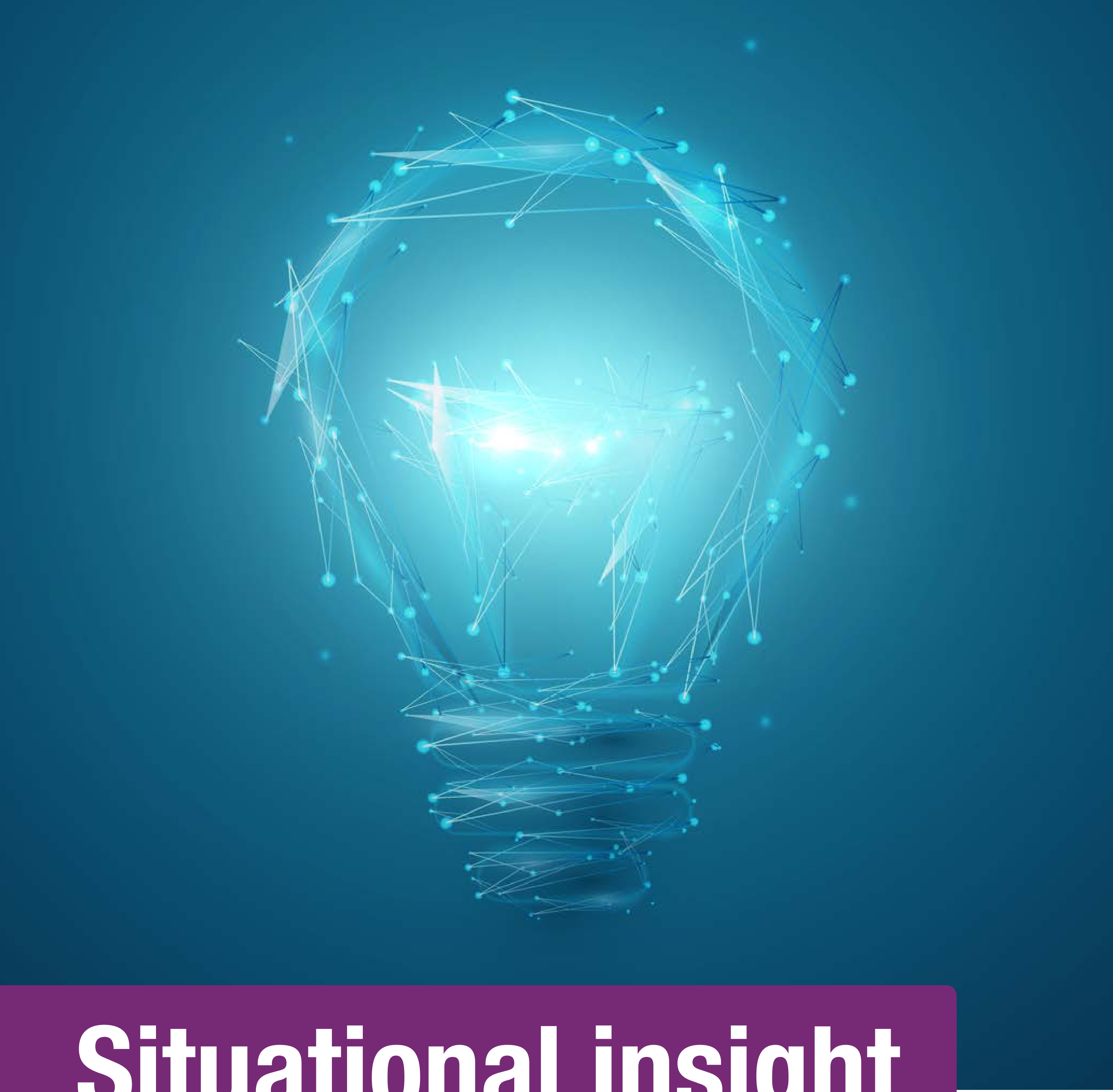
	Priorities	Attitudes	Source networks	Narrative risks	Worst case
Policy objectives	From 4.1	From 4.1	Which networks are threatening your policy goals with disinformation?	What aspects of your narrative(s) are vulnerable to disinformation?	What are some of the worst case scenarios / risks if disinformation spreads?
Influencers	From 4.1	From 4.1	Which influencer networks are spreading / engaging with disinformation?	What aspects of your narrative(s) are vulnerable to disinformation?	What are some of the worst case scenarios / risks if disinformation spreads?
Audiences	From 4.1	From 4.1	Which audiences are spreading / engaging with disinformation?	What aspects of your narrative(s) are vulnerable to disinformation?	What are some of the worst case scenarios / risks if disinformation spreads?



Summary

Digital monitoring should be focused on your key priorities, influencers and audiences. There are a number of government units that provide analysis and insight products that may be relevant to you. There are also many free and paid tools that can be used to support analysis. You should use combinations of these tools to create a monitoring toolkit that suits your needs.

The purpose of digital monitoring in relation to disinformation is ultimately to help you to **reduce vulnerabilities and plan for risk**. This kind of focused planning can help give you an early warning if disinformation appears within your priority policy areas or among key influencers and audiences.
The knowledge that you develop in these steps should be operationalised in the next step:
creation of **insight**.



Situational insight

This section will help you answer the following question:

- What is **insight** in the context of disinformation and how should it be used to support a timely response to disinformation?

By the end of this section, you will be familiar with the basic steps required to produce an insight briefing on disinformation for relevant people in your organisation.

5.1 Turning monitoring into insight

Monitoring becomes valuable when it is turned into **insight.** Insight is a form of analysis that turns **interesting data** into **actionable data**. It answers the question, ‘So what?’ At its core, insight is about understanding audiences to support communication planning. Insight should be used to:

- baseline/benchmark over time to show change;
- identify emerging trends and provide early warning of threats;
- understand how disinformation is distributed to key audiences;
- generate hypotheses and recommendations; and
- provide support for developing and targeting messages and campaigns, including pre-clearance of lines.

Insight usually takes the form of reports that are circulated daily, weekly or ad hoc depending on need. Much of the data can be drawn automatically from the monitoring toolbox or dashboard that you developed in the previous section. A good insight report can be as short as one or two pages: **put the most important information at the top and get to the ‘so what’ quickly**. Bear in mind that your insight product might be the first time that people in your organisation are exposed to digital monitoring data as a basis for analysing disinformation. It should be usable as a briefing for special advisers, policy advisers, senior staff and ministers, so explain things clearly by avoiding jargon and using images where helpful.

A disinformation insight product should at a minimum include:

- key insights and takeouts: a top line summary including a short commentary explaining the ‘so what’ and setting out your recommendations for action; and
- sections on key themes and issues covering:
  - relevant outputs from your department on priority issues, for example a ministerial announcement;
  - examples of disinformation relating to these outputs, including where and how it is circulating;
  - key interactions and engagements, for example is the disinformation being dealt with organically, is it being picked up by journalists and influencers and if so which ones?;
  - trends and changes in attitudes (and influencers and audiences) over time (this can be combined with any polling data you have); and
  - your commentary and recommendations for a response.

Your analysis and recommendations should provide as much clarity as possible on the following questions. Note that answers in previous steps will help you to fill in these fields.

From Attitudes and Narrative risks (section 4.3)

What are the main disinformation narratives and arguments?	Outline narrative(s), show examples
What is objectively false about them?	Clarify disinformation dimension

From Source networks (section 4.3)

Who is circulating the disinformation?	Visible sources
What networks/communities are they part of?	Friends, followers, influencers – use network analysis if necessary
Is the disinformation singular or part of a pattern?	Assess prevalence and scope
Who are the apparent targets of the narratives? **(primary and secondary audiences)**	Analyse hashtags, mentions and shares; consider audience types
What is the level of engagement with the disinformation?	Engagement
What are your initial recommendations?	Suggest actions

The goal of a disinformation insight product is to share the early warning signals you have gleaned from digital media monitoring with the people who need a situational briefing. As with all monitoring, it can be used in long-term planning, for example in an annual report or as part of a campaign evaluation, for ad hoc issues, or produced on a weekly or even daily basis.

Producing a disinformation insight product will support a proactive, accurate and speedy response. It will also help you to gain the internal approval you need to make a public response.

Summary

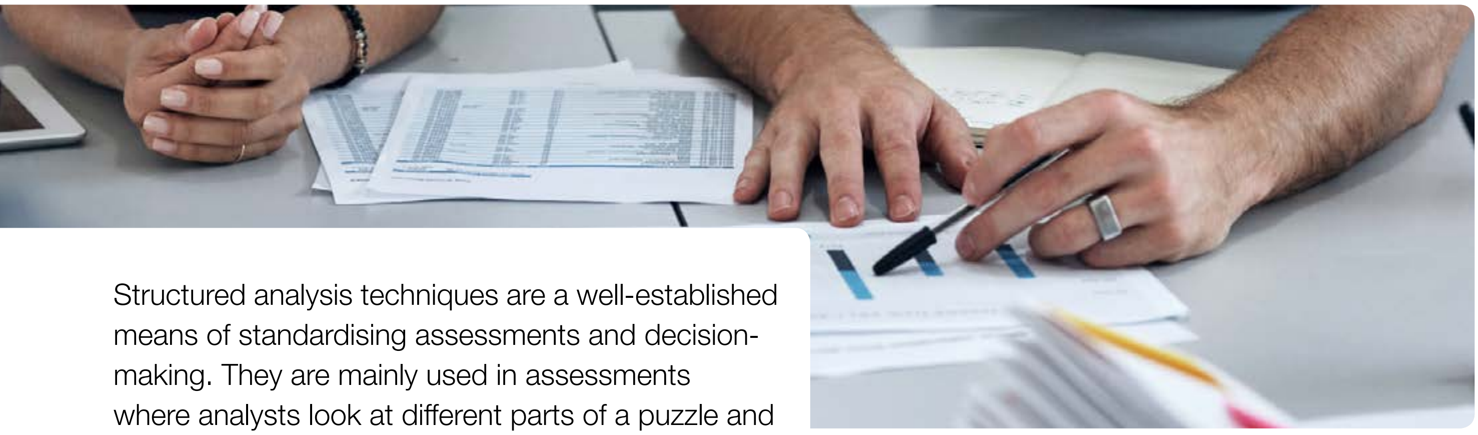
Situational insight is a means of gathering the data you have collected through digital media monitoring, providing a briefing to those who need to be involved in the issue. They should be short, clear and to the point. A basic example of an insight product is given in annex C.

Impact analysis

This section will help you answer the following questions:

- What is the likely **goal** of the disinformation?
- What is the likely **impact** of the disinformation?
- What is the likely **reach** of the disinformation?
- How should I **prioritise** the disinformation?

The following section provides a number of structured analysis techniques which cover a series of questions that can help to guide your assessment of the goals, impact and reach of potential disinformation you have identified through monitoring and insight. This can be used to help you decide whether to act and, if so, how.



Structured analysis techniques are a well-established means of standardising assessments and decision-making. They are mainly used in assessments where analysts look at different parts of a puzzle and need to share the same process and language. We draw upon simplified versions of these techniques here because handling disinformation should not be based on a gut feeling. You need to follow a structured, coherent process using a common language that leads to consistent decisions.

The structured analysis techniques are backed up by the 'uncertainty yardstick', which is used to provide a standard means of expressing risk. Further structured analysis techniques for more complex cases are given in annex D.

Qualitative term	Shortened version	Probability range
Highly unlikely	HU	Less than 10%
Unlikely	U	15-20%
Realistic probability	RP	25-50%
Likely	L	55-70%
Highly likely	HL	75-85%
Almost certain	AC	More than 90%

6.1 What is the likely goal of the disinformation?

Drawing upon the previous monitoring and insight, consider the following questions. Use the uncertainty yardstick to grade the likelihood of your hypotheses if necessary.

What is the intention of the disinformation?	From section 3.1
What disinformation techniques are being used?	From section 3.2
What is the observable effect?	Describe
Who benefits?	Actors, audiences
Who is disadvantaged?	Actors, audiences
Is action required, and if so what kinds of response?	From section 5.1

6.2 What is the likely impact of the disinformation?

Based on the above analysis, you should be able to make a reasoned assessment of the likely impact of the disinformation.

Does it affect the ability of your organisation to do its job?	Does it affect the people who depend upon your services?	Does it pose a significant risk to the general public?
Ability to deliver services	Key stakeholders	National security
Reputation	Key audiences	Public safety
Policy areas/goals	Niche audiences	Public health
Individual staff/staff safety	Vulnerable audiences	Climate of debate

6.3 What is the likely reach of the disinformation?

You should make an assessment of how extensively you believe the disinformation will be engaged with. Is it likely to disappear within a few hours or does it have the potential to become tomorrow's headlines?

Exposure/reach	Likelihood
Little interest: very limited circulation and engagement.	
Filter bubble: some engagement within niche audiences with similar worldview / automated circulation.	
Trending: some discussion online, may include open debate and rebuttals.	
Minor story: some reporting on mainstream media.	
Headline story: affecting day-to-day operations.	

6.4 How should I prioritise the disinformation?

Once the previous steps are completed, you should be able to assign a priority level to the disinformation. Is the disinformation likely to become part of a major cross-governmental crisis, like the Skripal poisoning, or is it enough simply to monitor developments?

Below are three example priorities: high, medium and low. You may need to develop your own criteria for prioritising disinformation based on your specific needs and experiences. The principle is that the goal, impact and reach should inform how urgently you prioritise the case.

	Description	Actions	Audiences	Tools
High	The disinformation has the potential to affect national security and has a high likelihood of making headlines. It requires immediate attention and escalation.	Make senior staff, SpAds/policy advisers and other parts of government aware of issue and its priority. Share insight and analysis. Prepare quickly for a cross-Whitehall response.	- Senior staff - Wider government	- Share insight - Briefings - Prioritise short-term communications

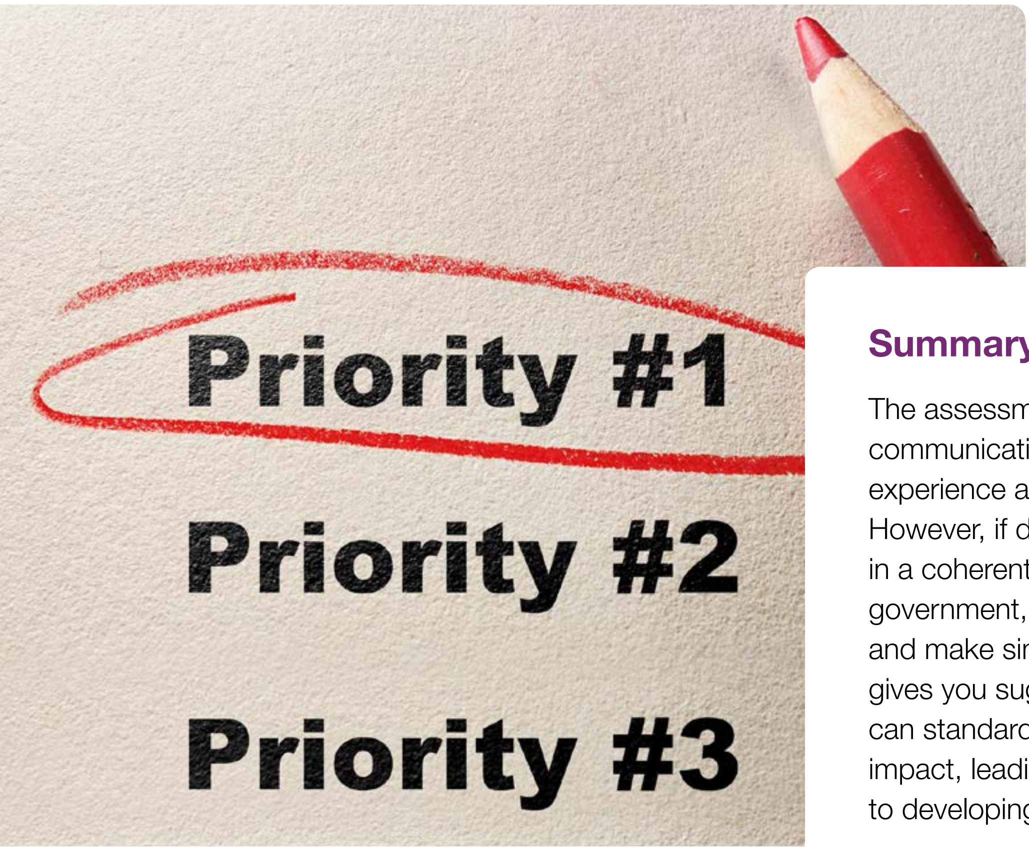
Example: Following the poisoning of two UK residents in Salisbury, a disinformation campaign began around the incident, spread by Russian news sources. Early warnings from digital media enabled the production of briefings for senior staff across government to understand the scale and impact of the disinformation.

	Description	Actions	Audiences	Tools
Medium	The disinformation has the potential to negatively affect a policy area, departmental reputation or a large stakeholder group and is trending online. It requires a response.	Make senior staff and SpAds/policy advisers aware of the issue. Share insight and analysis within department. Investigate the issue and prepare press lines based on known facts.	- Senior staff - Policy advisers	- Insight - Briefings - Press lines - Prioritise short and medium-term communications

Example: A trade press with limited circulation misleadingly claims that a recent parliamentary vote determined that animals have no feelings. Early warning assessment highlights a risk that the narrative may be picked up by mainstream press. Insight, briefings and press lines are prepared either to proactively correct the story or to prepare for possible mainstream interest in policy area.

	Description	Actions	Audiences	Tools
Low	The disinformation has the potential to affect the climate of debate and has limited circulation. The debate should be routinely followed but intervention is unnecessary/ undesirable.	Share insight and analysis within media department. Investigate the issue and prepare press lines/narratives based on known facts. Conduct a baseline analysis of debate and track any changes.	- Communications officers	- Insight - Press lines - Baseline analysis - Prioritise medium and long-term communications

Example: A conspiracy theory has emerged holding the government responsible for a major public safety incident. The theory is only being circulated by fringe groups known for anti-government sentiment, and runs counter to current mainstream debates. Insight and press lines are prepared, but no response is made for the time being. The area is monitored and baseline analysis is used to spot any sudden changes in the climate of debate.



Summary

The assessment of risk and impact in communication work is often the result of experience and a qualified ‘gut feeling’. However, if disinformation is to be tackled in a coherent and consistent way across government, we need to use common tools and make similar assessments. This section gives you suggestions for approaches that can standardise the assessment of risk and impact, leading to a priorities-based approach to developing a response.

Strategic communication

This section will help you answer the following questions:

- What should a public **response** to disinformation look like?
- What is the **sign-off process?**
- **What are the available options for responding**, whether short-term/reactive, medium-term/proactive options or long-term/strategic?

The development of a response needs to follow certain key principles of GCS communications. These include the style of response, the sign-off process, and the response strategy, including content creation over different timeframes.

7.1 What should a public response to disinformation look like?

Not all disinformation has to be responded to. In many circumstances, public opinion will self-correct. Any public response to disinformation that you do decide to make should represent **the truth, well told**. In order to have an immediate and lasting impact your response should be:

Counter-brand, not counter narrative

Countering individual narratives can be ineffective and in many cases has the impact of amplifying or entrenching the falsehood. Generally first impressions are the most resilient, and audiences do not always later remember that a particular piece of disinformation was false. Information overload leads people to take shortcuts in determining the trustworthiness of messages, and familiar themes or messages can be appealing even if they are false. This means a more nuanced and strategic approach than rebuttal is required. This can focus on **framing the tactic of disinformation**, contextualising and outwardly communicating the **motives** or errors of the actor/adversary and not replying directly to their message. It can also focus on providing an **alternative vision** to any that the disinformation narrative has provided. You should develop and stick to a strong, shared narrative so that all communications are coherent, in contrast to a potential multiplicity of disinformation narratives.

Accurate and values-driven

Government and public sector communications must exemplify the values we seek to uphold: truthfulness, openness, fairness and accuracy. Communicating in this way will enable us to build and maintain trust with our audiences.

Timely

The speed and agility of your response is crucial in countering disinformation. This can mean working to faster deadlines than is usual and developing protocols for responding that balance speed with formal approval from senior officials and ministers.

It can also mean knowing when to wait, for example for more information to come to light. Rather than simply producing a **fast** response, think in terms of a timely response.

Edgy

Disinformation narratives often have an impact because they are sensational and attention-grabbing. Your communications will need to be edgy and interesting enough to compete. While remaining true to the principles above, you should consider stepping outside the usual ‘tick box’ government responses and creating an approach or a narrative that will carry in a crowded information space.

Work with friendly influencers

Who is the most credible deliverer of your messages? Third party actors can be a valuable means of building bridges to sceptical audiences, particularly if they are seen as an objective source of credible information.

Case study: Counter-Daesh Global Coalition communications

Counter-Daesh communications originally focused on countering Daesh’s propaganda by rebutting and refuting their claims, but quickly realised it was more effective to expose Daesh’s false narratives of life under Daesh. Moving into a proactive posture, the cell launched whole-of-coalition campaigns like ‘Take Daesh Down’ and increased the positive messaging focused on life after Daesh illuminating important international and grassroots stabilisation efforts in Iraq and Syria.

7.2 The sign-off process

You will need to develop a sign-off process suited to your organisational setup. Ask the following questions:

Who will sign off content?

This needs to include the minimum number of people who absolutely need to review and sign off on your content, for example your **Head of News and/or Head of Communications**, followed by the relevant Special Adviser. You should secure delegates for each of these who will be able to respond on their behalf if they are absent. If you have been creating and sharing situational insight in the form of monitoring reports – as set out in the situational insight section – this will help people to understand the context in advance. They will already have an understanding of disinformation affecting your organisation or its policy areas, which will help when you need to build a case to respond and when you want to clear content quickly.

How quickly should content be signed off?

You should explain to those people signing off your content that they need to do so quickly in order for it to have the required impact. This could well mean providing much quicker deadlines than they are used to in the normal course of business, for example compared to signing off a press release, so again it will help you to establish likely timescales in advance. In a crisis situation the response time required to sign of content could be **within an hour**. To provide an appropriate deadline in individual instances you will need to judge the severity of the incident accurately – this toolkit gives you the tools to do this.

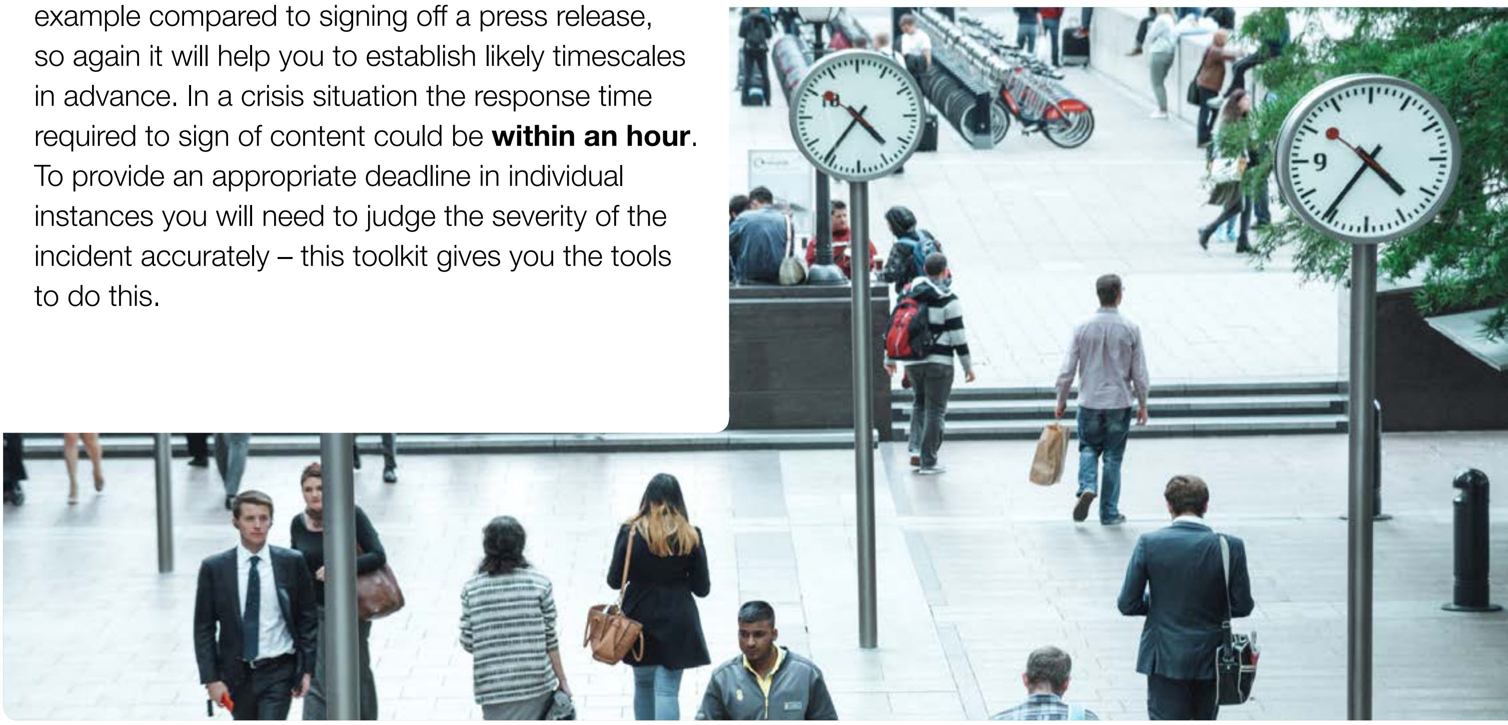
Can lines be pre-cleared?

If insight is already available into an emerging disinformation trend, it may be possible to pre-clear some press lines before they are needed. For example, some government departments have weekly routines to pre-clear lines via their subject experts several days before an event is likely to make the news, in case of negative reporting.

Case study: Counter-Daesh Global Coalition communications

Terrorist networks like Daesh are agile, swift and proactive in their messaging. The cell and its partners had to become agile and swift to get the advantage in the information environment.

Because the cell spent time and resources on developing a cross-government and international stakeholder network early, they built trust and credibility to message on behalf of partners early in the formation of the coalition. This approach helped them to remove roadblocks and layers of the process that might exist for other communications efforts.



7.3 Response strategies

When you have conducted your risk assessment, you will have reached a conclusion about the priority of the disinformation. This will enable you to consider a range of communicative tools which you can then tailor to relevant target groups. Generally, **the higher the priority, the more focus should be placed on short-term reactive responses, at least initially.** Note that a combination of short, medium and long-term approaches may be necessary, depending on the priority of the issue. You should use the **OASIS model** to plan your communication activities (see annex E).

	Action	Target groups	Tools
Short-term Reactive	The disinformation requires an immediate response. Use the FACT model to help identify a rapid communications response to correct or counter disinformation in accordance with the established facts (see Annex E)	- Traditional media (journalists/editors) - Stakeholders and influencers - Social media platforms - Key audiences	- Holding statement - Press statement - Minister statement - Brief journalists - Q&A - Paid advertisement - Search engine optimisation (SEO) - Expose actors via friendly influencers

Example: Disinformation has reached the mainstream press. The response is to brief journalists and request an amendment to published stories, place a Q&A on GOV.UK and use SEO to ensure government information is the highest ranked article on Google.

	Action	Target groups	Tools
Medium-term Proactive	The disinformation requires a considered response. Use a combination of communications to assert own values/brands. Tie proactive measures with your normal everyday communications and work with stakeholders/ influencers to create consensus around your position.	- Traditional media (journalists/editors) - Stakeholders and influencers - Social media platforms - Wide audiences	- Campaign, narrative and brand development - Community outreach, dialogue and engagement - Facilitate network, stakeholders and influencers - Workshops/training

Example: Disinformation is engaged with on social media and closed chat rooms. The response is to develop the profile of the issue through brand and narrative development, and engage with a variety of intermediaries with these materials.

	Action	Target groups	Tools
Long-term Strategic	The disinformation requires a coherent, sustained response to create long-term change. Develop and assert strategic narratives in relation to an issue by shaping the information space to promote your own position and deter others (raising the threshold).	- Traditional media (journalists/ editors) - Young up-and-comers - Stakeholders and influencers - Social media platforms - Wide audiences	- Campaign, narrative and brand engagement - Programme funding e.g. for participatory content - Talent spotting and influencer support/creation - Facilitate network, stakeholders and influencers - Workshops/training - Contingency planning

Example: Disinformation is engaged with by fringe groups as a form of conspiracy theory. The response is to look to emerging voices within these fringe groups and to provide them with training and opportunities. Another response is a public information campaign combined with public participation in content creation, such as in the form of a youth competition.

	Action	Target groups	Tools
Ignore	The disinformation is unlikely to have a major impact or receive widespread attention. It does not require intervention but can be monitored if necessary.	- Monitor those concerned if necessary - Record data on the case and document your assessment for future reference	- Use monitoring and insight templates to record events

Example: Disinformation is occurring but engagement levels are low. No response is necessary but the case is logged and a small number of accounts are added to general monitoring routines.

Summary

When developing a response to disinformation, you should consider the **style of communication**, the **routines by which you approve messaging**, and the **timeframes of your communication activities**.

Track outcomes

This section will help you answer the following questions:

- How should I **record and share** information about the disinformation campaign?
- How can I **evaluate** my actions and understand the **lessons learned?**

Many of the basic questions that you need to answer are listed below. An example scoresheet is included in annex F.

8.1 Recording and sharing information

Tracking outcomes in relation to disinformation refers to two tasks:

- documenting and sharing data on cases of disinformation; and
- assessing the effect of your decision-making and actions.

Keep in mind that you will not be attempting to track the outcomes of the disinformation, but rather the effectiveness and relevance of your efforts. This is crucial for ensuring that countermeasures are on point and congruent with analysis and long-term objectives.

In the course of identifying and responding to disinformation, much data is already recorded on a case. This data needs to be paired with the conclusions of the evaluation to provide a record of the full process of response.

Recognise disinformation: Provide a bottom-line overview of the disinformation techniques used in the disinformation, including visual examples.

- What was the goal of the disinformation?
- What disinformation techniques were used?
- How were the disinformation techniques combined to achieve an impact?

Early warning: Consider your preparatory work and the extent to which it supported your efforts to handle disinformation.

- Is your digital monitoring sufficiently focused on your priorities?

Situational insight: Once you have identified disinformation, consider how well your initial analysis and situational briefing supported your team's response.

- Were we able to offer an accurate and timely briefing to colleagues?
- Did we make any incorrect assumptions? On what basis?

Impact analysis: Consider your assessment of the likely goals, impact and reach of the disinformation.

- Was the disinformation prioritised correctly, based on goals, impact and reach?

Strategic communication: Provide an overview of the communicative responses you took, broken down into actions, target groups and tools.

Track outcomes: Collect this information in a dossier together with your assessments of the actions you took.

- What was the impact of your efforts to handle the disinformation?
- What lessons can be learned from this case?

Annex A: RECOGNISE DISINFORMATION

Glossary of disinformation techniques

Technique	Example
ASTROTURFING (I) Falsely attributing a message or an organisation to an organic grassroots movement to create false credibility.	A source pays or plants information that appears to originate organically or as a grassroots movement.
BANDWAGON EFFECT (S) A cognitive effect where beliefs increase in strength because they are shared by others.	A person is more willing to share an article when seeing it is shared by many people.
BOT (I, T) Automated computer software that performs repetitive tasks along a set of algorithms. - **IMPERSONATOR BOTS (I, T)** Bots which mimic natural user characteristics to give the impression of a real person. - **SPAMMER BOTS (I, R, T)** Bots which post repeat content with high frequency to overload the information environment.	Bots can be used to amplify disinformation or to skew online discussion by producing posts and comments on social media forums and other similar tasks – sometimes they focus on quantity and speed (spammer bots); other times they attempt to mimic organic user behaviour (impersonator bots) – bots can also be used for hacking and to spread malware.
BOTNET (I, T) A network of hijacked computers used to execute commands.	Infests personal computers with malware, contribute to DDoS attacks, and distributing phishing attacks.
CHEERLEADING (R) The overwhelming promotion of positive messages.	A dissenting opinion is crowded out by positive messages perpetuated by an abundance of commentators cheerleading the 'right' opinion.
DARK ADS (F, T) Targeted advertisement based on an individual user's psychographic profile, 'dark' insofar as they are only visible to targeted users.	An advertisement containing false information is targeted to social media users with personality traits deemed susceptible to this messaging, with the goal of shaping their opinions in a specific direction.
DDoS ATTACKS (T) Distributed Denial of Service (DDoS) is a cyber-attack where multiple IP addresses are used to disrupt services of a host connected to the internet.	A DDoS attack is conducted to bring down a government website during a crisis, to deny citizens access to reliable information.

DEEPFAKES (F, I, T) Use of digital technology to fabricate facial movements and voice, sometimes in real time.	A fabricated video of a politician shows them saying something outrageous or incriminating, with the goal of undermining confidence in government.
ECHO CHAMBER (S) A situation where certain ideas are reinforced by repetition within a social space online.	Creation of internet sub-groups, often along ideological lines, where people engage with like-minded people, which reinforces pre-existing beliefs.
FAKE NEWS (F) Deliberate disinformation disguised as news.	A non-journalist fabricates a news story to influence public opinion and to undermine the credibility of mainstream media, which is published on a private platform.
FAKE PLATFORM (I) Identity of a web platform is disguised to promote fabricated content.	A web platform is designed to appear like an official site, with the goal of creating the appearance of a credible source of information.
FILTER BUBBLE (I, T) Algorithms which personalise and customise a user's experience on social media platforms might entrap the user in a bubble of his or her own making.	The social media flow of a user interested in Brexit gradually adapts to consumed content to eventually only show information in favour of Brexit.
FLOODING (T) The overflowing of a target media system with high-volume, multi-channel disinformation.	Multiple commentators, both in the form of bots and real users, make an overwhelming amount of posts with nonsense content to crows out legitimate information.
FORGERY (F, I) Product or content is wholly or partly fabricated to falsely ascribe the identity of the source.	A false document with an official-looking government heading is produced to discredit the government.
HACKING Use of illegitimate means to unlawfully gain access to, or otherwise disturb the function of, a platform.	An actor illegitimately claims access to a network from which private information, such as emails, is extracted.
HIJACKING (S, T) Unlawful seizure of a computer or an account.	A website, hashtag, meme, event or social movement is taken over by an adversary or someone else for a different purpose.
LAUNDERING (F, I) The process of passing of disinformation as legitimate information by gradually distorting it and obscuring its true origin.	A false quote is referenced through multiple fake media channels until the original source is obscured and the quote is accepted as real by legitimate actors.

LEAKING (S, T) Disseminating unlawfully obtained information.	Unlawfully obtained emails are leaked to compromise individual actors or to undermine public confidence.
MALIGN RHETORIC (R) Lingual ruses aimed at undermining reasonable and legitimate debate and silencing opinions. - **NAME CALLING (R)** A classic propaganda technique based on abusive or insulting language directed against a person or a group. - **AD HOMINEM (R)** Argumentative strategy focused on attacking the person making the argument rather than the content of the argument itself. - **WHATABOUTERY (R)** A rhetorical maneouvre which discredits an opponent's position by accusing them about unrelated issues. - **GISH GALLOP (R)** A debate tactic focused on drowning the opponent in an overwhelming amount of weak arguments which require great effort to rebut as a whole. - **TRANSFER (R)** A classic propaganda technique based on transferring blame or responsibility to associate arguments with admired or despised categories of thought. - **STRAWMAN (R)** A form or argument which targets and refutes an argument that has not been present in the discussion.	A combination of different rhetorical moves is applied in online conversation to ridicule and diminish other opinions.
MANIPULATION (F) Alteration of content to change its meaning.	An image is cropped to only show some of the participating parties in an incident.
MISAPPROPRIATION (I) Falsely ascribing an argument or a position to another's name.	A public figure is incorrectly cited or falsely attributed as a source.
PHISHING (I, T) A method to unlawfully obtain information online via malware distributed over emails or web platforms.	Malicious links are distributed via email which lead to phishing sites.
POINT AND SHRIEK (S) Exploitation of sensitivity to perceived injustices in society to create outrage.	A commentator diverts from a real issue at hand by pointing out the audacity of a make-belief incident which play on pre-existing social grievances.

POTEMKIN VILLAGE (I, R) A smoke-screen of institutions and/or platforms established to deceive audiences.	A complex network of fake think tanks is established to disseminate disinformation which seems legitimate due to the perceived legitimacy of the network.
RAIDING (S, T) Temporarily disrupting a platform, event, or conversation by a sudden show of force.	Several automated accounts are coordinated to disrupt a conversation by temporarily spamming nonsense messages.
SATIRE AND PARODY (R, S) Ridiculing and humouring of individuals, narratives or opinions to undermine their legitimacy.	A public figure is ridiculed using memes where non-factual opinions are ascribed to the public figure.
SHILLING (I) To give credibility to a person or a message without disclosing intentions or relationships.	An actor endorses certain content while appearing to be neutral but is in fact a dedicated propagandist.
SOCKPUPPETS (I, R, T) Use of digital technology to disguise identity, to play both sides of a debate.	A user creates two or more social media accounts under opposing identities i.e. one pro-fox hunting, one against, with the aim of playing the identities against one another.
SPIRAL OF SILENCE (S) The decrease in audibility of deviant opinions due to non-conforming beliefs.	A person with non-conforming minority beliefs is less willing to share his or her opinions.
SYMBOLIC ACTION (S) Refer to acts that carry symbolic value in the sense that they signal something to an audience to create a response.	A user plays on universally shared symbolic cues e.g. terrorist attacks to create a climate of fear.
TAINTING (F, S, T) Leaked contents are tainted with forgeries.	Leaked documents are distributed together with carefully placed fakes.
TERRORISM (R, S) Imagery from real-world events is used to make political claims.	Images of violence are used to support false narratives, with the aim of creating a climate of fear or justifying a political argument.
TROLLING (I, R, S) Deliberate commenting on internet forums to provoke and engage other users in argument.	Social media users deliberately post provocative comments to create emotional outrage in other users.
WOOZLE EFFECT (R) Self-perpetuating evidence by citation.	A false source is cited repeatedly to the point where it is believed to be true because of its repeated citation.

Browser extensions

CHROME browser extensions
(https://chrome.google.com/webstore)

Awesome Screenshot: capture and annotate: captures all or part of webpage or record screen as video – includes useful tools such as blur sensitive info, add comments.

Evernote Web Clipper: documents your research processes by saving webpages and screenshots, highlight key info, add graphics and text on saved items.

Google Translate: translates words, phrases or websites to more than 100 different languages.

OSINT resources

Inteltechniques (www.inteltechniques.com): locates personal info about any target using different search tools and automated analysis.

Metacrawler (www.metacrawler.com/): metasearch engine which accepts a single search request from the user – extends the search coverage of the topic and allows more information to be found by sending multiple queries to several other search engines.

OSoMe tools (https://osome.iuni.iu.edu/tools/): tools developed by Indiana University that let you analyse trends, maps and networks.

SimilarWeb (www.similarweb.com): a competitive intelligence tool that collects data from various sources and categorises events, keywords etc; generates and exports graphs, tables, and other visuals based on collected data.

The Search Engine List (www.thesearchenginelist.com): provides search engines in different categories, such as all-purpose search engines, blogs, meta search, multi media, news, open source, and visual search engines.

Toddington (www.toddington.com/resources): provides search tools and resources within different categories, such as news and journalism, username search, webpage analysis, and social media.

Image and video search

Amnesty International´s Youtube DataViewer (https://citizenevidence.amnestyusa.org/): identifies where an image or video appears online.

Berify (www.berify.com): upload an image or video and find out if the image or video is distributed at other websites – notifies you when someone uses your images.

Google Image (www.images.google.com): find similar images, webpages where an image has been published.

Jeffrey's Image Metadata Viewer (http://exif.regex.info/exif.cgi): gives you image data, such as when and where a picture was taken (also called Exif reader).

Labnol Reverse Image Search (www.labnol.org): upload an image and search on Google to verify the source.

TinEye (https://tineye.com/): find out where an image appears online; discovers modified or edited versions of an image.

Social media monitoring

Agora Pulse (www.agorapulse.com): synchronises your social media accounts around the clock, offers unlimited reports and graphics of performance analytics, retains all your account data, compares your page with others on key metrics.

Botometer (https://botometer.iuni.iu.edu/#!/): decides whether the account is a bot by analysing its tweets, its followers and when and where tweets are published.

Facebook for developers (https://developers.facebook.com/docs/graph-api/overview/): use the graph API which is the primary way to get data into and out of the Facebook platform.

Foller.me (https://foller.me/): gathers information about a specific Twitter user; conducts automatized analyses based on tweet's contents on topics, hashtags, mentions, attitudes, activity time.

Followerwonk (https://followerwonk.com/): helps you explore your social graph – find out who is following you, their location and when they tweet; connect with influencers; compare your graph with others.

Hootsuite (https://hootsuite.com/): social media listening tool with specific search terms in real-time – this can be useful for tracking mentions of your brand, products, or relevant keywords you are interested in.

Iconossquare (https://pro.iconosquare.com/) : effectively manage conversations and your social media accounts; make communication plans.

Jollor (www.jollor.com): monitors and analyses social media data – identifies key influencers and offers unlimited reports and downloadable charts for measuring performance (integrates with Instagram and YouTube).

Social Searcher (https://www.social-searcher.com/): monitors public social mentions on social networks and web – quickly find what people are saying about an issue.

Sprout Social (www.sproutsocial.com): a popular and user friendly social media management software – contains tools such as social performance reporting, advanced social analytics, social monitoring and listening tools, and advanced social listening (at the moment does not include visual networks such as YouTube).

Twitterfall (https://twitterfall.com/): collects tweets based on real-time tweet searches.

Twitter for developers (https://developer.twitter.com): stream Twitter data to enable analysis in real-time or back in time; use different API filters to find out more about key topics, breaking news etc.

Network analysis

Alexa Internet (https://www.alexa.com/): provides various tools based on commercial web traffic data, such as keyword research tools, competitive analysis tools, audience analysis tools and much more.

Analyst´s Notebook (www.ibm.com): provides visual analysis tools with focus on identifying and disrupting criminal, cyber and fraudulent threats – connected network visualisations, social network analysis, and geospatial or temporal views to uncover hidden connections and patterns in data.

Crimson Hexagon (https://www.crimsonhexagon.com/): social media monitoring and analysis platform that gives you access to over one trillion consumer conversations from social media – also provides many other tools such as advanced image analytics.

Hoaxy (https://hoaxy.iuni.iu.edu/): visualizes the spread of articles online (Twitter is currently the only social network tracked by Hoaxy, and only publicly posted tweets appear in the visualizations).

Maltego (https://www.paterva.com/web7/index.php): focuses on providing a library of transforms for discovery of data from open sources – this information is then displayed on a node-based graph suited for performing link analysis.

Mediacloud (https://mediacloud.org/): open source platform for studying media ecosystems – it chooses a set of media sources and uncovers the feeds; each feed is trawled to determine if any stories have been added; all content is then extracted of each relevant story.

Other

Automating OSINT (https://register.automatingosint.com/): open source intelligence training course – learn how to code and automatically extract and analyse data from webpages and social media.

PropOrNot (http://www.propornot.com/p/the-list.html): gathers and exposes Russian efforts to influence US opinion using propaganda.

Quetext (https://www.quetext.com/): plagiarism checker tool that looks for duplicate content online.

Junk News Aggregator (https://newsaggregator.oii.ox.ac.uk/about.php): evaluates the spread of junk news on Facebook to identify junk news sources that publish misleading, deceptive or incorrect information purporting to be real news – the aggregator shows junk posts along with how many reactions they received.

Annex C: SITUATIONAL INSIGHT

Disinformation insight report template	
Key insights	Summary of the top 3 points your stakeholders should know including any recommendations for action.
Event summary	Concise explanation of the issue.
Disinformation narratives	Identification of each false narrative, plus reference by source and date.
Disinformation examples	Any visual material, e.g images of key Twitter posts, or other supporting material e.g. quotes from press statements.
Other points of note	Anything else of relevance to the issue.
Trends over time	Whether/how the disinformation has changed and a brief analysis of how.
Recommendations	Recommendations for response and outline of suggested strategy.

Annex D: IMPACT ANALYSIS

This annex presents three advanced structured analytic techniques useful for analysing disinformation on a more advanced level. These techniques will expand your analytical toolset in cases where you need to test hypotheses in a more rigorous way. During a major crisis such as the Salisbury poisoning, for example, these techniques would have been used to weigh up evidence of Kremlin involvement against alternative explanations.

Key Assumptions Check (KAC)

A key assumption is a piece of information that analysts accept as true and forms the basis of their assessment. With reference to disinformation this could range from a political position of an influencer, the attributed source of a message, the composition of a specific target audience or the reach of a social media platform. Unstated assumptions often underpin analysis. A KAC articulates and reviews these assumptions to ensure that analysis is not based on a faulty premise. A KAC can also help you to develop indicators that would cause you to abandon an assumption, which can be useful for re-directing resources.

KAC step-by-step:

1. Define and document your current reasoning (analytic line).

2. Identify and articulate all premises which are accepted as true for your reasoning to be valid.

3. Challenge each assumption by asking why it 'must' be true – How confident are you that the assumption is correct, and what explains your degree of confidence?

4. Refine your key assumptions to those that must be true for your line of reasoning to work and consider under what conditions these assumptions may not hold up – How would disproving a key assumption alter your line of reasoning?

Example: Using KAC to assess the likely reach of disinformation about a major public safety incident

Analytic line

There is little risk of disinformation related to the incident. Authorities are using their own communications in a timely and appropriate manner, and legacy media is reporting live on developments. The information space is saturated by credible sources, and it seems unlikely that disinformation will perpetuate the debate.

Key assumptions	Assessment
Legacy media is primary source of information for target audiences;	Possible but not likely given rapid developments of social media – highly likely social media coverage will be prominent;
Relevant target audiences are known to us; and	Niche audiences exist, and it is possible for them to circulate information unbeknownst to us; and
In case of disinformation, it will likely critique the management of the incident.	Likely, but disinformation could equally well distract from management of the incident or relate to the incident in any other way.

Key Assumption Check

We should not dismiss the possibility of disinformation related to the incident, and we should closely monitor social media channels where fringe groups may start rumours, spread falsehoods and establish hostile narratives that can worsen the situation on the ground and diminish trust in the authorities.

Quality of Information Check (QIC)

Key assumptions are only part of your assessment of any given situation. Evidence constitutes another element which should be properly examined and reviewed in complex cases. A QIC will help you to evaluate the completeness and soundness of the evidence you base your assessment on. This involves both weighing the validity and reliability of a source, as well as reflecting on your interpretation of certain pieces of information. This is an ongoing process that should be revisited periodically during an analytical process.

A QIC helps you to avoid anchoring judgement on weak information as well as differentiating between what we know and what we do not know. This is important for identifying information gaps and for detecting possible deception and denial strategies by adversaries dealing with disinformation. It will also help you understand how much confidence you can place in your analytical judgment.

QIC step-by-step:

1. Map and plot key pieces of available information and systematically review sources for accuracy.

2. Identify pieces of information that appear most relevant to you and check for sufficient corroboration; flag pieces of information which are not deemed reliable and valid and provide motivation for your judgement; indicate a level of confidence for each source.

3. Consider multiple interpretations of ambiguous information and provide caveats where appropriate.

4. Document your findings as they may be useful for both other analysts and as input to an analysis of competing hypotheses (see next page).

Example: Using QIC to map disinformation

Disinformation about public safety incident					
Source	**Information**	**Link**	**Date**	**Comment**	**Reliability rating**
UK national news site	Live news reporting on incident	www.example.com	14 June 2017	Congruent with official statements	B1
International news site	Current events report with panel discussion	www.example.com	14 June 2017	Fairly congruent with official statement but with elements of opinion journalism	C3
International news site	Article on incident	www.example.com	14 June 2017	Contradicts official statements – claim higher death toll and slower response	D5
Twitter	Multiple accounts engaging under #publicsafety	www.example.com	14 June 2017	Narrative centred on injured people and how official sources have not commented – mix of users, some exhibit automated behaviour	E6

Source reliability: A (reliable), B (usually reliable), C (fairly reliable), D (not usually reliable), E (unreliable), F (reliability unknown)

Information reliability: 1 (confirmed by independent sources), 2 (probably true), 3 (possibly true), 4 (doubtful), 5 (improbable), 6 (cannot be determined)

Analysis of Competing Hypotheses (ACH)

The ACH method identifies all reasonable alternative explanations for an observed phenomenon and then assesses the relevance of evidence against possible explanations. This prevents premature conclusions and ensures that different possibilities are explored and evaluated equally before further actions are developed.

ACH helps analysts avoid heuristic shortcuts such as confirmation bias (tendency to interpret information in a way that confirms already held beliefs), selective perception (letting expectations affect perception), ambiguity effect (avoiding options where information is missing), anchoring (relying on past references and experiences), and world-view backfire-effect (discounting information that does not fit within your understanding of the world).

ACH can be used by a single analyst to generate a range of possible explanations but is most effective when multiple analysts challenge each other's hypotheses.

ACH step-by-step:

1. Generate different possible hypotheses to explain the case under analysis (either individually or through brainstorming with a group) – do not dismiss any hypothesis at this stage, explore every hypothesis you generate.
2. Identify and document significant evidence relevant to the case and to the different hypotheses – build on evidence examined in your QIC (see above).
3. Prepare a matrix of competing hypotheses where hypotheses and evidence are plotted on the different axes; assess each piece of evidence in relation to each hypothesis to determine if they are consistent, inconsistent or not applicable; focus should be on disproving rather than proving hypotheses.
4. Analyse the results and revisit your original hypotheses, evidence and other assumptions (the matrix can be refined and revisited for better results if time permits) – make sure to consider if there are pieces of evidence not being seen which would be expected for a certain hypothesis to be true and contemplate on why it has not been included (maybe you have been biased in your selection?).

Example: Using ACH to assess the goals of disinformation

Disinformation about public safety incident					
			H1	H2	H3
			Extremist group disseminating disinformation to rally support	Foreign actor disseminating disinformation to undermine government	Private individuals disseminating disinformation for the sake of attention
	Inconsistency score ->		-2	-1	-1
	Evidence	Weight			
E1	Disinformation on casualties	Medium	I	I	C
E2	Disinformation on government response	High	C	C	N
E3	Disinformation on missing people	Low	N	N	C
E4	Bots used	High	I	C	I

Explanation: ACH matrix on disinformation about the public safety incident. After noticing disinformation being spread about the incident on social media, a group of communicators use ACH to examine possible explanations of the disinformation to determine the risk of the disinformation reaching a wider audience and in extension have an impact on the function of a specific department. Three hypotheses are generated through a brainstorming session and evaluated against four pieces of evidence obtained through media monitoring.

The process allows the communicators to see that the available evidence is more consistent with a foreign actor or private individuals as the disseminators of disinformation in this case, perhaps a combination of both. Since H2 is consistent with both pieces of evidence weighed as ‘high’ (E2 and E4), the unit will proceed with the hypothesis that there is a foreign actor involved. This information will be useful for designing a response. The matrix can be updated when more evidence is available or if other hypotheses are generated later in the process.

Additional resources: This annex draws heavily on the US Government document ‘A Tradecraft Primer: Structured Analytic Techniques for Improving Intelligence Analysis’ and Heuer and Phearson ‘Structured Analytic Techniques for Intelligence Analysts’.

Annex E: STRATEGIC COMMUNICATION

The FACT model

Misinformation and disinformation: a rapid response guide for government media professionals.
The FACT model has been developed by HMG's Rapid Response Unit (RRU) for straightforward application in everyday communications activity. It consists of 4 key steps, which can be tailored to suit departmental capability.

	Essential For teams with no online media monitoring capability	**Intermediate** For teams with some online media monitoring capability	**Advanced** For teams with established online media monitoring capabilities
FIND misinformation or disinformation, through continuous media monitoring	Subscribe to MMU Updates. Sign-up to Google Alerts. Use free tools to monitor Twitter.	Actively search for misleading social media posts and online articles.	Identify longer-term narratives and trends (through in-depth analysis and social listening tools), following best practice set out in the GCS RESIST counter disinformation toolkit.
ASSESS the risk of the inaccurate or misleading narrative	Use MMU Updates to: – judge if stories are misleading/inaccurate – identify key influencers – measure scale of interaction	Use tools to analyse engagement (retweets, shares, views, comments, reactions).	Use the GCS RESIST disinformation toolkit to calculate the long term level of risk.
CREATE HMG content to counter this risk	Simple content, including: – departmental blog – GOV.UK post – reactive social media posts Share existing content.	More engaging content: – videos – images/graphics	Multi-channel content that resonates with the affected audience: – videos and images – op-eds – long-term campaigns – influencer collaboration
TARGET this content at relevant audiences	Organic social media targeting Direct response to posts/articles. Contact publisher/author directly.	Social media advertising activity targeted at relevant audiences.	Search advertising targeted at people seeking information relating to specific topics. Different content and messaging targeted to segmented audiences across social media and relevant digital media channels.

The OASIS model

All planned government communication activities and campaigns use the OASIS model to ensure effective and efficient communications. The OASIS model has the benefit of not only providing a coherent framework for government communications, but also of contextualizing individual communication activities to wider campaigns and strategic narratives.

	OASIS	RESIST
Objectives	Set out what the communications activity is intending to achieve based on policy aims. Develop communication objectives which are achievable, measurable and focused on outcomes rather than outputs.	Contextualize your objectives to your knowledge of the disinformation issue at hand (early warning, situation insight, impact analysis) to formulate clear objectives related to desired changes in attitudes and behaviours. Align disinformation objectives to wider policy objectives.
Audience insight	Who are the audiences of the campaign and how do you need to influence them to achieve your objectives? What opportunities/ barriers exits?	Through your early warning and situational insights analyses you will have a clear understanding of the target audiences of disinformation. You will target your activities accordingly but not necessarily to the same audiences. Consider how the audiences have been affected by disinformation and what challenges this poses for your response.
Strategy/ideas	Set your approach in relation to position/messaging; channels; partners/influencers. Design communication plan with audience journey and test approach to assess effectiveness.	The response level derived at using the impact analysis method should guide your strategy by providing insight into the kind of response needed. It will give insight into how communications should be structured to prevent negative impact.
Implementation	Set out how your communications should be delivered and develop a clear plan that allocates resources and provides a timescale for delivery.	The tools of your designated response level should guide the implementation considerations of your communications plan when deciding how to allocate resources and deliver using the suggested tools.
Scoring/ evaluation	Monitor outputs, outtakes and outcomes through the campaign and evaluate upon completion to discern the effectiveness of your communication activities.	Scoring/evaluation of your planned communications serves a different purpose from tracking outcomes of the RESIST model, as it focused on the communication activities specifically rather than your response to disinformation as a whole.

Annex F: TRACK OUTCOMES

RESIST – Track outcomes template

Name: Date: Department: Unit:	Case: Date of discovery: Responsible:

Short description of the case:

Recognise

Perceived goal/objective		RESIST 3.1
Disinformation technique(s)		RESIST 3.2
Combination of techniques		RESIST 3.3

Early Warning

Means of discovery		RESIST 4.2
Digital monitoring used (yes/no)		Describe reason
Relevance of threats and vulnerability assessment		RESIST 4.1
Need for assessment update? (yes/no)		Describe new needs

Situational insight		
Assessment of disinformation: - narratives/arguments - falsehoods - circulation - networks/communities - targets - initial recommendations		RESIST 5.1
Did initial analysis support response (yes/no)		Describe reason
Incorrect assumptions made		Describe
Briefing relevance		

Impact analysis		
Objectives: - observable effect - beneficiary - disadvantaged - required action		RESIST 6.1
Affected parties		RESIST 6.2 (list)
Exposure /reach		RESIST 6.3 (list)
Priority level		RESIST 6.4
Priority level accuracy		RESIST 6.4

Strategic communication		
Response(s) taken		RESIST 7.3
Target groups selected		RESIST 7.3
Tools used		RESIST 7.3
Sign-off experience		RESIST 7.2
Timings		Timing from decision to release of product

Track outcomes		
Impact		What was the impact of your efforts on TAs?
Lessons learned		
Will an in-depth evaluation using OASIS be conducted? (yes/no)		

Further resources

Contact details

Sign up for the Research, Information and Communications Unit RDAT disinformation reports and request advice: at DisinformationAnalysisTeam@homeoffice.x.gsi.gov.uk

Find out more about MMU social media research agency:
mmu@cabinetoffice.gov.uk

Contact the Rapid Response Unit:
rru@cabinetoffice.gov.uk

The FCO's Open Source Unit analyse and report on international open source material: osu@fco.gov.uk

The National Security Communications Team deliver the Government's UK counter-disinformation communications strategy and run the GCS training programme: Nat-sec-comms@cabinetoffice.gov.uk

GCS International offer disinformation training internationally:
GCSI@cabinetoffice.gov.uk

Further reading

Althuis, J., Haiden, L. (eds) (2018) Fake New: A Roadmap NATO Strategic Communications Centre of Excellence and King's Centre for Strategic Communications https://www.stratcomcoe.org/fake-news-roadmap

Atlantic Council (2018) Disinfo Portal https://disinfoportal.org/

Bjola, C., Pamment, J. (eds) (2018) Countering Online Propaganda and Extremism: The Dark Side of Digital Diplomacy

Bodine-Baron, E., Helmus, T., Radin, A., Treyger, E. (2018) Countering Russian Social Media Influence, RAND https://www.rand.org/content/dam/rand/pubs/research_reports/RR2700/RR2740/RAND_RR2740.pdf

Cook, J., Lewandowsky, S. (2012) The Debunking handbook http://www.skepticalscience.com/docs/Debunking_Handbook.pdf

European Commission (2018) A draft code of practice on online disinformation (1st draft) https://ec.europa.eu/digital-single-market/en/news/draft-code-practice-online-disinformation

European Values Centre (2018) 2018 Ranking of countermeasures by the EU28 to the Kremlin's subversion operations, Kremlin Watch Report https://www.kremlinwatch.eu/userfiles/2018-ranking-of-countermeasures-by-the-eu28-to-the-kremlin-s-subversion-operations.pdf

Fly, J., Rosenberger, L., Salvo, D. (2018) Policy Blueprint for Countering Authoritarian Interference in Democracies no. 27 The German Marshall Fund of the United States https://www.intelligence.senate.gov/sites/default/files/documents/a-lrosenberger-080118.PDF

Fried, D., Polyakova, A. (2018) Democratic Defense Against Disinformation, The Atlantic Council http://www.atlanticcouncil.org/images/publications/Democratic_Defense_Against_Disinformation_FINAL.pdf

Full Fact (2018) Tackling misinformation in an open society https://fullfact.org/media/uploads/full_fact_tackling_misinformation_in_an_open_society.pdf

Hindman, M., Barash, V. (2018) Disinformation, 'Fake News' and Influence Campaigns on Twitter. Knight Foundation https://kf-site-production.s3.amazonaws.com/media_elements/files/000/000/238/original/KF-DisinformationReport-final2.pdf

Hug, A. (eds) (2017) The Information Battle: How governments in the former Soviet Union promote their agendas AND attack their opponents abroad, the Foreign Policy Centre https://fpc.org.uk/wp-content/uploads/2017/03/1801.pdf

Jackson, L., Thomas, T., Laity, M., Nimmo, B. (2015) Information at War: From China's Three Warfares to NATO's Narratives, the Legatum Institute
http://www.lse.ac.uk/iga/assets/documents/arena/archives/information-at-war-from-china-s-three-warfares-to-nato-s-narratives-pdf.pdf

Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K., Hobbs, W., Mattsson, C. (2017) Combating Fake News: An Agenda for Research and Action
https://shorensteincenter.org/wp-content/uploads/2017/05/Combating-Fake-News-Agenda-for-Research-1.pdf

Lewandowsky, S., Ecker, U., Seifert, C., Schwarz, N., Cook, J. (2012) Misinformation and Its Correction: Continued Influence and Successful Debiasing, Psychological Science in the Public Interest 13, no. 3 (December 2012): 106–31, https://doi.org/10.1177/1529100612451018

Lewis, R. (2018) Alternative Influence: Broadcasting the Reactionary Right on YouTube, Data & Society Research Institute
https://datasociety.net/wp-content/uploads/2018/09/DS_Alternative_Influence.pdf

Lord Strang (1963) The Unavowable Information Services of Her Majesty's Government Overseas, report on the Foreign Office Information Research Department, CAB 301/399
http://discovery.nationalarchives.gov.uk/details/r/C16747784?fbclid=IwAR2ZWFE2xPU8_bj3t_P6su89bRnYFWxZbZcJkWo96i9mKgmtYZrlgd8o644

Lucas, E., Pomeranzev, P. (2016) Winning the Information War: Techniques and Counter-strategies to Russian Propaganda in Central and Eastern Europe, Center for European Policy Analysis
https://cepa.ecms.pl/files/?id_plik=2706

MSB (Swedish Civil Contingencies Agency) (2018) Countering information influence activities: a handbook for communicators
https://www.msb.se/sv/Produkter--tjanster/Publikationer/Publikationer-fran-MSB/Countering-information-influence-activities--A-handbook-for-communicators/

Nissen, T. (2016) Social Media's Role in 'Hybrid Strategies', NATO Strategic Communications Centre of Excellence
https://www.stratcomcoe.org/social-medias-role-hybrid-strategies-author-thomas-elkjer-nissen

Pamment, J., Nothhaft, H., Agardh-Twetman, H., & Fjällhed, A. (2018) Countering Information Influence Activities
https://www.msb.se/sv/Produkter--tjanster/Publikationer/Publikationer-fran-MSB/Countering-Information-Influence-Activities-The-State-of-the-Art-research-report-/

Paul, C., Matthews, M. (2016) The Russian 'Firehose of falsehood' Propaganda Model – why it might work and options to counter it, RAND
https://www.rand.org/content/dam/rand/pubs/perspectives/PE100/PE198/RAND_PE198.pdf

Philips, W. (2018) The Oxygen of Amplification – Better Practices for Reporting on Extremists, Antagonists, and Manipulators Online, Data & Society Research Institute
https://datasociety.net/wp-content/uploads/2018/05/FULLREPORT_Oxygen_of_Amplification_DS.pdf

Polyakova, A., Boyer, S. (2018) The Future of Political Warfare: Russia the West, and the Coming Age of Global Digital Competition, Brookings Institute
https://www.brookings.edu/wp-content/uploads/2018/03/the-future-of-political-warfare.pdf

Sanovich, S. (2017) Computational Propaganda in Russia - The Origins of Digital Misinformation, Working Paper, Computational Propaganda Research Project, Oxford Internet Institute
https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-Russia.pdf

The Swedish Institute of International Affairs (2017) New Approaches and the Way Forward in Strategic Communications, Stratcom Stockholm
https://www.ui.se/globalassets/ui.se-eng/news/2017/conference-report-stratcom-1-december-2017.pdf

Treverton, G., Thvedt, A., Chen, A., Lee, K., McCue, M. (2018) Addressing Hybrid Threats, Swedish Defence University
https://www.hybridcoe.fi/wp-content/uploads/2018/05/Treverton-AddressingHybridThreats.pdf

Víchová, V., Janda, J. (2017) The Prague Manual, Europe Values Centre
https://www.europeanvalues.net/wp-content/uploads/2018/07/Prague-Manual.pdf

Waltzman, R. (2017) The Weaponization of Information - The Need for Cognitive Security, RAND
https://www.rand.org/pubs/testimonies/CT473.html

Wanless, A., Berk, M. (2018) Participatory Propaganda: the Engagement of Audiences in Spread of Persuasive Communications, Social Media & Social Order, Culture Conflict 2.0, Research Council of Norway
https://lageneralista.com/wp-content/uploads/2018/03/A-Participatory-Propaganda-Model-.pdf

About the authors

James Pamment is Head of the Department of Strategic Communication at Lund University and a senior analyst at the Centre for Asymmetric Threats Studies (CATS) at the Swedish National Defence University. The Lund University team consists of Henrik Twetman, Alicia Fjällhed, Howard Nothhaft, Helena Engelson and Emma Rönngren.

The Government Communication Service (GCS) is the professional body for people working in communication roles across government. Our aim is to deliver world-class communications that support Ministers' priorities, improve people's lives and enable the effective operation of our public service.

gcs.civilservice.gov.uk | @ukgovcomms

Designed by Design102

design102 Find out more at design102.co.uk
Design that makes a difference